

EQUIP-Tanzania Impact Evaluation

Final Baseline Technical Report, Volume I: Results and Discussion



Gunilla Pettersson, Georgina Rawle, Rachel Outhred, Stephanie Brockerhoff, Gabrielle Wills, Dita Nugroho, Paul Jasper, Andrej Kveder and Adrian Beavis

Acknowledgements

We are very grateful to the Ministry of Education and Vocational Training (MoEVT): Minister for Education, Dr Shukuru Jumanne Kawambwa; Permanent Secretary for Education, Professor Sifuni Mchome; Commissioner for Education, Professor Eustella Bhalalusesa; Assistant Director Diversity Unit, Dr Sayi; Director of Primary Education, Ms Sarah Mlaki; and Education Management Information System Manager, Mr Hango, and to the Prime Minister's Office for Regional Administration and Local Government (PMO-RALG): Permanent Secretary of PMO-RALG, Mr Jumanne Sagini; Director of Basic Education, Mr Makali; Assistant Director of Education Coordination, Ms Nkwama and Data Officer, Ms Asnat for their support to the impact evaluation, advice on the impact evaluation design, sharing of data and for providing research permits.

We would like to thank the Department for International Development (DFID), particularly DFID Education Adviser Mr Atfield and the former and current DFID Results Adviser in Tanzania Mr Cockerill and Ms Forgan for useful discussions, advice on the impact evaluation design and feedback on earlier drafts.

The Impact Evaluation Reference Group provided invaluable technical advice on the impact evaluation design at its first meeting and consisted of: Commissioner for Education, MoEVT, Professor Eustella Bhalalusesa; Principal Education Officer, MoEVT, Ms Manori; Associate Dean Academic, School of Education, University of Dar es Salaam, Dr Annate Komba; Lecturer and Coordinator, School of Education, University of Dar es Salaam, Dr Blackson Kanukisya; Former Dean, Faculty of Education, University College of Education, Dar es Salaam, Professor Kitilla Mkumbo; DFID Education Adviser, Mr Atfield; DFID Results Adviser, Mr Cockerill; Monitoring and Evaluation Specialist, EQUIP-T Managing Agent, Ms Gordon; Manager Uwazi and Experimental Interventions at TWAVEZA, Dr Schipper; and Senior National Education Specialist, OPM Impact Evaluation Team and University of Dar es Salaam, Professor Herme Moshia.

Thanks also go to the Managing Agent Cambridge Education: National Coordinator, Mr Aslett; Deputy National Coordinator, Mr Ogondiek; Deputy National Coordinator – Technical, Dr Tao; Operations Manager, Ms Mroz-Davies; Monitoring & Evaluation Specialist, Ms Gordon; and all of the EQUIP-T technical team leaders, for sharing information on the EQUIP-T programme design and being available for interviews.

For advice on the development of the pupil learning assessment we are grateful to Dr Msonde, Executive Secretary at the National Examinations Council of Tanzania; Ms Hyera, Kiswahili Expert at the Tanzanian Institute for Education; Ms Batchelder, Ms Mulcahy-Dunn and Mr Brombacher, EGRA/EGMA Specialists at the Research Triangle Institute (RTI). We would also like to thank David Bruns, USAID Tanzania, for useful insights and sharing of Big Results Now results. Dr David Johnson, Department of Education, University of Oxford generously shared the teacher development needs assessments (TDNA) developed for Nigeria and for GEP to serve as a basis for the development of the TDNAs for Tanzania for the impact evaluation. Special thanks go to the national experts on the pupil test and TDNA instrument development team: Mr Zombwe, Dr Rugeihyamu and Dr Mrikaria.

We are exceptionally grateful to UWEZO Tanzania, who generously shared their knowledge of instrument development and fieldwork in Tanzania and UWEZO data: Country Coordinator, Ms Mgalla; Manager Uwazi and Experimental Interventions, Dr Schipper, and former UWEZO Programme Officer – Research, Mr Kilala. Our thanks also to Ms Zubeida Masabo, UNICEF, for sharing information on UNICEF's education programmes in Tanzania.

Dr Caine Rolleston, Institute of Education, University of London reviewed the interim and final drafts of this baseline report and provided invaluable feedback and numerous suggestions for improvements.

The OPM team responsible for the design of the impact evaluation and analysis of the baseline data consisted of: Georgina Rawle (Project Manager, Senior Education Specialist and Lead of Instrument Development); Gunilla Pettersson (Education Lead); Paud Murphy (Evaluation Director); Andrej Kveder (Survey Design Lead); Anna Haas (Qualitative Lead Design Stage); Rachel Outhred (Qualitative Lead Implementation and Analysis Stage); Sean O’Leary (Impact Evaluation Specialist); Stephanie Brockerhoff (Qualitative Researcher); and Prof. Herme Moshia (Senior National Education Specialist). The core team was supported by a group of specialist advisers: Patrick Ward (Senior Evaluation Adviser); Ingo W. Outes Leon (Impact Evaluation Adviser), Juan Muñoz (Senior Sampling Adviser), Saltanat Rasulova (Qualitative Researcher), Ian MacAuslan (Quality Assurance), Joshua McGrane (Rasch Modelling Specialist), Gabrielle Wills (Education Researcher), Dita Nugroho (Education Researcher), Paul Jasper (Quantitative Methods), and Adrian Beavis (Senior Education Researcher).

The quantitative baseline fieldwork was conducted by the OPM Tanzania office survey team, which included: Marta Moratti (Survey Manager); Said Aboud (Survey Fieldwork Director); Shafique Arif (Data Management and Fieldwork Specialist); Maria Fibaek (Assistant Survey Manager); Ignatus Jacob (Fieldwork Specialist); Erick Kilala (Data Management Specialist) and Goodluck Lugina (Fieldwork Coordinator). The fieldwork was successfully completed thanks to the hard work of the field enumerators who, under challenging conditions, visited 200 primary schools in 25 districts located in different parts of Tanzania. Special thanks to: Grace Mwambuluma, Lameck Jackson, Everine Swai, Mary D. Mping’wa, Mary Msafiri, Andrew Madembwe, Catherine Sayo, Muhsin Abdallah, Christina Jasper, Noela Lyimo, Mwajuma Ramadhani, Leah Mabeyo, Benedict Mande, Robert Sizya, Ester Kayinga, Jonas Gervas, Magreth S. Kamwela, Godrick Fransis, Juliana Nyange, Stanley Kalokola, Flavia Stanslaus, Joseph Sanga, Hamisi Karisha, Violeth Lemy, Regina Kafanabo, Innocent Tibaikana, Kanisia Kombo, Theresia Robson, Grace Mturo, Naomi Silayo, Josephat Lite, Levina Peter, Doreen W. Mossi, John Wilson, Rehemaely Makotha, Matilda Ndimbo, Sylvester Michael, Samwel Mande, Monica Joseph, Irene Lukumay, Eddna Chandeu, John Lelo, Sakina Mgaya, Matondo Mokili, and Sussie Berya.

The qualitative fieldwork was conducted by a team of national qualitative researchers working with the OPM qualitative research team, and was completed due to their dedication and great efforts. We would like to thank Mary David, Elizabeth Mabella, Rehema Makota, Levina Peters, Ezbon Simon, and Allen Rwegasila.

Needless to say, during the field work regional education officers, district education officers, ward education coordinators, community leaders, head teachers, teachers, pupils and parents played a crucial role by giving their time to allow the impact evaluation baseline research to be conducted and we would like to extend a big thank you to them.

Executive summary

Background

The Education Quality Improvement Programme in Tanzania (EQUIP-T) is a four year, Government of Tanzania programme, funded by the UK Department for International Development (DFID). It targets some of the most educationally disadvantaged regions in Tanzania to increase the quality of primary education and improve pupil learning outcomes, in particular for girls.

This report presents the baseline results from the quantitative survey and qualitative fieldwork conducted for the independent impact evaluation (IE) of the EQUIP-T programme.

The key purposes of the IE study are to: i) generate evidence on the impact of EQUIP-T on primary pupil learning outcomes, including any differential impacts for girls and boys; ii) examine perceptions of effectiveness of different EQUIP-T components; iii) provide evidence on the fiscal affordability of scaling up EQUIP-T post-2018¹; and iv) communicate evidence generated by the IE to policy-makers and key education stakeholders.

The EQUIP-T programme will seek to improve the quality of education, especially for girls, in seven regions in Tanzania and has five components (the IE study covers the first four):

1. Enhanced professional capacity and performance of teachers;
2. Enhanced school leadership and management skills;
3. Strengthened systems that support the district and regional management of education;
4. Strengthened community participation and demand for accountability; and
5. Strengthened learning and dissemination of results.

Together, changes in these five outputs are intended to reduce constraints on pupil learning and thereby contribute to better-quality education (outcome) and ultimately improved pupil learning (impact).

More on the IE purposes and baseline objectives and the EQUIP-T programme

- Objectives of the baseline analysis: Section 1.1 and Annex section A.2.1.6 in Volume I.
- Overview of the EQUIP-T programme: Section 1.2, Annex C and Annex D in Volume I.

Impact evaluation design and methods

The IE takes a mixed methods approach where qualitative and quantitative methods are integrated so as to strengthen the robustness and depth of the research findings. The baseline approach consists of three main components to allow the IE to: i) capture the situation prior to the start of EQUIP-T so that changes can be measured during the follow-up data collection rounds; ii) develop an expanded programme theory of change (TOC) to help inform possible programme adjustments; and iii) provide an assessment of the education situation in some of the most educationally disadvantaged regions in Tanzania to the government and other education stakeholders. This includes:

¹ The IE will also examine whether EQUIP-T scale-up would be affordable given the likely future fiscal context through a separate fiscal study (see Volume II, Annex N).

- Quantitative survey of 100 government primary schools in 17 programme treatment districts and 100 schools in eight control districts in 2014, 2016 and 2018² covering:
 - 2,987 standard three pupils;
 - 681 interviews, 510 teacher development needs assessments (TDNAs) in Kiswahili and 505 TDNAs in mathematics for teachers who teach standards one to three;
 - 569 TDNAs in mathematics for teachers who teach standards four to seven;
 - 200 head teacher interviews and school record checks; and
 - 397 standard two lesson observations in Kiswahili and mathematics;
- Qualitative fieldwork in nine research sites that overlap with a sub-set of the quantitative survey schools, in 2014, 2016 and 2018, consisting of key informant interviews (KIIs) and focus group discussions (FGDs) with head teachers, teachers, pupils, parents, school committee (SC) members, region, district and ward education officials and EQUIP-T programme staff; and
- A mapping of the causal mechanisms, and assessment of the strength of assumptions, underpinning the programme theory of change (TOC) using qualitative and quantitative IE baseline data as well as national and international evidence.

The baseline quantitative survey took place from the end of March to mid-May 2014, followed by the qualitative fieldwork which started in June 2014 and was completed by mid-August.

Selecting the samples

The quantitative survey uses a quasi-experimental design with multi-stage sampling, while the qualitative research uses a small purposive sample to collect the baseline data. The quantitative school sample was drawn from the group of programme districts in which no programme similar to EQUIP-T is in operation and that will not receive partial EQUIP-T interventions. Control districts were matched to the treatment districts using a propensity score matching (PSM) technique. In the next stage, a random sample of 100 government primary schools was drawn from the programme treatment districts, and 100 schools from the control districts were matched to selected treatment schools using PSM. At the final stage, random samples of pupils and teachers were drawn in the 200 sample schools. To select the qualitative research sites, in the first stage, three districts were selected from the programme treatment districts using purposive 'typical case sampling' (sampling an average district in terms of education outcomes) and 'extreme case sampling' (sampling high and low performing districts with respect to education inputs, outputs and outcomes). In the second stage, three schools within each of the three selected programme districts were sampled also using 'typical case sampling' and 'extreme case sampling'.

Generalisability and representativeness of the baseline results

The quantitative baseline results presented in Volume I of this report are representative of government primary schools, and of standard three pupils and teachers in these schools, in the 17 programme treatment districts in the five regions where the EQUIP-T programme started implementation in July 2014. The 17 programme treatment districts covered by the IE appear to be fairly similar to the remaining 12 programme districts in terms of key contextual characteristics, but the programme treatment districts as a whole are on average significantly different from other districts in Tanzania. This means that the quantitative results in this report are not generalizable

² Computer-Assisted Personal Interviewing (CAPI) was used to collect the data except for the teacher development needs assessment (TDNA), which was administered on paper as this takes the form of mock pupil tests that teachers mark

beyond the programme treatment districts, although they may provide indications in cases of districts highly similar to the programme treatment districts.

The baseline analysis of the programme TOC, also presented in Volume I of this report, draws on both the quantitative and qualitative baseline results, and is the first step in the theory-based aspect of the evaluation. Following future data collection rounds, which seek to understand the reasons and context in which components of EQUIP-T lead to impacts, this will allow some speculation about the likely impact of EQUIP-T if it were to be scaled up to other parts of the country.

Risk to the impact evaluation achieving its objectives

As for any IE, there are a range of risks to achieving its purpose, as well as a related set of strategies to try to mitigate these. Some risks were identified at the design stage, for example it is possible that the time-period of the evaluation (four years) may not be sufficient to detect the impact of the programme on pupil learning if change happens very slowly. More recently, a potentially serious risk has emerged of contamination of the control districts by the Literacy and Numeracy Education Support Programme (LANES). LANES has similar aims to EQUIP-T, will be implemented over a similar period, and plans to train a selected number of teachers from all primary schools which are not already covered by an early grade support programme (so this will exclude EQUIP-T programme districts). The first stage mitigation of this risk is to seek an optimal sequencing of the LANES programme, through discussion with the Government and the LANES supervising entity, as well requesting regular updates on its implementation. Prior to the design of the mid-line research, this risk will be reviewed as it may be necessary to substantially alter the design of the IE in order to meet its objectives.

More on IE design and methods

- Overall IE and mixed methods approach: Section 1.3.1 in Volume I and Section 2 in Volume II.
- Survey instruments and research tools: Section 1.3.2 in Volume I, Sections 3.3, 4.5 and Annex J in Volume II.
- Sampling strategies and sample sizes for the IE baseline: Section 1.3.3 in Volume I and Sections 3.2 and 4.4 in Volume II.
- Generalisability and representativeness of baseline results: Section 1.3.4 in Volume I.
- Risk to the impact evaluation achieving its objectives: Section 3.7 and 4.8 in Volume II.

The baseline results for programme treatment districts summarised next are typical of the 17 disadvantaged districts in which the EQUIP-T programme is operating.

Results: Pupil learning in programme treatment districts

Pupil learning levels in the programme treatment districts are very low overall, and lower than the national average

At the baseline, pupils in the programme treatment districts on average perform much worse than the national average on the primary school leaving examination (PSLE). Specifically:

- Close to half of schools in programme treatment districts did not even manage a 25% PSLE pass rate, while only one-fifth of schools nationally are in this situation.
- There is also a huge gap at the opposite end of the performance spectrum, with merely 5% of schools in the programme treatment districts achieving a pass rate of 75% or above, compared to 20% of primary schools nationally.

The majority of standard three pupils have fallen considerably behind curriculum expectations in both Kiswahili and mathematics

The IE baseline survey tested standard three pupils in Kiswahili and mathematics at standard one and standard two national curriculum levels, using an oral test that was administered on a one-to-one basis³.

- Only 12% of pupils in standard three are currently achieving the expected standard two level skills in core areas of the Kiswahili curriculum and 6% of pupils for mathematics.
- About 24% of pupils have acquired ‘emerging’ standard two level skills in Kiswahili and 30% in mathematics.
- 26% of pupils are achieving at standard one or emerging standard one level in Kiswahili and 57% of pupils in mathematics.
- Most alarming is that 38% of pupils have not yet acquired even emerging standard one level skills in Kiswahili. A substantially smaller but also important group of pupils (7%) have not yet achieved even emerging skills at standard one level in mathematics.

Pupils who are poorer perform significantly worse in Kiswahili

- Pupils from poorer households are significantly more likely to have acquired less than standard one skills in Kiswahili and are also less likely to have achieved standard two skills than their richer peers.
- However, there are no significant differences in the proportions of poorer and richer pupils achieving at the different curriculum levels in mathematics.

Pupils who do not speak the language of instruction (Kiswahili) at home perform significantly worse in both Kiswahili and mathematics

- For Kiswahili, almost 20% of pupils who speak Kiswahili at home are achieving at standard two level compared to only 10% of pupils who speak other local languages at home. At the other end of the performance distribution, 42% of pupils from non-Kiswahili-speaking homes achieve below standard one level in Kiswahili compared to ‘only’ 26% of pupils from Kiswahili-speaking homes.
- For mathematics, almost 11% of pupils who speak Kiswahili at home are achieving at standard two level compared to just 5% of pupils who speak other local languages at home. However, for mathematics the proportions of pupils that achieve below standard one level are notably smaller for both groups of pupils: 9% of pupils who speak other local languages at home achieve below standard one level compared to 2% of pupils from Kiswahili-speaking homes.

Girls perform significantly worse than boys in mathematics, but not in Kiswahili

Comparing the distribution of Kiswahili performance for boys and girls reveals no significant performance differences by gender. However, boys tend to perform better than girls in mathematics, with the proportion of boys who are achieving at standard two level in mathematics being twice as high as the proportion of girls.

³ The IE pupil tests were adapted from an early grade reading assessment (3Rs-EGRA) and an early grade mathematics assessment (3Rs-EGMA) which were recently used to collect baseline data for the Government’s Big Results Now (BRN) Education programme (RTI 2014).

Are stakeholder perceptions of learning consistent with actual pupil performance?

The qualitative research explored perceptions of different education stakeholders on the comparative academic performance of boys and girls. Head teachers, teachers and parents often perceived that girls perform better than boys. They based their judgements on the positive behaviours displayed by girls, including not being involved in negative peer groups and sitting at the front of the classroom.

These perceptions are not consistent with the IE baseline results presented above that show boys and girls performing similarly in Kiswahili and boys tending to perform significantly better than girls in mathematics. This underlines the importance of schools monitoring pupil performance (see below) and, once they do, to share pupils' academic progress with parents.

More on pupil learning

- Pupil learning Kiswahili: Section 2.1 in Volume I.
- Pupil learning in mathematics: Section 2.2 in Volume I.
- Perceptions of pupil learning: Section 2.3 in Volume I.

Results: Characteristics of pupils, head teachers and teachers and schools in programme treatment districts

Pupil poverty status, language spoken at home, gender and age

- One-third of pupils belong to a household predicted to fall below the national poverty line.
- The large majority of pupils (77%) speak a local language other than Kiswahili at home.
- 52% of standard three pupils are girls.
- The mean pupil age is 10 years (the appropriate age for attending standard three is 9 to 10 years). However, 6.5% of pupils were under-age and 37% were over-age which could be due to children not starting school at the correct school-entry age, repetition or both.

Differences in a typical day for boys and girls in case study communities

The qualitative research explored a typical day for a boy and for a girl. Across the nine case studies, girls generally had more responsibilities for chores in the home and in most of the case study schools boys mentioned having time to play after school whereas girls did not. However, girls in all three districts talked about reading and studying at home but boys only did so in two of the three districts.

Head teacher and teacher gender, academic and professional qualifications and experience

- More than half of teachers (55%) are female but only 16% of head teachers are female. This suggests that there may be some gender bias in favour of men in the promotion of teachers into head posts since female and male teachers are similarly qualified and experienced.
- The vast majority of teachers (76%) have completed Form 4 and 10% have completed Form 6 or higher. However, there is a group of teachers (14%) who have only completed primary schooling, the same level of schooling they are teaching. Among head teachers, 72% have a Form 4 qualification and 28% have Form 6 or higher.
- Nearly all teachers (99.5%) and all head teachers hold a professional education qualification. Most common by far is to have a Certificate in Education.

- The average teacher has spent 16 years working as a teacher and eight years working in his/her current school; however, there are some teachers with very long tenure and experience, so these averages mask low levels of experience and tenure among many teachers. Close to one-fifth (19%) of teachers have been at their current school for less than one year, and a similar proportion are due to retire before the end of EQUIP-T.
- Almost two-fifths (38%) of head teachers have been at their current school for under one year. The reason for this exceptionally high turnover is not clear, but some respondents suggested that this is the result of a policy to post more capable head teachers to improve certain schools.

Characteristics of the average school

The primary schools in the programme treatment districts are on average far from meeting the government's recommended basic standards for primary education (MoEVT 2009a), reflecting the fact that the programme is targeting some of the most educationally disadvantaged regions in Tanzania.

- The average pupil–teacher ratio (PTR) was 54:1 compared to the recommended benchmark of 40:1. However, there was substantial variation across schools, with PTRs ranging from 31:1 to 83:1.
- On average, there were 78 pupils per classroom in use compared to the recommended national benchmark of 45 pupils per classroom.
- 48% of schools have at least one class coming in at a second shift, that is, they cater to separate groups of pupils during the school day. From the IE baseline data it appears that the key constraint is the lack of classrooms although subject specialisation means that in some cases there may also be insufficient numbers of teachers.
- The vast majority of schools (96%) had a functional toilet, but there was on average 74 pupils per functional toilet. This falls short of the government's recommended standard of at least one toilet per 20 girls and one toilet per 25 boys.
- Furthermore, only 32% of schools had drinking water available on the day of the survey compared to the recommended standard of at least five litres per pupil per day for each school.

More on pupil, head teacher, teacher and school characteristics

- Pupil background characteristics and daily routine: Section 3.1 in Volume I.
- Head teacher and teacher characteristics: Section 3.3 in Volume I.
- Characteristics of schools: Section 3.2 in Volume I.

Results: School leadership and management in programme treatment districts

One of the key outputs of the EQUIP-T programme is enhanced school leadership and management (SLM). The IE baseline results on SLM in the programme treatment districts are presented before the other programme outputs because effective SLM is a critical foundation for improving school performance.

Head teachers' understanding and implementation of their role and responsibilities is weak⁴

Within the case study schools, head teachers understanding and reported implementation of their leadership role and responsibilities (including: Shaping a vision of academic success for all pupils; creating a climate hospitable to education; cultivating leadership in others; improving instruction; and managing people, data and processes) were generally poor.

The IE baseline survey found that just over one in ten (11%) of the head teachers had attended any in-service training (INSET) on SLM in the last two school years (2012 and 2013).

Head teacher absenteeism is common

A necessary condition for improving the efficacy with which head teachers lead and manage schools is that they are present in school. But head teacher absenteeism is very high. The head count on the day of the survey indicates that 16% of head were absent, and school records show that 64% were absent on one or more of the previous five days. Some of these absences may be legitimate, for instance, head teachers sometimes need to attend meetings at the district level.

Whole school development plans (WSDPs) are rarely available and typically incomplete, but most schools appear to manage capitation grants

- Despite the importance placed on WSDPs, only 37% of head teachers reported that the school had a WSDP, but only 21% of head teachers were able to show the physical document.
- Where WSDPs were available, their comprehensiveness was limited. Specifically, three elements: A budget, teaching and learning objectives, and baseline data and targets are considered important features, but only 2% of the schools had WSDPs that included all these three elements.
- In contrast to the relatively small proportion of schools that had WSDPs, especially comprehensive ones, over 83% of schools were identified by the enumerators as having complete records on per capita grants in 2012 and in 2013.

Teacher performance management is typically weak

- Head teachers consider checking lesson plans as the most important means of assessing teachers (37%) followed by observation of teaching performance in class (24%) and using pupil academic results (18%). Only 4% of head teacher cited punctuality and attendance, despite there being extremely high levels of teacher absence from classrooms in particular.
- Nearly three-quarters of teachers reported that they had not had a one-on-one meeting with the head teacher, assistant head teacher or academic master to discuss their performance or professional development needs in the previous school year (2013).
- Moreover, the use rewards, financial or non-financial, for good teacher performance was only reported by 35% of head teachers.
- Some head teachers in the qualitative study reported that they do not feel empowered to deal with teacher disciplinary matters, and view ward officials as having authority on this.

More on SLM

- Head teachers' roles and responsibilities: Section 4.1 in Volume I.

⁴ The qualitative study explored five areas of school leadership responsibility that are mapped against the four competencies in the National School Leadership Competencies Framework (NSLCF) developed under the EQUIP-T programme.

- Head teacher absenteeism: Section 4.2 in Volume I.
- Head teacher INSET and community support: Section 4.3 in Volume I.
- Whole school development planning: Section 4.4 in Volume I.
- Teacher performance management: Section 4.5 in Volume I.

Results: Teacher capacity, performance, motivation and morale in programme treatment districts

The programme aims to improve teacher capacity, performance, motivation and morale through the development of a Teacher Competency Framework (TCF), school-based INSET, the development of a teacher performance management system (TPMS), and through a Teacher Morale Toolkit.

Teacher subject knowledge at the lower levels of the Kiswahili and mathematics curricula is relatively strong

- Teachers scored 58% on average in the Kiswahili TDNA.
- Teachers performed considerably better on the questions drawn from the standard one to four level questions in the Kiswahili curriculum compared with those from standards five to seven.
- The contents of the mathematics TDNA focus on questions from the upper standards of the primary curriculum because pre-test results showed that the vast majority of teachers found the questions from the lower levels of the curriculum very easy.
- Teachers scored 59% on average in the mathematics TDNA, but this masks a wide range of teacher performance.
- Teachers who teach mathematics to upper-primary school pupils demonstrated much stronger subject knowledge than their colleagues who teach the lower primary level.

The qualitative research found that teachers felt confident in their knowledge of the subject in which they specialised, but many teachers say they are now teaching subject areas in which they have not specialised. In the case study schools, the biggest gap in 'overall' teacher knowledge, which comprises curriculum, subject and pedagogic knowledge, was reported to be curriculum knowledge. Due to the lack of teacher training, teachers were unaware of the content of the new syllabus, did not know what subject matter should be taught to which grade level pupil, and were unclear how various textbooks relate to the curriculum they are required to teach.

Most teachers show gender balance in classroom interactions with pupils but interact more with pupils at the front of the classroom than in other parts

- Just over half (54%) of the observed teacher interactions with pupils were gender balanced, while 30% of teachers interacted more with boys than girls and 16% of teachers interacted more with girls than with boys. This suggests that gender balance in teaching practice is not an acute problem, but clearly there is scope for improvement. There are more classrooms where boys receive more attention from the teacher than girls than vice-versa.
- Teacher interactions were most common with pupils seated at the front of the classroom (42% of all interactions) compared to the middle (30%) and back (28%) areas. This suggests that teachers interact less with pupils seated towards the middle and back of the classroom.

Only a small group of teachers demonstrate use of effective teaching skills in the classroom

According to the qualitative study, most teachers were unable to identify pedagogy as a knowledge required of a teacher and tended to focus on subject and curriculum knowledge during the FGDs. Moreover, only a few teachers displayed elementary knowledge of pedagogy and discussed child-centred teaching and participatory methods. This finding is consistent with the IE quantitative baseline results on pedagogy based on lesson observations:

During the introductory stage of the observed lessons:

- 23% of teachers clearly stated learning objectives and 22% checked knowledge of prior work covered, and only a very small proportion (7%) of teachers specifically stated what new skills or knowledge the pupils should have acquired by the end of the lesson.
- In the middle stage of the observed lessons:
- The vast majority of observed teachers (85%) were found to relate well to pupils, including conveying enthusiasm, using encouragement and praise to give positive feedback.
- Only 4% of teachers frequently encouraged individual pupils to ask questions and explain ideas, whereas 79% of teachers never did this.
- 11% of teachers frequently asked pupils open-ended questions, but 66% never did.
- 12% of teachers frequently commented on or probed pupil answers frequently, but a large group of teachers (51%) never displayed this behaviour.
- Merely 6% of teachers frequently asked pupils to carry out activities in pairs or in groups whereas the vast majority (77%) of teachers did not do this.

At the concluding stage of the observed lessons:

- 22% of teachers checked if pupils had acquired the new skills or knowledge set out in the lesson introduction and 21% of teachers drew the whole class together to summarise what materials/topics had been covered and directed pupils to the next stage of the topic.

Among the small group of teachers (9%) who 'frequently demonstrated' seven or more effective teaching behaviours, more than two-thirds also displayed gender and spatial balance in their interactions with pupils. This suggests that there is a small group of teachers who use both effective and gender-balanced teaching behaviours, but also a very large group of teachers who do neither.

- When it comes to assessment of pupil progress, slightly more than one-quarter (27%) of teachers were able to show that they carried out at least two different types of pupil assessment the previous five school days.
- At the same time, 31% of teachers were unable to show that they had carried out any pupil assessment over the same period.

The IE baseline survey also recorded if teachers switch between Kiswahili (language of instruction) and a local language other than Kiswahili during teaching.

- Language switching was only observed in 4% of the observed lessons and in the interviews a similar proportion of teachers reported that they teach in a language other than Kiswahili or switch languages during lessons.

Teacher absenteeism from school and classrooms is a major problem with adverse consequences for instructional time

The EQUIP-T programme will not only seek to improve teachers' subject knowledge and teaching behaviours but also to improve teacher motivation and morale so as to increase teacher attendance, level of effort, commitment to the job, and time on task.

From the qualitative research, the regional and district education officials indicated that low levels of societal respect for the teaching profession combined with the low qualifications to join the profession have resulted in a large number of teachers in the workforce who are either unlikely to respond to efforts to increase teacher motivation and morale or who are only likely to respond to tangible and extrinsic incentive mechanisms.

- More than one in 10 (12%) teachers were absent from school on the day of the survey. Other sources for Tanzania report even higher teacher absence rates (World Bank 2012).
- Further, of those teachers who were present on the day of the survey, almost two-thirds (63%) arrived late.
- Most detrimentally of all to pupil learning, classroom absence is extremely high. Among teachers present at school and timetabled to teach 66.8% were absent from their lesson.
- To enable comparison with another study for Tanzania, the 2010 Service Delivery Indicator Survey (SDIS), a second, comparable IE baseline measure of classroom absenteeism was computed and using this classroom absenteeism measured by the IE was 67% compared to 50% for rural schools in the SDIS (68% for urban schools) (World Bank 2012).

While schools may be open for instruction and instructional time formally timetabled, the actual amount of instructional time received by pupils ultimately depends on whether teachers are in the classroom.

- If an adjustment is made to account for classroom absenteeism, pupils only receive 77 minutes of mathematics instruction per week, compared to the mean 219 minutes timetabled, and 78 minutes of Kiswahili instruction per week compared to the mean 219 minutes timetabled.

Such large losses of instructional time will have large, adverse consequences for pupil learning.

From the qualitative study, teacher absenteeism and lateness was most often attributed to a lack of teacher housing near the school. The second most commonly reason cited as reducing teacher motivation and morale and contributing to absenteeism were low salaries and late salary payments with teachers across all the case study schools engaging in other income-generating activities (IGAs).

More on teacher capacity, performance, motivation and morale

- Teacher subject knowledge: Section 5.1 in Volume I.
- Teacher pedagogy: Section 5.2 in Volume I.
- Teacher motivation and morale: Section 5.3 in Volume I.

Results: Community participation and accountability in education in programme treatment districts

Component 4 of the EQUIP-T programme aims to increase community participation and accountability in education by providing resources, training and information.

Communication mechanisms between schools and communities are often not working well

- At the baseline, only 49% of schools had a noticeboard and of the existing noticeboards few communicated relevant education information. Only 8.1% displayed information on the WSDP, the school budget or school capitation grants, 21% displayed academic results, 17% displayed pupil or teacher attendance information, and 16% displayed information on events.

The qualitative research found that the main form of communication between schools and parents is through letters given to pupils to take home to their parents, although this is often unreliable. Other mechanisms that were frequently mentioned were village meetings and school meetings.

- In the vast majority of surveyed schools, whole-school parent-teacher meetings are taking place. Two-thirds of head teachers were able to produce meetings from a previous meeting.
- Only 14% of schools reporting having some type of formal parent-teacher group in place.
- A quarter of all surveyed teachers said that they did not report pupil progress to parents in 2013. Of those teachers who said they did, 9% reported doing so only once and 88% twice during the year.

The qualitative research found that monitoring the academic progress of children was largely perceived as being the responsibility of parents. Effective communication was widely viewed as a key factor leading to a good relationship between the school and community. However, views on who should ensure that communication takes place and that this is done in an effective way differed across the various stakeholders.

SCs generally understand their role and responsibilities, but capacity to effectively enact these is currently weak

- All schools in the programme treatment districts had a SC and 91% of the schools could show minutes from a SC meeting.

The qualitative research found that SCs have a very broad understanding of their roles and responsibilities including dealing with pupil and teacher absenteeism, maintaining infrastructure, ensuring proper use of financial resources and persuading parents to contribute financially among other things. Most SCs did not mention representing parents as a core responsibility.

Typically, SCs described their role as a problem solver mediating between teachers, parents and pupils and engaging with the village council. SCs are highly aware that their power is limited and that they can only call meetings with parents. In order to call meetings with the community, they have to liaise with the village council, which then initiates these meetings. The relationship between the SC and the village council is thus key to the ability of the SC to fulfil its role. SCs also see the creation of a good relationship between teachers and parents as one of their core functions.

SCs generally felt that they were doing a good job fulfilling their responsibilities, even in cases where parents reported being unhappy with the performance of the SC. SCs themselves reported facing a number of problems in conducting their job, including a lack of support from the village council and district. Other problems mentioned were a lack of funds for transport to attend meetings at the district and ward level, difficulties when talking to parents who are unaware of the importance of education, a lack of support from parents for ideas for IGAs, non-responsiveness when asking for financial contributions from parents, late payments by parents, a general lack of funds and lack of attendance at meetings. Some teachers and parents perceived the SC as being at a disadvantage when dealing with education matters in general, and in particular with questions surrounding schools' financial resources.

Many school leaders do not feel supported by the SC and community

- Around half of head teachers (55%) rated SC support to their school as either good or very good, whereas only 20% rated wider community support equally positively.

Across all nine case study schools, parents and village chairpersons reported providing some form of support to schools, either financial or in-kind. However, parents felt that they did not always know what happened with their contributions, which in some cases led to an unwillingness to contribute further. Notably, head teachers, teachers and at times also SC members seemed to only consider financial contributions as proper community support to schools.

- On the day of the school survey, a head count found that one-third (33%) of pupils in standards one to three were absent.

The extent of pupil absence can serve as an indirect indicator of community support for schools, as high rates of absence in the early grades may indicate a lack of parental valuation of education, although there may be other reasons too. However, in the case study schools where the relationship between the school and the community was described as good, pupil absenteeism was reported to be lower. Teachers cited a lack of awareness of the importance of education as a reason for low pupil attendance, while parents felt that this was only true for certain parts of the community. Teachers generally blamed parents for pupil absenteeism and said it was parents' responsibility to ensure that children attend school. While parents generally agreed that it was their responsibility too, they also reported having to rely on information from the school about the non-attendance of their children.

Perceptions of the ability of parents to hold head teachers and teachers to account for education delivery vary across stakeholders

Teachers and head teachers across all case study schools felt that parents frequently came to school to complain if they were unhappy with the state of schooling. On the other hand, parents in several case study schools discussed the repercussions of trying to hold teachers and head teachers to account. Parents felt that teachers would indirectly punish the children of parents who demanded better education, branding them as trouble makers, and would no longer provide them with the support they needed. These parents were therefore sceptical of their ability to hold schools to account.

The qualitative study also highlighted multiple sources of perceived conflicts between communities and schools over child work at school, corporal punishment, sexual harassment of girls by teachers, and the use of community contributions and revenue from IGAs. This strongly suggests that accountability relationships are not working well in many cases.

More on community participation and accountability in education

- Communication mechanisms between schools and communities are often not working well: Section 6.1 in Volume I.
- SCs understand their role and responsibilities, but the capacity to effectively enact these is currently weak: Section 6.2 in Volume I.
- Many school leaders do not feel supported by the SC and community: Section 6.3.1 in Volume I.
- Perceptions of the ability of parents to hold head teachers and teachers to account for education delivery vary across stakeholders: Section 6.3.3 in Volume I.

Results: District and regional education management in programme treatment districts

The objective of Component 3 of the EQUIP-T programme is to strengthen systems and human resource capacity to support sub-national management of education. The programme is designed to do this by providing training and mentoring to district and regional officials on education planning, budgeting, Education Management Information System (EMIS) analysis, public financial management, and monitoring and evaluation (M&E).

Regional, district and ward-level education officials consider insufficient resources the most important constraint to fulfilling their responsibilities

Overwhelmingly, regional education officers (REOs), district education officers (DEOs) and ward education coordinators (WECs)⁵ cited insufficient resources as the biggest constraint to fulfilling their responsibilities. The resources REOs, DEOs and WECs perceive as being required to do their jobs include funds to support transportation and teaching and learning materials. For WECs, office space, housing, desks and stationery were also cited as necessary requirements to fulfil their responsibilities. However, school level and community stakeholders tended to provide less tangible political motivations as explanations for poor support for education at the regional, district and ward levels. Such explanations included education officials seeking allowances, attending their own training and events and maintaining the hierarchy, rather than focusing on education service delivery.

The flow of funds from central to district to school level is inconsistent and unpredictable

- Officially schools are supposed to receive TZS 10,000 per primary pupil in the form of capitation grants annually. However, on average, schools only received around TZS 3,800 per pupil in 2013⁶. This finding is consistent with other evidence from Tanzania (DataVision 2013). From the qualitative research, regional, district and ward-level officials did not consider the failure to release capitation grants in full and on time as a capacity problem, but rather as a matter of diversion of education funds to other areas, lack of accountability, and vested interests.

District offices provide support to schools through district school inspectors and ward education coordinators

The qualitative study found that while WECs perceived their roles to include providing academic support to teachers and support to school leaders, the majority of teachers and head teachers felt that lack of academic and leadership support was a primary factor in declining levels of education service delivery. By contrast, the main responsibilities cited by school and community stakeholders regarding WEC support for school improvement was disciplinary actions against absent teachers, engagement with the community regarding pupil absences, and dealing with community reports of difficulties at the school.

- Almost all schools (99%) were visited by WECs in 2013, and just under one-third (35%) of teachers were visited individually by WECs who observed their teaching or held a meeting

⁵ Since the baseline research took place, WECs have been retitled as Ward Education Coordinators.

⁶ It should be noted that these estimates are indicative only as records on capitation grant amounts, and in particular date received, were in many cases incomplete.

with them. Schools that were visited by the WEC were visited an average of 6.6 times in 2013, but the frequency of visits varied considerably between schools.

As part of the IE baseline survey, head teachers were also asked how they rate the support of the WEC to their school, from very poor to very good. Overall, 80% of head teachers in the programme treatment districts rated WEC support as either good or very good. However, in case study schools where the WEC visited regularly, there was no evidence that the head teacher and teachers necessarily functioned more effectively in their roles.

- District school inspectors (DSIs) visited schools less frequently, with two-thirds of schools (63%) and less than one-quarter of teachers (24%) being visited individually by DSIs in 2013. Visits by the DSI averaged 1.5 times for the schools that were visited.

In the qualitative study, the different stakeholders identified weaknesses across all stages of the education planning cycle, including very low levels of data quality in the education sector, incomplete and late disbursement of funds, misuse of resources, and a lack of personnel and physical resources. By contrast, strengths identified by education officials tended to focus on specific events or people, rather than systematic strengths.

More on district and regional education management

- Constraints to regional, district and ward-level education officials fulfilling their responsibilities: Sections 7.1 and 7.2 in Volume I.
- The flow of funds from central to district to school level: Section 7.3 in Volume I.
- District offices' support to schools: Section 7.4 in Volume I.

Conclusions based on the impact evaluation baseline results and the expanded programme theory of change

In addition to presenting baseline results for areas in which the programme seeks to affect change, the IE baseline report expands the programme TOC by mapping out the programme interventions and expected changes, how they interlink and the assumptions underpinning these links. A combination of the IE baseline results as well as Tanzanian and international literature was then used to assess the assumptions, the relevance of programme components, and to identify possible areas for programme adjustment.

The expanded programme TOC is shown in Figure 1. The blue hexagons indicate programme interventions and the white hexagons changes expected to stem from these interventions. The relative strengths of the explicit and implicit assumptions underpinning the links in the TOC are shown as arrows, where green arrows indicate assumptions that are likely to hold based on available evidence, orange arrows mean there is some evidence that the assumptions are likely to hold, but they may only do so under specific conditions, and red arrows indicate weak assumptions for which there is no or only little existing evidence.

Programme relevance and the relative strengths of the assumptions underpinning the expanded programme TOC are summarised below for each programme component.

Pupil learning

Programme relevance. The IE baseline results on low pupil learning levels confirm that the core objective of the programme, 'to provide better-quality education, especially for girls', is highly relevant, and that large-scale change will be needed. It may be important to ensure that pupils who do not speak Kiswahili at home are visible as a target group as well as girls.

Strength of programme assumptions and risks. While most of the links between the planned programme interventions and the four changes directly linked to improving school quality in the expanded TOC, appear to be supported by existing evidence, the IE baseline study finds two important potential weaknesses. First, on the school-readiness intervention, while there is strong evidence that pre-school classes can increase school-readiness on a wide range of important measures, these do not include sustainable gains in language acquisition, which is clearly critical in this context. Second, on improving the learning environment for girls, the assumptions required for the Form 4 scholarship programme to lead to more female teachers into rural schools are numerous, and at best only likely to hold in certain contexts. It also seems likely that any change stemming from this intervention is likely to occur outside the programme life cycle.

School leadership and management

Programme relevance. The core aim of the second programme component is to ensure that head teachers lead schools effectively, so that schools meet quality standards, including those set for teacher performance. The evidence supporting the link in the TOC between head teachers leading schools effectively and improving teacher performance (which in turn is linked to improving education quality) appears fairly strong in principle. The IE baseline study finds that head teachers' capacity to lead and manage schools is generally weak across a range of leadership competency areas including teacher management, and that head teachers were not typically aware of many basic school leadership responsibilities. Given this baseline context, the areas of programme intervention seem highly relevant.

Strength of programme assumptions and risks. The links between the planned programme interventions⁷ under this component and improving school leadership were found to be supported in principle by existing evidence, with the exception of leadership professional development where the extent to which this is likely to hold is mixed/contextual, partly depending on the scope and quality of training, as well as the readiness of head teachers to receive training and then apply new skills. Another factor that could potentially undermine the effectiveness of head teacher training is the high level of head teacher turnover.

Teacher performance

Programme relevance. Improving teacher performance is the core aim of the first component of the EQUIP-T programme. This is a highly relevant objective given the potentially strong link between teacher performance and improving school quality, and because the IE baseline findings show low levels of teacher performance in critical areas including pedagogy and also low levels of teacher motivation and morale associated with high levels of absenteeism, which suggests that the areas of programme intervention in this component are highly relevant.

Strength of programme assumptions and risks. While the link between professional development training and improving teacher performance is supported by the literature⁸, the success of the teacher morale and motivation intervention will depend heavily on intervention design and context. It seems questionable at this point whether the teacher morale toolkit and implementation of a PMS will be able to address many of the systemic inhibiting factors identified by the qualitative research including lack of teacher housing and low salaries. Given the clear need for behaviour change among the majority of teachers, this intervention is critical to programme impact on pupil

⁷ Development of PMS, HTs develop WSDPs, School quality framework developed, HT trained in leadership (including gender)

⁸ Assuming various design and contextual conditions.

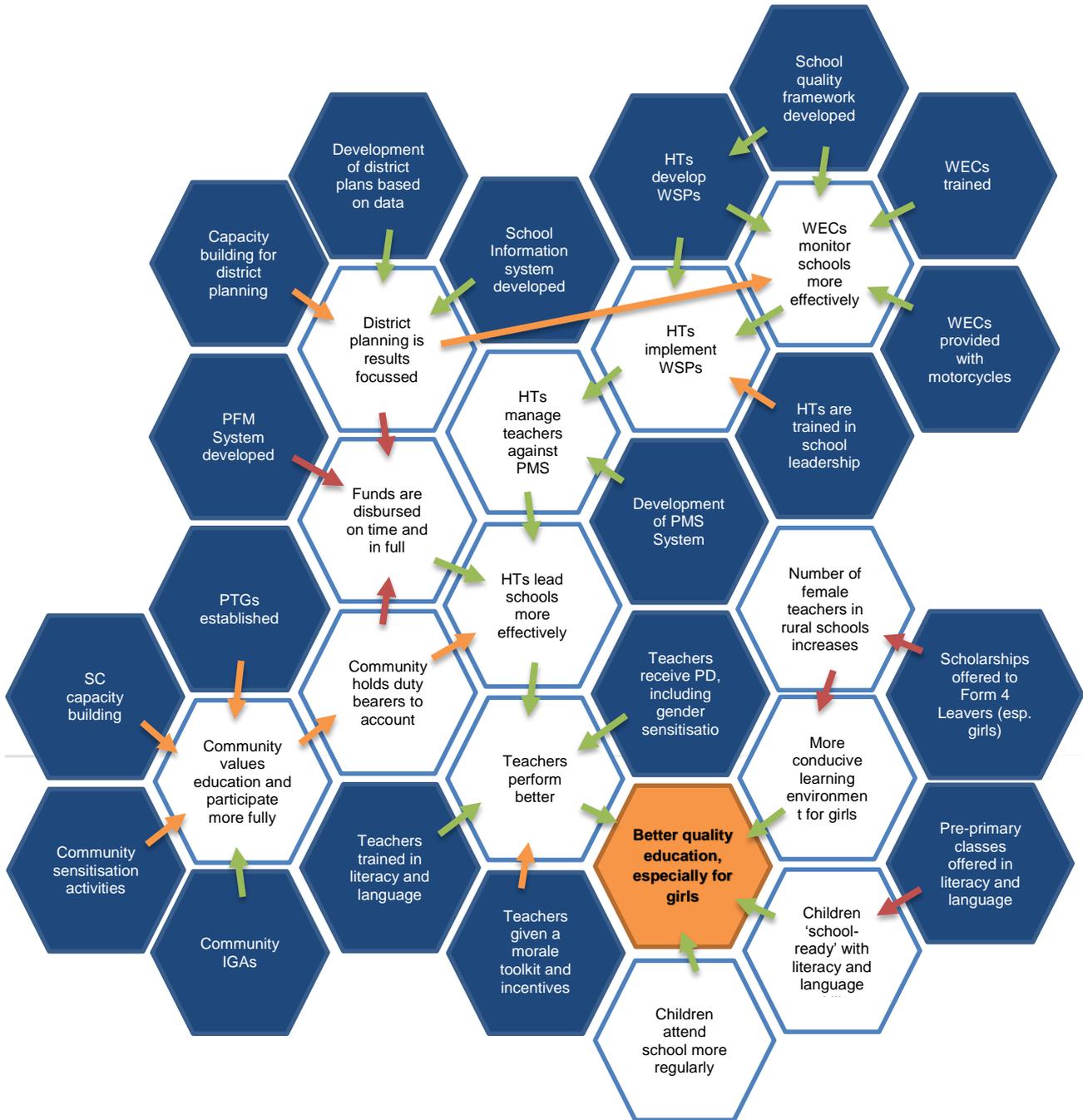
learning, and seems at risk. Another factor that could undermine the effectiveness of this component, especially in relation to the professional development intervention, is the high level of teacher turnover found at baseline.

Community participation and accountability

Programme relevance. The objective of the fourth programme component is to inform, equip and empower communities to engage more effectively in school operations and outcomes. This component targets two key changes: Communities valuing education and participating more fully in schools; and communities holding duty bearers (including the head teacher) to account for school performance. In the TOC, the link between holding duty bearers to account and improving head teacher leadership is the main pathway to improving the quality of education for this component, and is supported by mixed/contextual evidence. The IE baseline results indicate that communities are generally not well informed, equipped or empowered currently. Taken together, this affirms the relevance of this component's objectives.

Strength of programme assumptions and risks. All of the links between the planned interventions and communities valuing education and participating more fully, are supported by mixed/contextual evidence, except the IGA intervention, where there is evidence of a stronger link. Community-led IGAs have been instigated in many of the case study communities already, and it was a common view that the capital invested, when available, ultimately generated funds which were used for education purposes. The broader research highlighted the risk of a lack of sustainability of IGAs after the initial investment, and also the potential for disruption and conflict due to shifting power dynamics linked to resources.

Figure 1 Complete, expanded EQUIP-T programme theory of change



District and regional education management

Programme relevance. The core objective of the third programme component is to ensure that regions and districts have the capacity to effectively manage the education system. Three key changes are sought: For WECs to monitor schools more effectively; for funds to be disbursed to schools on time and in full; and for district education planning and implementation to be results-based. In the TOC the links between the first two changes and improving head teacher leadership are the main pathways to improving the quality of education for this component and are supported by fairly strong evidence. Given this, and the weaknesses in district and regional management observed at baseline both through the qualitative and quantitative findings, these changes would appear to be relevant programme objectives.

Strength of programme assumptions and risks. While the majority of links between the planned interventions and the desired changes were found to be supported by existing evidence, this is not the case for the interventions designed to lead to funds are being disbursed on time and in full to schools. There is little contextual evidence in particular that having results-focused district planning, communities holding duty bearers at district-level to account, and improvements in the PFM system, will be sufficient to ensure that funds will flow to schools. More generally, the IE study identifies the risk of political elites and vested interests disrupting or blocking certain elements of the district support programme.

More on the expanded programme TOC and conclusions from the IE baseline study

- Assessing and expanding the programme TOC: Chapter 8 in Volume I.
- Conclusions: Chapter 9 in Volume I.

Table of contents

Acknowledgements	ii
Executive summary	iv
List of figures and tables	xxv
List of abbreviations	xxvii
PART A: Background	1
1 Introduction	1
1.1 Objectives of the baseline analysis	1
1.2 Overview of the EQUIP-T programme	2
1.3 Overview of the impact evaluation	3
1.3.1 Overview of impact evaluation design	4
1.3.2 Instruments used for the impact evaluation baseline	5
1.3.3 Sampling strategies and sample sizes for the impact evaluation baseline	6
1.3.4 Generalisability of baseline results	8
1.4 Pointers on the interpretation of baseline results	9
1.4.1 Weighted estimates for programme treatment districts	9
1.4.2 How to read the figures and tables in Part B	9
1.5 Structure of this volume	10
PART B: Baseline results for programme treatment districts	11
2 Pupil learning outcomes	11
2.1 Pupil learning in Kiswahili	13
2.1.1 Are standard three pupils achieving at the expected curriculum level in Kiswahili?	13
2.1.2 Key results from the Kiswahili raw test score analysis	15
2.2 Pupil learning in mathematics	18
2.2.1 Are standard three pupils performing at the expected curriculum level in mathematics?	18
2.2.2 Key results from the mathematics raw test score analysis	19
2.3 Perceptions of pupil learning	20
3 Context: Characteristics of pupils, schools and teachers	21
3.1 Background characteristics and daily routine of pupils	21
3.2 School characteristics	23
3.3 Head teacher and teacher characteristics	25
4 EQUIP-T Component 2: School leadership and management	28
4.1 Head teacher leadership: Roles and responsibilities	28
4.2 Head teacher absenteeism	31
4.3 Head teacher capacity development: in-service teaching training and support networks	32
4.4 School management: Whole school planning and financial management systems	33
4.5 School management and information systems: Teacher performance management	35
4.6 School management and information systems: Protecting instructional time	39
5 EQUIP-T Component 1: Teacher professional capacity, performance, motivation and morale	40
5.1 Teacher professional capacity: subject knowledge	41

5.1.1	Teacher subject knowledge in Kiswahili	41
5.1.2	Teacher subject knowledge in mathematics	42
5.2	Teacher performance: Pedagogy	45
5.2.1	Gender balance in teacher interaction with pupils	45
5.2.2	Spatial balance in teacher interaction with pupils	46
5.2.3	Teacher performance: Teaching behaviours	47
5.2.4	Teacher performance: Assessment of pupil work	50
5.3	Teacher motivation and morale	51
5.3.1	Teacher motivation and morale: Absence and punctuality	51
5.3.2	Teacher motivation and morale: Job satisfaction and valuation by others	54
5.3.3	Factors that influence teacher motivation and morale	54
5.3.4	What schools do to increase teacher motivation and morale	55
6	EQUIP-T Component 4: Community participation and accountability	56
6.1	Informed communities – communication between schools and communities	57
6.1.1	Mechanisms of communication	57
6.1.2	Reporting on academic progress	59
6.1.3	Relationship between schools and communities	60
6.2	Equipped communities – capacity of communities and school committees	61
6.2.1	Understanding of the roles and responsibilities of the school committee	62
6.2.2	Strengths and weaknesses of school committees	63
6.3	Empowered communities – engagement and accountability	64
6.3.1	Supporting schools	64
6.3.2	School attendance	65
6.3.3	Perceived ability of the community/parents to hold school management and teachers to account for the delivery of education	67
7	EQUIP-T Component 3: District and regional management	68
7.1	Current district and regional management of primary education	68
7.2	Physical resources	69
7.3	Transparency and information on key financial and budgetary procedures	69
7.3.1	Capitation grants	70
7.4	Support systems and school improvement systems	71
7.4.1	Ward Education Coordinator and District School Inspector visits	72
7.4.2	In-kind resources received by schools	75
7.5	Key tasks across the policy cycle	76
PART C: Expanding and assessing the EQUIP-T theory of change		78
8	The programme theory of change: relevance, causal mechanisms and underlying assumptions	78
8.1	EQUIP-T Impact: Improving learning outcomes and education quality, especially for girls	79
8.2	EQUIP-T component 1: Improving teacher professional capacity, performance, motivation and morale	82
8.3	EQUIP-T component 2: Improving school leadership and management	93
8.4	EQUIP-T component 3: Strengthening district and regional education management	99
8.5	EQUIP-T component 4: Strengthening community participation and accountability	105
8.6	The complete expanded theory of change	109
PART D: Conclusions and next steps going forward		112

9	Conclusion	112
9.1	Pupil learning and quality of education	112
9.2	Teacher performance	113
9.3	School leadership and management	114
9.4	District and regional education management	115
9.5	Community participation and accountability	116
10	Follow-up quantitative survey and qualitative research	117
	Bibliography	119
Annex A	Original and agreed terms of reference	127
A.1	Original terms of reference	127
A.2	Agreed terms of reference	134
A.2.1	Our understanding of the agreed terms of reference	134
A.2.2	Changes to the impact evaluation design since the technical proposal	138
Annex B	Impact Evaluation Districts	138
Annex C	Constraints underpinning the EQUIP-T programme theory of change	140
Annex D	EQUIP-T programme summary description	141
D.1	Component 1: improving teacher performance	141
D.2	Component 2: school leadership and management	141
D.3	Component 3: district planning and management	143
D.4	Component 4: community participation and accountability	144

List of figures and tables

Figure 1 Complete, expanded EQUIP-T programme theory of change	xx
Figure 2 EQUIP-T programme TOC.....	2
Figure 3 IE districts	7
Figure 4 Distribution of the PSLE pass rate, 2013.....	12
Figure 5 Distribution of pupils by curriculum-linked performance band in Kiswahili	14
Figure 6 Oral reading speed from a familiar word list	16
Figure 7 Oral reading speed from a story passage.....	17
Figure 8 Distribution of pupils by curriculum-linked performance band in mathematics	18
Figure 9 Mean test scores on different mathematics skill areas by home language	19
Figure 10 Mean test scores on different mathematics skill areas by gender.....	20
Figure 11 Background characteristics of pupils in standard three.....	22
Figure 12 Background characteristics of schools	24
Figure 13 Overview of school infrastructure	25
Figure 14 Head teacher and teacher retirement during the EQUIP-T programme.....	27
Figure 15 Absenteeism of head teacher.....	31
Figure 16 Absenteeism of head teacher in previous five days.....	32
Figure 17 INSET on development planning and management for head teachers.....	33
Figure 18 WSDPs	34
Figure 19 Capitation grant record keeping	35
Figure 20 Most common teacher performance management practice reported by head teachers.....	36
Figure 21 Teacher performance management practices reported by teachers.....	37
Figure 22 Staff meetings held in the last 60 days.....	38
Figure 23 Rewards for good teacher performance and teacher ratings of school support	38
Figure 24 Meeting instructional time requirements	39
Figure 25 TDNA results in Kiswahili	42
Figure 26 TDNA results in mathematics.....	43
Figure 27 Gender and spatial balance in teacher interaction with pupils	46
Figure 28 Distribution of teacher–pupil interactions across the classroom	47
Figure 29 Observed teaching behaviours during introductory and concluding lesson stages	48
Figure 30 Observed teaching behaviours during middle stages of lessons	49
Figure 31 Variety in types of pupil assessments used by teachers.....	50
Figure 32 Absence and punctuality of teachers.....	52
Figure 33 Official requirements, timetabled and estimated actual instructional time in mathematics and Kiswahili for standards 1 and 2	53
Figure 34 Key features of EQUIP-T Component 4, Community participation and accountability ...	56
Figure 35 Use of school notice boards to communicate with communities	57
Figure 36 Availability of minutes from parent-teacher meetings and topics discussed.....	59
Figure 37 Teacher reporting of individual pupil progress to parents	60
Figure 38 Availability of SC meeting minutes and topics discussed	62
Figure 39 Head teacher rating of SC and community support to school.....	64
Figure 40 Standards 1–3 pupil absence rates from school on day of survey	66
Figure 41 Amount of per pupil capitation grant payments received meets expectation.....	71
Figure 42 District visits to school and individual meetings with teachers	73
Figure 43 Frequency of district visits to school and meetings with teachers	74
Figure 44 Head teacher rating of WEC support to school, by frequency of visit	74
Figure 45 Proportion of schools receiving different types of in-kind resources	76
Figure 46 Causal mechanism: Better quality education, especially for girls.....	79
Figure 47 Causal mechanism: Improving school readiness, the learning environment for girls and teacher performance.....	83
Figure 48 Causal mechanism: Improving school leadership and management	94
Figure 49 Causal mechanisms: Strengthening district and regional management.....	100
Figure 50 Causal mechanism: Strengthening community participation and accountability	105
Figure 51 Complete expanded TOC.....	110

Table 1 Core activity areas under EQUIP-T components one to four	3
Table 2 Overview of IE instruments and respondents	5
Table 3 Summary of quantitative survey sample sizes (actual at baseline)	6
Table 4 Summary of qualitative survey sample sizes (actual at baseline)	8
Table 5 Characteristics of programme treatment districts compared to other districts in Tanzania ¹	8
Table 6 Background characteristics of head teachers and teachers.....	25
Table 7 School leadership responsibilities: Qualitative findings from case study schools	30
Table 8 TDNA results in mathematics by standards taught by teachers ¹	44
Table 9 Summary of strength of assumptions underpinning the causal links.....	80
Table 10 Summary of causal links and strength of underpinning assumptions for EQUIP-T component one	84
Table 11 Summary of causal links and strength of underpinning assumptions.....	94
Table 12 Summary of causal links and strength of underpinning assumptions.....	101
Table 13 Summary of causal links and strength of underpinning assumptions.....	106
Annex table 1 What the impact evaluation will measure.....	137
Annex table 2 Impact evaluation regions and districts	139
Annex table 3 EQUIP-T component 1 core activities	141
Annex table 4 EQUIP-T component 2 core activities	142
Annex table 5 EQUIP-T component 3 core activities	144
Annex table 6 EQUIP-T component 4 core activities	145
Annex figure 1 Constraints on children’s capability to learn to their full potential	140
Annex figure 2 National School Leadership Competencies Framework	143

List of abbreviations

BRN	Big Results Now
CAPI	Computer-Assisted Personal Interviewing
CI	Confidence interval
DEO	District Education Officer
DFID	Department for International Development
DSI	District School Inspector
EGMA	Early Grade Mathematics Assessment
EGRA	Early Grade Reading Assessment
EMIS	Education Management Information System
EQUIP-T	Education Quality Improvement Programme in Tanzania
FGD	Focus Group Discussion
HH	Household
ICT	Information and Communication Technology
IE	Impact Evaluation
IEC	Information, Education and Communication
ICC	Intra Cluster Correlation
IGA	Income-Generating Activity
INSET	In-service Teacher Training
KII	Key Informant Interview
LANES	Literacy and Numeracy Education Support Programme
M&E	Monitoring and Evaluation
MA	Managing Agent
MoEVT	Ministry of Education and Vocational Training
NECTA	National Examinations Council of Tanzania
NSLCF	National School Leadership Competencies Framework
OPM	Oxford Policy Management
PFM	Public Financial Management
PMO-RALG	Prime Minister's Office Regional Administration and Local Government

PSA	Programme Support Activities
PSLE	Primary School Leaving Examination
PSM	Propensity Score Matching
PSM-DID	Propensity Score Matching-Differences in Differences
PTG	Parents–Teachers Group
PTR	Pupil–Teacher Ratio
RCT	Randomised Control Trial
REO	Regional Education Officer
RTI	Research Triangle Institute International
SACMEQ	Southern African Consortium for the Measurement of Education Quality
SC	School Committee
SDIS	Service Delivery Indicators Survey
SES	Standardised Effect Size
SEQAS	Specialist Evaluation and Quality Assurance Services
SLM	School Leadership and Management
SP	Service Providers
TA	Technical Assistance
TCF	Teacher Competency Framework
TDNA	Teacher Development Needs Assessment
TIE	Tanzania Institute of Education
TOC	Theory of Change
TOR	Terms of Reference
TPMS	Teacher Performance Management System
UNESCO	United Nations Educational, Scientific and Cultural Organisation
UNICEF	United Nations Children's Fund
WEC	Ward Education Coordinator
WSDP	Whole School Development Plan

PART A: Background

1 Introduction

This final baseline report presents findings from a baseline survey and qualitative fieldwork conducted for the IE of the EQUIP-T. The purpose of the IE is to:

- Generate evidence on the impact of EQUIP-T on learning outcomes for pupils in primary education, including any differential impacts for girls and boys;
- Provide evidence on the fiscal affordability of scaling up EQUIP-T;⁹
- Assess perceptions of effectiveness of different EQUIP-T components; and
- Communicate evidence generated by the IE to policy-makers and key education stakeholders.

For full details on the scope of the IE, see the terms of reference (TOR) in Annex A.

This draft baseline report is organised into two volumes. Volume I presents an overview of the evaluation, the baseline results for programme treatment districts, and an assessment of the EQUIP-T programme TOC. It is designed to be accessible to all readers. Volume II covers the technical and methodological detail, and is intended for those interested in methods, detailed statistical results, and the detailed qualitative background paper.

1.1 Objectives of the baseline analysis

The overall objectives of the baseline analysis are to:

- **Capture the education situation, and perceptions of the situation, prior to the EQUIP-T programme's start** so that changes can be measured during the follow-up data collection rounds, and any impact attributable to the programme assessed;
- **Provide evidence to inform the EQUIP-T programme TOC** and potential programme adjustments at an early stage; and
- **Provide an assessment of the status of primary education in some of the most educationally disadvantaged regions in Tanzania** to the government and other stakeholders in the education sector.

The two key parts of this report are: i) the integrated quantitative and qualitative baseline results for programme treatment districts in areas the EQUIP-T programme will seek to influence – pupil learning, SLM, teachers' professional capacity, performance, motivation and morale, community participation and accountability, and district and regional management; and ii) a mapping and analysis of an expanded programme TOC, drawing out inter-linkages between expected changes and assessing the core underlying assumptions.

This report also includes the quantitative results from the IE control districts for the set of impact indicators that will be used to assess impact after the follow-up rounds of the survey. These are reported as part of a methodological discussion on measuring impact in Volume II (Section 3.4 and 3.5).

⁹ The fiscal affordability study is not part of this baseline report, as this study will be carried out much later into programme implementation. A summary of the fiscal study concept note is in Volume II of this report (Annex N).

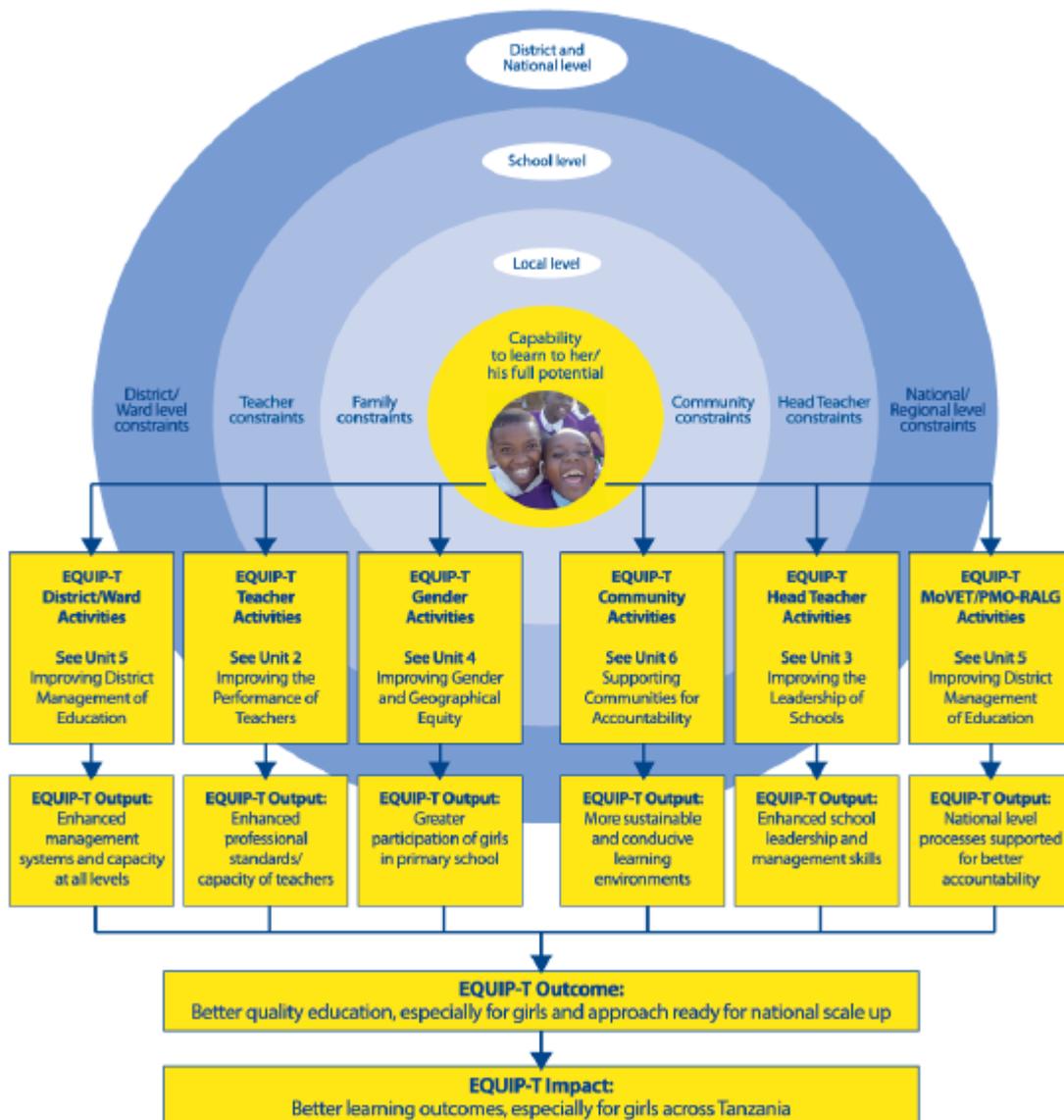
The baseline quantitative survey took place from the end of March to mid-May 2014, followed by the qualitative fieldwork which started in June 2014 and was completed by mid-August.

The main audience for this report is DFID, the EQUIP-T MA, the government and the broader group of education stakeholders in Tanzania. This report will be supplemented by two stand-alone policy briefs on teacher capacity and performance and pupil learning.

1.2 Overview of the EQUIP-T programme

EQUIP-T is a four-year, DFID-funded, Government of Tanzania programme that will seek to improve the quality of education, especially for girls, in seven regions in Tanzania (Cambridge Education 2014b). The programme was designed based on a TOC captured in Figure 2. It identifies six groups of constraints acting on pupils' capability to learn to their full potential.

Figure 2 EQUIP-T programme TOC



Source: Cambridge Education (2014b).

The programme's overarching theory is that by reducing or removing these constraints, the quality of education and pupil learning will improve. The programme has identified multiple constraints at

each level (see Annex B) and has grouped its interventions into five components (reduced from the six shown in Figure 2), each related to a set of constraints. Each component is linked to a programme output. Gender is a cross-cutting theme, and gender-specific interventions are included under each component. The five outputs are:

- Output 1: enhanced professional capacity and performance of teachers;
- Output 2: enhanced SLM skills;
- Output 3: strengthened systems that support the district and regional management of education;
- Output 4: strengthened community participation and demand for accountability; and
- Output 5: strengthened learning and dissemination of results.

The five components are closely related to each other, and taken together they are intended to reduce the constraints on pupils' capability to learn at various levels (community, school, district and regional), and so contribute to better-quality education (outcome) and improved pupil learning (impact).

The IE will cover the first four outputs and key aspects of the related TOC. Table 1 contains a list of the core activities planned for each of these components to give a broad picture of the main interventions. For further details on the programme, and in particular the links between the identified constraints, component objectives, inputs, activities and expected outputs, see Annex C.

Table 1 Core activity areas under EQUIP-T components one to four

C1: Improving teacher performance	C2: SLM	C3: District planning and management	C4: Community participation and accountability
TCF	Whole school planning	Review of district management systems	Improving communication mechanisms
Teacher professional development	School quality framework	Strengthening planning and management capacity	Community engagement in education planning
Teacher performance management	School information system	Strengthening of EMIS system	School IGAs
Teacher Morale Toolkit	Performance management system for school leaders	Strengthening of financial planning, budgeting and management	Building the capacity of SCs
School-readiness programme		Strengthening of M&E systems	Parent-teacher group formation
Teacher training scholarships for rural candidates			

Source: Cambridge Education 2014b, pp. 32–43.

1.3 Overview of the impact evaluation

This section gives an overview of the main features of the overall IE and then discusses key elements of the baseline round. The aim is to give enough information to set the context for the results in Part B while the IE technical design and details are discussed in the methods chapters in Volume II.

1.3.1 Overview of impact evaluation design

The IE aims to measure the impact of EQUIP-T overall on pupil learning, SLM, and teachers' professional capacity and performance through the quantitative research. Perceptions of changes in district and regional education management influenced by the EQUIP-T programme, and of the effectiveness of EQUIP-T activities to increase community participation and demand for accountability in education, will be explored through the qualitative research informed by data from the quantitative component. The qualitative research will also explore stakeholder views on the pathways contained in the programme TOC.¹⁰

The IE will collect qualitative and quantitative data at three points in time: 2014 (baseline), 2016 (midline) and 2018 (endline).

The IE has a mixed methods design that is well suited to the evaluation of a complex programme like EQUIP-T that covers multiple levels of the education system. It combines quantitative and qualitative research components to yield an in-depth understanding of key EQUIP-T impacts and channels of programme influence.¹¹ Specifically, the IE integrates methodologies for better measurement, sequences information for better analysis, and will merge the qualitative and quantitative findings into a single, mixed methods report for each round of the IE.

The EQUIP-T programme TOC informs the IE as a whole, but is particularly important for the qualitative component because it permits stronger, though not full, generalisation and attribution of impact. Specifically, the programme TOC is used to map out the programme's causal chain and contextual assumptions that must hold for programme activities to lead to the intended impact (White 2009).

The IE will use the quantitative component to measure programme impact for selected outcomes and outputs, but also qualitative data to conduct 'rigorous factual analysis' on whether the expected links in the TOC causal chains hold and whether the TOC assumptions are valid over time (White 2009). This rigorous factual analysis will only be possible for some of the links in the causal chain, given resource constraints, which means that the IE will not constitute a theory-based evaluation to the full extent set out in White (2009). Rather the baseline quantitative and qualitative results will assist in further focusing subsequent rounds of qualitative data collection on specific issues in the context and on causal pathways that are considered by stakeholders and researchers to be the most significant. To this extent, the evaluation will be based in theory and thus amenable to a certain amount of speculation about what would have happened if different components of EQUIP-T were scaled up.

Overall then, a key strength of this mixed-methods design is that it provides for iteration, development and deepening of lines of enquiry through the life of the evaluation. The quantitative component will allow rigorous attribution of changes in selected outcomes to EQUIP-T as a whole, while the qualitative methods allow for the probing and exploration of the relationships between any such changes and EQUIP-T. This means seeking to understand why the programme had an impact or otherwise or identifying heterogeneous impact, using evidence on the strengths and weaknesses of linkages in the causal chain. The qualitative research will also seek to explore

¹⁰ Strictly speaking, since the IE research only takes place in a selected number of programme districts (see Section 1.3.3 for details), the IE approach could be better described as testing the EQUIP-T model rather than testing the EQUIP-T programme.

¹¹ The IE will also examine whether EQUIP-T scale-up would be affordable given the likely future fiscal context through its fiscal affordability component (this separate study is not discussed in this report).

changes in outcomes not amenable to quantification (for instance, understanding of responsibilities), as well as to contextualise any such changes.

1.3.2 Instruments used for the impact evaluation baseline

While the quantitative survey focuses on collecting data in a representative sample of programme treatment schools, the qualitative research collects data in a small purposive sample of districts, wards, communities and schools and from EQUIP-T staff at the national, regional and district levels.

Table 2 gives an overview of the types of qualitative and quantitative instruments used to collect the IE baseline data. Details on these instruments can be found in the methods chapters in Volume II.

Table 2 Overview of IE instruments and respondents

Respondents	Quantitative survey instruments (200 government primary schools)	Qualitative research tools (small purposive sample)	
		FGDs	KIIs
Pupils	Kiswahili test; Mathematics test	X	
Parents	Poverty scorecard	X	
Community leader			X
SC		X	
Teachers	Teacher interview; TDNA in Kiswahili; TDNA in mathematics; Lesson observation	X	
Head teachers	Head teacher interview and school record data collection		X
District officials			X
Ward officials			X
Regional officials			X
Academics			X
EQUIP-T programme staff			X

Source: OPM IE team.

To measure the quantitative indicators, the IE administered eight different survey instruments. The respondents were standard three pupils, their parents, teachers and head teachers. Enumerators also observed standard two Kiswahili and mathematics lessons. Computer-assisted personal interviewing (CAPI) was used to collect the data except for in the case of the TDNA instrument, which was administered on paper as this takes the form of mock pupil tests that teachers mark.

The qualitative component of the IE used KIIs to collect information from head teachers, community leaders, regional, district and ward education officials, academics and EQUIP-T programme staff. FGDs were used to elicit the views of pupils, parents, teachers and SC members. All of the KIIs and FGDs utilised structured and unstructured methodologies. The structured methods allowed for the testing of pre-specified hypotheses and the unstructured methodologies for unanticipated or context-specific information to be captured and for new hypotheses to be developed.

1.3.3 Sampling strategies and sample sizes for the impact evaluation baseline

This section summarises the core features of the sampling strategy for the IE baseline. It describes the key principles but leaves the full technical details to the methods chapters in Volume II. The quantitative survey uses a quasi-experimental design with multi-stage sampling, while the qualitative research uses a small purposive sample.

The quantitative sample was drawn by starting with a group of 17 ‘pure’ treatment districts. This group is a subset of the purposive sample of 29 EQUIP-T programme districts selected by the MA (excluding Lindi and Mara, which are not covered by the score of the IE). The ‘pure’ group are treatment districts in which no programme similar to EQUIP-T is in operation and/or that will not receive partial EQUIP-T interventions. Figure 3 shows the 17 treatment districts that are part of the IE, as well as the additional 12 treatment districts excluded from the IE sample mainly because other programmes are operating there.

From these 17 ‘pure’ treatment districts, a random sample of 100 government primary schools was drawn. During the fieldwork, enumerators then drew random samples of pupils and teachers from those present on the day of the survey within the sampled schools. The actual sample sizes for treatment units are shown in the first column of Table 3.

In order to evaluate the impact of the programme in future rounds, the IE requires an appropriate control group that does not benefit from the programme that can be used as a counterfactual to the treatment group. The IE uses a PSM technique to construct an appropriate control group of districts and schools (see the quantitative methods chapter in Volume II).¹² The eight control districts are visible in Figure 3 and Table 3 summarises the baseline sample sizes for each sampled unit in the control districts.

Table 3 Summary of quantitative survey sample sizes (actual at baseline)

Sampling unit	Treatment	Control	Total
Districts total	17	8	25
Schools total	100	100	200
St. 3 pupils per school tested in Kiswahili and mathematics	1,497	1,490	2,987
Poverty scorecard for tested pupils’ parents	1,443	1,450	2,893
Teachers of stds 1–3 Kiswahili and mathematics interviewed	329	352	681
Teachers of stds 1–3 administered Kiswahili TDNA	247	263	510
Teachers of stds 1–3 administered mathematics TDNA	246	259	505
Teachers of stds 4–7 administered mathematics TDNA	283	281	564
St. 2 lesson observations in Kiswahili and mathematics	199	198	397

Sources: IE baseline survey. Notes: (1) Lesson observations were not randomly sampled, but were instead chosen by convenience based on the school timetable in operation on the day of the survey; (2) Information was not available from four teachers on which standards they were assigned to teach, while 26 teachers taught standards one to three *and*

¹² Enumerators followed an identical procedure to that used in the treatment schools to draw random samples of control group pupils and teachers.

stage, three schools within each of the three selected programme districts were sampled using 'typical case sampling' and 'extreme case sampling' so as to ensure a diverse collection of cases.

Table 4 Summary of qualitative survey sample sizes (actual at baseline)

Sampling unit	Regions total	Districts total	Schools total	FGDs with std. 3 pupils	FGDs with parents of std. 3 pupils	FGDs with SC members	FGDs with teachers	KIIs with HTs	KIIs with community leaders	KIIs with REOs	KIIs with DEOs	KIIs with WECs ¹	KIIs with EQUIP-T regional staff	KIIs with EQUIP-T national staff
Total	3	3	9	9	9	9	9	9	9	2	2	9	3	6

Note: (1) Since this baseline research took place, WECs have been retitled as Ward Education Officers.

1.3.4 Generalisability of baseline results

The baseline quantitative results presented in Volume I of this report are statistically representative of the 17 'pure' treatment districts. These 'pure' treatment districts that are part of the IE appear to be fairly similar on average to the remaining 12 programme districts in terms of key contextual characteristics (see the first two columns of Table 5 for a comparison). However, the programme treatment districts are on average significantly different from other districts in Tanzania, as can be seen from the last two columns of Table 5. The programme treatment districts are significantly more rural; have a significantly higher PTR; have a significantly smaller proportion of households with electricity and books at home; there is a significantly smaller proportion of children in school; and there is a larger proportion of mothers with no education compared to other districts in Tanzania. This means that the baseline quantitative results in this report are not generalisable beyond the 17 programme treatment districts, although they may provide indications in cases of districts highly similar to these treatment districts.

Table 5 Characteristics of programme treatment districts compared to other districts in Tanzania¹

Characteristic	Treatment IE sample (part of EQUIP-T at BL)	Not in IE treatment sample (part of EQUIP-T at BL)	Part of EQUIP-T at BL	All other districts (Not part of EQUIP-T at BL ²)
% rural villages	90.9*	94.8	92.5***	73.0
PTR	51.3	51.4	51.3***	47.1
% children in school	82.9	84.4	83.5***	88.7
% mothers no education	32.4***	24.3	29.0***	18.8
% households with electricity	8.0	10.6	9.1***	17.6
% households with books at home	11.4*	14.7	12.8***	17.1
Number of Districts	17	12	29	137

Sources: IE baseline survey, UWEZO 2013. Note: (1) Statistically significant differences between groups are marked with asterisks: *significant at the 10% level; **significant at the 5% level; ***significant at the 1% level. The group comparisons are between columns 1 and 2, and between columns 3 and 4. (2) Lindi and Mara districts are not in EQUIP-T at baseline (BL), so they are included in column four only. They will be included in EQUIP-T in 2015.

The qualitative TOC analysis, also presented in Volume I, draws on both the quantitative and qualitative baseline results and is the first step in the theory-based aspect of the evaluation. Following future data collection rounds, which seek to understand the reasons and context in which components of EQUIP-T lead to impacts, this will allow some speculation about the likely impact of EQUIP-T if it were to be scaled up to other parts of the country (see Section 1.3.1 for more details).

1.4 Pointers on the interpretation of baseline results

One of the most important aims of the baseline IE research is to provide information which is potentially useful to the MA, DFID and the government for informing any potential adjustments of the programme design, and to understand the current education situation in some of the most educationally disadvantaged districts in the country. For this reason, Volume I of this report presents the quantitative **results from the programme treatment districts** (appropriately weighted, see below, so they can paint a picture of the typical situation in these disadvantaged areas). In addition, Volume I contains the TOC analysis which relies on baseline quantitative and qualitative results from the programme districts, as well as a broader literature review. This is highly relevant to all stakeholders interested in EQUIP-T as it draws some indicative conclusions on the strength of various assumptions underpinning the programme's causal pathways.

The baseline **results from the control districts** are less informative for a broad audience at this stage, because they are not statistically representative of the control districts overall (since they were selected using PSM). The main purpose of the control group is its use in estimating programme impact, once data is available from the 2016 and 2018 rounds. The control group estimates are thus presented in Volume II, as part of a section on impact estimation methodology (Section 3.5). This section shows that there are statistically significant differences between unweighted treatment and control group estimates for roughly half of the 22 impact indicators, as well as some of the pupil and teacher background characteristics. These results are not unexpected given that the treatment and control groups were not randomly selected from the same overall population. In order to estimate programme impact in 2016 and 2018, the treatment and control group samples will be matched using PSM to achieve balanced samples, and then a difference-in-difference approach applied to estimate impact. An illustration of the matching methodology is in the section.

1.4.1 Weighted estimates for programme treatment districts

To provide estimates of key quantitative indicators that are representative of the 17 programme treatment areas, the observed values were analysed using survey weights (see the quantitative methods chapter in Volume II). This is necessary because, although the units of analysis (pupils, teachers, and schools) were randomly sampled, they were not sampled with equal probability and so un-weighted averages, for example, would be misleading. Survey weights are used to account for this. The lesson observation results are an exception because lessons were not sampled at the school level and are thus considered 'school characteristics' so that school weights are applied

1.4.2 How to read the figures and tables in Part B

All the quantitative figures and tables show the mean estimate, as a summary indicator. To give information on the spread of results, some of the tables and figures contain percentile estimates. These are estimated values below which a given percentage of observations lie. For example, the 10th percentile (labelled P10 in the tables) is the estimated value below which 10% of observations lie. Similarly, the 90th percentile (labelled P90) is the estimated value below which 90% of observations lie.

In some cases, the bar charts with mean estimates also show 95% confidence intervals, which is the range of values for which it can be stated that there is a 95% probability that the estimated confidence interval encompasses the true value of the population parameter.

All figures and tables display the sample size (N), i.e. the number of respondents who answered a particular question, for each indicator.

In some tables, the estimates have asterisks, which indicate a statistically significant difference between the groups shown: *significant at 10% level **significant at 5% level ***significant at 1% level. The more asterisks are shown, the more likely it is that the observed difference is due to real differences between the groups rather than to chance because of who was interviewed or tested.

Some tables and figures compare estimates for different sub-populations, for example male and female pupils. It is important to remember that statistically significant differences between results for sub-populations do not necessarily indicate a causal relationship.

1.5 Structure of this volume

This is Volume I of the Interim EQUIP-T IE Baseline Report, which is accompanied by *Volume II: Methods and Technical Annexes*. It is structured as follows.

Part A describes the objective of the baseline analysis and provides an overview of the EQUIP-T programme to be evaluated and a summary description of the IE including descriptions of the survey instruments, sampling strategies and sample sizes, and a brief discussion of the generalisability of results. It concludes with some pointers on how to interpret the baseline estimates.

Part B presents the quantitative and qualitative baseline results for the programme treatment districts (the results for the control districts are in Volume II).

Part C presents an expanded programme TOC, which adds detail to the EQUIP-T TOC discussed in this chapter. It then analyses the causal links within the expanded TOC, drawing on wider literature as well as the quantitative and qualitative baseline findings.

Part D concludes.

The annexes of this volume contain the following: the original and agreed TOR (Annex A), a table of the IE districts (Annex B), a schematic representation of the constraints underpinning the EQUIP-T TOC (Annex B) and a programme summary description (Annex C).

PART B: Baseline results for programme treatment districts

Part B presents the mixed methods results from the IE baseline for **programme treatment districts (the estimates for the control group districts are in Volume II)**. The key intended impact of the EQUIP-T programme is to improve pupil learning outcomes. Therefore, the baseline findings on pupil learning overall and by learning domain, and for different groups (poorer and richer pupils, boys and girls, and pupils who speak Kiswahili compared to a local language other than Kiswahili at home) are reported and discussed first. Next, findings on pupil, school and head teacher and teacher characteristics are provided in order to set out the programme context. This is followed by findings on each of the four outputs the programme will seek to change and that are covered by the IE: SLM; teacher professional capacity, performance, motivation and morale; community participation and accountability; and district and regional education management.

2 Pupil learning outcomes

The overall goal of the EQUIP-T programme is to improve learning outcomes and education quality, especially for girls. The programme is starting many of its interventions at the lower primary school level, including, for example, in-service teacher training programmes that will initially be delivered to teachers of standards one to three. For this reason, it expects to see measurable improvements in pupil learning in early standards, as well as the narrowing of gender gaps in achievement, within the programme timeframe.¹⁴ The scale of the expected change is not articulated in the programme documents at this stage, although the signal from the original Intervention Summary for EQUIP-T (DFID 2013, p. 38) is ambitious. It proposes more than a doubling of the proportion of early standard pupils with basic literacy skills over the programme period.

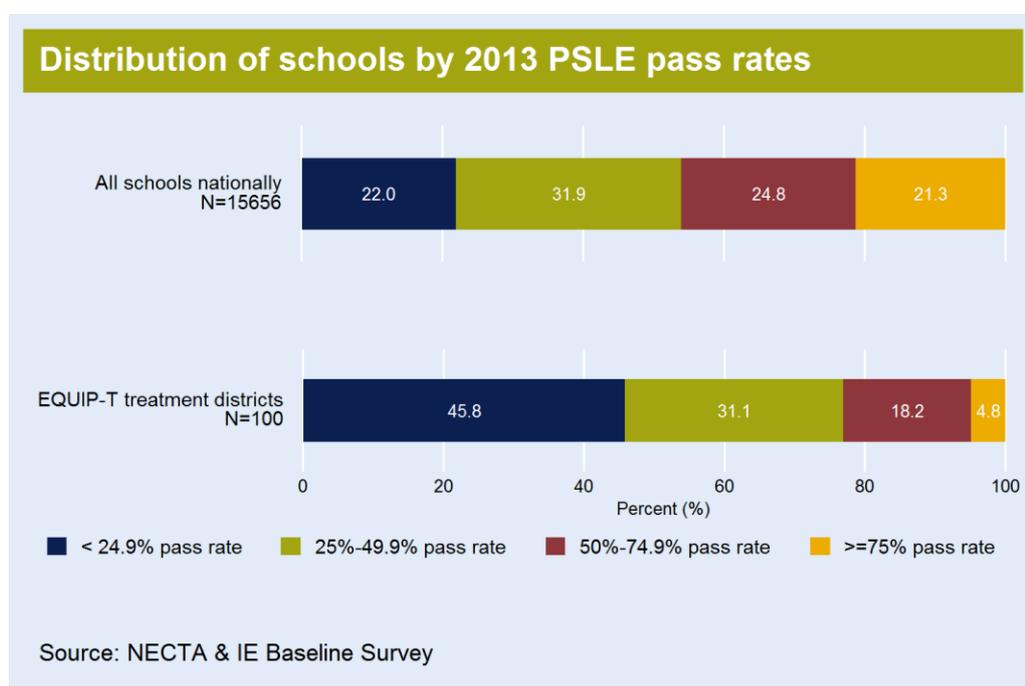
In the medium to longer term, the programme expects to drive up achievement over the whole primary school cycle. The core programme logic is that the early interventions will equip pupils with stronger foundational skills; these pupils will then move up the system and will be in a better position to benefit from later interventions targeted at upper-primary level. Ultimately, one important indicator of overall programme success is the PSLE pass rate. Nationally, PSLE pass rates have been highly volatile over the past three years, dropping from 58% in 2011 to 31% in 2012 and then rising to 51% in 2013, which raises concern about how reliable a measure they are. Looking further back, the trend in pass rates has been strongly downwards, from a peak of 71% in 2006, generating a national debate on the quality of primary education.¹⁵

PSLE results for the programme treatment districts are markedly lower than the national picture. This is clear from the distribution of the 2013 PSLE pass rates in Figure 4. Close to half of schools in programme treatment districts did not even manage a 25% pass rate, while only about one-fifth of schools nationally are in this situation. At the opposite end of the performance spectrum, there is also a huge gap, revealing just how far behind much of the rest of the country the programme treatment districts are. More than 20% of primary schools nationally achieved a pass rate of 75% or above, compared with less than 5% of schools in treatment districts.

¹⁴ Indicators of learning achievement for lower primary school pupils appear in the output and impact sections of the EQUIP-T programme logframe.

¹⁵ See, for example, analysis and comment in Sumra and Katararo (2014). In this report, Annex G in Volume II reproduces a figure from this paper that shows PSLE pass rates from 2001 to 2013 based on government data.

Figure 4 Distribution of the PSLE pass rate, 2013



Notes: (1) IE baseline results are weighted estimates.

Measurement of pupil learning in the IE baseline survey

The IE baseline survey measures learning in the early standards because, given the programme design outlined above, it is reasonable to expect the greatest programme impact at this level. The IE baseline survey tested **standard three pupils in Kiswahili and mathematics at standard one and standard two curriculum levels**, using an oral test that was administered on a one-to-one basis.¹⁶ The IE pupil tests were adapted from an early grade reading assessment (3Rs-EGRA) and an early grade mathematics assessment (3Rs-EGMA), which were recently used to collect baseline data for the government's BRN Education programme (RTI 2014). BRN Education has similar aims to EQUIP-T, but will be implemented in different geographical areas.

Each of the pupil tests covers a number of different skill areas or subtests:

- Kiswahili: reading speed (syllables, familiar words, non-words and a reading passage), reading and listening comprehension, and writing (spelling and punctuating dictated sentences); and
- Mathematics: number comparison, number sequences, addition, subtraction, multiplication and word problems.

There are a couple of important points to bear in mind when interpreting the results presented in this chapter. The first concerns the results for the sub-populations defined by the poverty status of the pupil's household and language spoken at home (Kiswahili or another local language).¹⁷ Across many of the key learning indicators there are statistically significant differences between

¹⁶ Standard three pupils were chosen because the IE baseline survey took place in March/April and it was not informative to use standard two pupils who would only have had one term of standard two lessons.

¹⁷ The definition of the poverty and home language variable is given in Volume II, Annex F.

poorer and richer pupils, as well as between pupils with different home languages. Further analysis reveals that poverty and home language variables are strongly correlated with each other, but also that each is independently correlated with pupil learning levels.¹⁸ This means that non-Kiswahili speakers perform worse on the tests than Kiswahili speakers even after taking poverty into account and vice versa, i.e. poorer pupils perform worse on the tests than richer pupils even after taking home language into account. The second point to highlight that the pupils tested are those attending school on the day of the survey. As will be seen later in this report, pupil absence rates are high, and so it likely that if all pupils were tested the results may be similar or worse (assuming that on average pupils who are absent frequently will perform worse than their peers).

What this chapter contains

This chapter focuses on results from the baseline survey on the overall performance of pupils in treatment schools against national curriculum benchmarks in Kiswahili and mathematics. This is useful because it gives some insight into whether pupils are performing at, above or below the curriculum level expected. It also exposes gaps between higher and lower achieving groups in terms of specific skills. For each subject, it sorts pupils into five performance bands, each described by a set of competency statements from the national standard one and two curricula. This analysis is based on Rasch modelling (see Volume II, Annex I for technical details).¹⁹

For completeness, simple descriptive statistics of the raw test scores from each of the subtests are included in Volume II, Annex G. These include some comparison of the IE results with those from the BRN Education study, although there are important limits to comparability. This chapter presents selected highlights from the raw score analysis to illustrate key messages likely to be of greatest relevance to the programme.

The survey analysis is complemented by findings from the qualitative study on perceptions of learning among various stakeholder groups.

2.1 Pupil learning in Kiswahili

2.1.1 Are standard three pupils achieving at the expected curriculum level in Kiswahili?

To answer this question, pupils were estimated to be achieving at the level of one of five ordered performance bands: **Band 0**: below standard one level; **Band 1E**: emerging standard one level; **Band 1A**: achieving standard one level; **Band 2E**: emerging standard two level; and **Band 2A**: achieving standard two level. A list of competencies linked to each band can be found in Volume II, Annex I.

Before reading the results, there are three technical points to note. First, the expected levels used here are guided by the curriculum, and it is worth considering whether these may be over- or

¹⁸ To examine whether poverty and home language are independently correlated with pupil performance, a set of regressions was run including both as explanatory variables. These show that both variables have a significant, negative relationship with pupil performance in Kiswahili and mathematics.

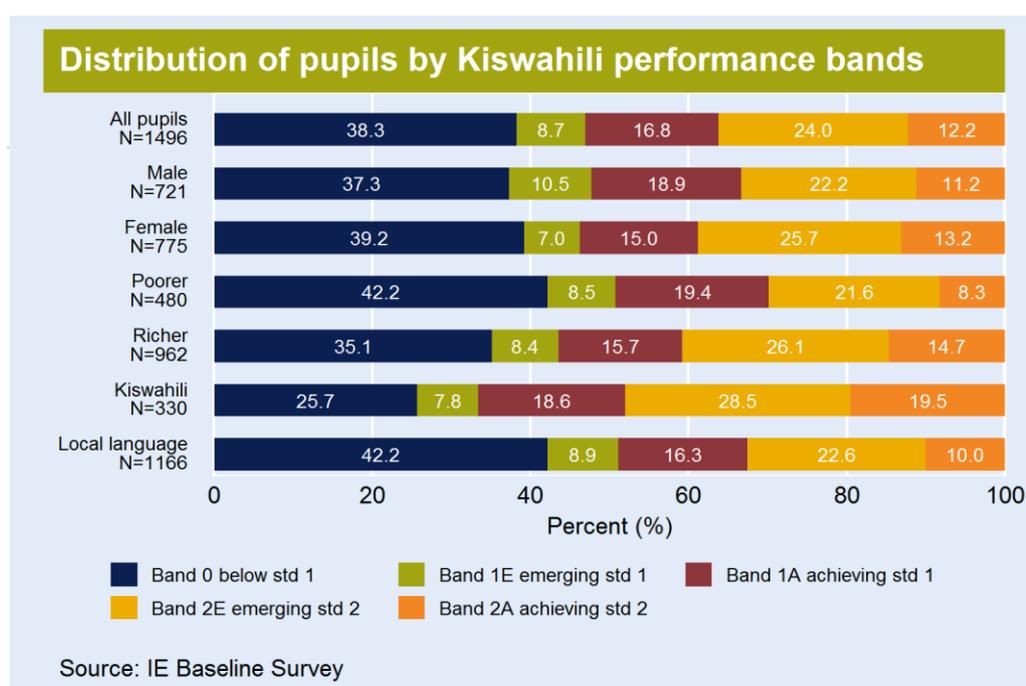
¹⁹ Rasch modelling was used to generate estimates of pupil 'ability' in Kiswahili and mathematics on an interval scale that is directly linked to criterion-referenced competencies in the national curricula. The Rasch model is probabilistic, and is a special case of an item response theory model. More details on the rationale and assumptions underpinning this methodology are in Volume II, Annex I (the pupil learning technical paper).

under-ambitious.²⁰ Research from other developing countries has found that over-ambitious curricula can in themselves be a cause of poor performance (UNESCO 2014, p. 281; Pritchett and Beatty 2012). The second point concerns interpretation. The term ‘achieving at band level’ means that pupils in a particular band are *more likely than not* to be able to demonstrate the skills linked to that performance band.²¹ Third, these results are based on five of the seven subtests in the Kiswahili test, with reading syllables and listening comprehension excluded. Volume II, Annex I explains why these subtests were left out of the Rasch analysis.

Results

Only about 12% of pupils in standard three are currently achieving standard two level skills (falling in band 2A) in core areas of the Kiswahili curriculum, as the top bar in Figure 5 reveals.²² A further quarter of pupils (24%) are estimated to have acquired ‘emerging’ standard two level skills (band 2E). One key difference in skills between these top two bands is reading speed. International research suggests that pupils need to be reading fluently at a minimum rate in order to comprehend what they are reading (Abadzi 2006). Reading comprehension is a core skill specified in the national standard two curriculum, although specific reading speeds are not cited.

Figure 5 Distribution of pupils by curriculum-linked performance band in Kiswahili



Notes: (1) Weighted estimates.

Just over 60% of standard three pupils in the programme treatment districts fall in the three lowest bands. This means that they have not even acquired emerging standard two skills in Kiswahili, and so have fallen at least one full standard behind. Most alarming of all is that nearly four pupils in

²⁰ ‘The term ‘over-ambitious’ here means that the standards set out in the curriculum may be inappropriate because pupils lack the required foundational skills, not that the standards are in principle necessarily inappropriate for pupils of the correct age for standard.

²¹ The results are derived from Rasch analysis, which is based on a probabilistic model of item response.

²² By design, the test is not able to identify pupils who are performing at standard 3 level, but the Rasch analysis did not indicate that there were many pupils with current ‘ability’ estimates above the most difficult items on the test.

every 10 have yet to acquire standard one level skills. This 'band 0' group is at least two standards behind the expected level.

Comparing the distribution of Kiswahili performance for various subgroups of pupils in Figure 5 reveals a similar pattern for boys and girls. There are, however, significant differences in the size of the top and bottom performing groups for the other subgroups (see supplementary tables in Volume II, Annex H for significance tests). Pupils from poorer households are significantly more likely to be in band 0 (less than standard one skills), and less likely to be in band 2A (achieving standard two skills), than their richer peers. Similarly significant, but much larger, gaps are evident between pupils who speak different home languages. Almost 20% of pupils who speak Kiswahili at home are achieving at standard two level (band 2A), compared with a mere 10% of pupils who speak other local languages at home. The reverse situation is apparent at the other end of the performance spectrum. Here, just over 40% of pupils from non-Kiswahili-speaking homes fall into band 0 (below standard one level), compared with one-quarter of pupils with Kiswahili-speaking backgrounds.

In summary, it appears that the majority of standard three pupils have fallen considerably behind curriculum expectations in Kiswahili, and that the learning gap is particularly acute for pupils from poorer and from non-Kiswahili-speaking homes. Kiswahili is the language of instruction in Tanzanian schools, and another recent study also found that the language a pupil speaks at home is correlated with learning levels, specifically whether this is different from the language of instruction (UWEZO 2011).

2.1.2 Key results from the Kiswahili raw test score analysis

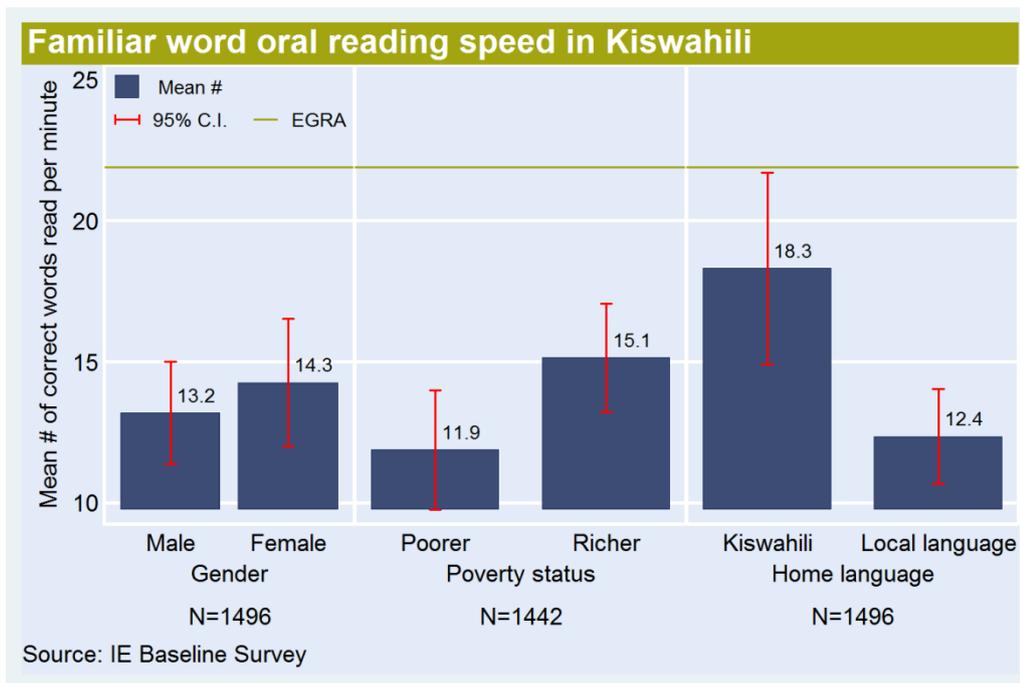
The raw test scores give an insight into pupil performance in different skill areas. This is particularly useful where there is a substantive benchmark to compare performance against. This is the case for oral reading speed, which is the focus of this section.²³

Oral reading speed: Standard three pupils in treatment districts are able to read an average of 21 words per minute from a short story and 14 words per minute from a list of familiar words. The IE familiar word subtest was very similar to the 3R's EGRA subtest, and so Figure 6 contrasts familiar word reading speeds for different groups of IE treatment pupils with the 3Rs EGRA estimate. The pupils in the 3R's sample markedly outperformed the IE programme treatment pupils by eight words per minute on average. This is not surprising given that the 3Rs sample was designed to be nationally representative, and the treatment districts are among the poorest and most educationally disadvantaged in the country.

Figure 6 also reveals that familiar word reading speed differs significantly between subgroups of treatment pupils divided by poverty status and by home language, but not by gender. Pupils from Kiswahili-speaking homes read familiar words at an average speed of 18 words per minute, six more words than pupils who speak other local languages at home. The gap in familiar word reading speed between poorer and richer pupils is three words per minute.²⁴

²³ For information on the marking scheme for the test, see Volume II, Annex I.

²⁴ It is useful to remember that this sample is not representative of Tanzania overall and that the gap between poorer and richer pupils would be larger if more advantaged areas were included.

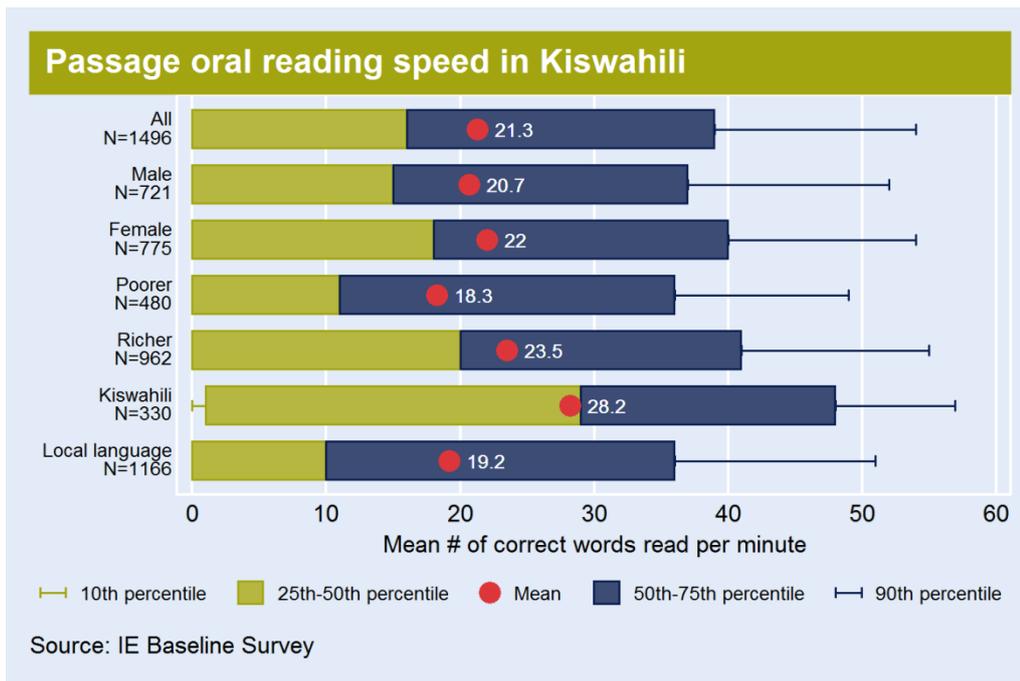
Figure 6 Oral reading speed from a familiar word list

Notes: (1) Weighted estimates.

As part of a recent target-setting exercise for the government's BRN Education programme, representatives of a range of education stakeholders in Tanzania agreed on benchmarks for standard two pupils of an oral reading speed of 50 words per minute. Figure 7 demonstrates that standard three pupils in the programme treatment districts are, on average, not able to read a story passage accurately at half this benchmark speed.²⁵ Another concern is the wide range of oral reading speeds shown in Figure 7. At one extreme, at least one-quarter of pupils were not able to read a single word of the story, while, at the other end, some 10% of pupils can read at speeds of 54 words per minute or more (i.e. above the benchmark). This is consistent with the message from the previous section, which revealed that a small minority of pupils are performing at a curriculum-appropriate level but a large majority have fallen well behind. The significantly stronger oral reading skills of pupils from Kiswahili-speaking homes compared with those who speak other local languages at home, and of pupils from poorer households compared with richer, are again evident.

²⁵ The passage was designed to be appropriate in vocabulary and structure for standard two pupils.

Figure 7 Oral reading speed from a story passage



Notes: (1) Weighted estimates.; (2) All of the bars, except the one marked 'Kiswahili', show that at least 25% of pupils scored zero, meaning that the 10th percentile equals the 25th percentile in this case (with a value of zero words per minute).

Reading comprehension: Pupils were asked five reading comprehension questions based on the passage they had read. On average, pupils were able to answer one question correctly (average score was 19%) and more than half of the pupils were unable to answer any questions correctly. The questions included simple fact-based recall questions, and so this suggests that reading comprehension skills are generally very weak. There may be various underlying reasons for this, including reading fluency and poor vocabulary.²⁶

Writing skills: Pupils were required to spell and punctuate correctly two short dictated sentences. Most pupils struggled with this task. Average scores were 39% for spelling and 30% for punctuation.

For all the subtests, the gap in Kiswahili skills related to home language group is consistently significant and large, with pupils from Kiswahili-speaking homes showing much stronger skills. The same applies to the gap in Kiswahili skills linked to household poverty groups, although the size of this gap is smaller than the 'home language' gap in all subtests. There are no statistically significant differences in Kiswahili subtest scores for boys and girls.

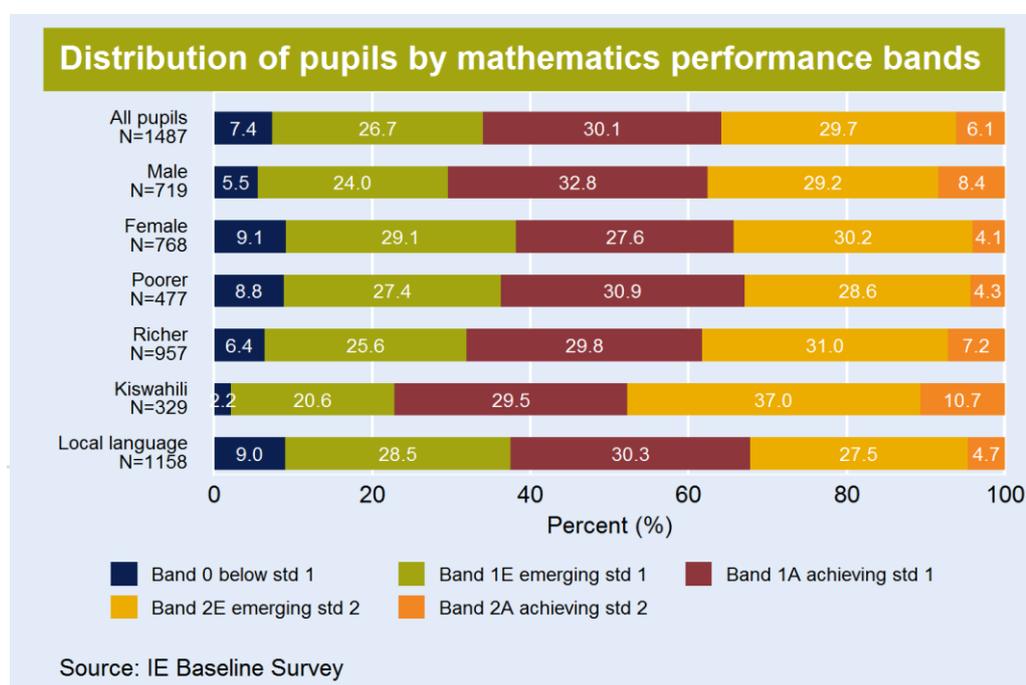
²⁶ The test also included a listening comprehension passage and questions, but diagnostic analysis of the item responses suggests this subtest is problematic and so results are not reported. See Volume II, Annex I for more information on this.

2.2 Pupil learning in mathematics

2.2.1 Are standard three pupils performing at the expected curriculum level in mathematics?

Five ordered performance bands frame the analysis of overall pupil mathematical performance: **Band 0**: below standard one level; **Band 1E**: emerging standard one level; **Band 1A**: achieving standard one level; **Band 2E**: emerging standard two level; and **Band 2A**: achieving standard two level. A list of mathematical competencies linked to each band can be found in Volume II, Annex I. The distribution of pupils in the programme treatment districts according to these five bands is shown in Figure 8.

Figure 8 Distribution of pupils by curriculum-linked performance band in mathematics



Notes: (1) Weighted estimates

The top bar Figure 8 illustrates that only a small percentage of standard three pupils (6%) are achieving at standard two level or above (band 2A). A much larger group of pupils, just under 30%, have emerging standard two skills (band 2E), while the majority of pupils (60%) are achieving at standard one level or at emerging standard one level (bands 1A and 1E). This latter group have fallen more than a year behind the level expected. This leaves a small, but critically important, group of pupils (7%) who are performing below the level of the test. This means that they have not yet achieved even emerging skills at standard one level in the skill areas tested.

Comparing the distribution of mathematics performance for the various subgroups of pupils, there are no significant differences in performance band estimates for poorer and richer pupils, which contrasts with the Kiswahili results discussed earlier. However, the proportion of boys who are achieving at standard two level in mathematics (band 2A) is twice as high as the proportion of girls. Moreover, there is a seven percentage point gap in the share of pupils who fall in the bottom band (below emerging standard one level) between pupils from Kiswahili-speaking backgrounds (2%) and pupils who speak other local languages at home (9%). See the supplementary tables in Volume II, Annex H for significance tests.

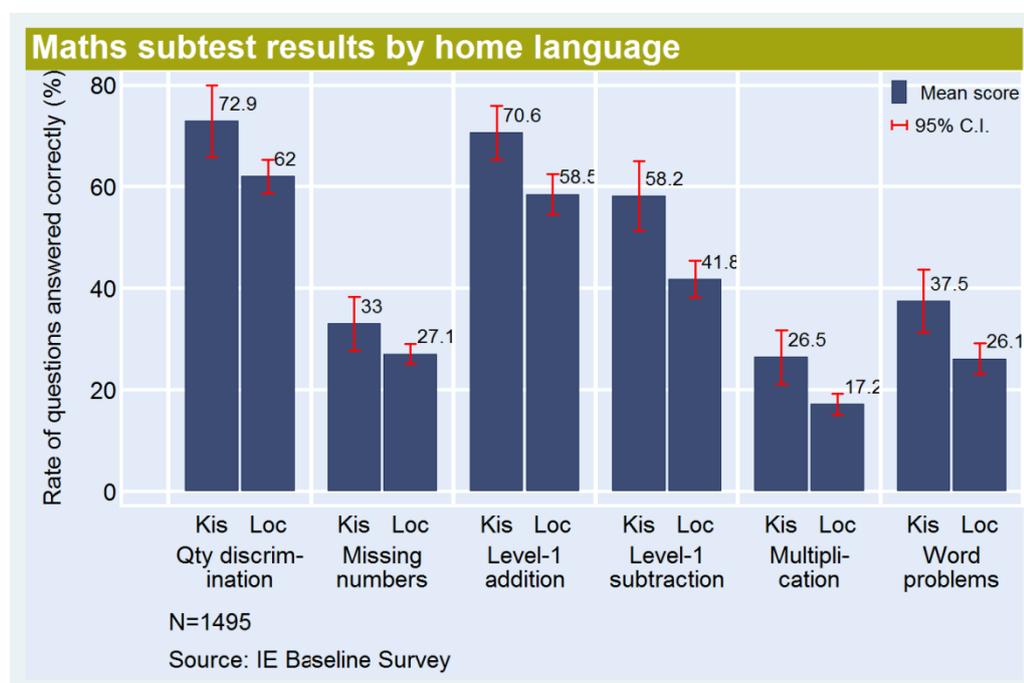
The majority of standard three pupils have fallen considerably behind curriculum expectations in mathematics, although to a lesser extent than for Kiswahili. Again the learning gap is particularly acute for pupils from non-Kiswahili-speaking homes. Mathematics is a hierarchical subject where strong foundational skills are essential for developing higher-order skills later on. Evidence from elsewhere in the region on mathematics performance suggests that learning gaps for disadvantaged groups tend to widen over time, and it can be very difficult to catch up (Spaull and Viljoen 2013²⁷). Unlike Kiswahili skills, there is a gender gap in mathematics skills with boys performing significantly better than girls.

2.2.2 Key results from the mathematics raw test score analysis

The raw test scores give an insight into pupil performance in different skill areas. Unlike for oral reading speed, there are no absolute benchmarks available for arithmetic skills. Instead, this section focuses on comparing the performance of pupils from different home language backgrounds, and of boys and girls, in the different skill areas. For information on the marking scheme for the test, see Volume II, Annex I.

Pupils from Kiswahili-speaking homes have consistently stronger mathematical skills than their peers who speak other local languages at homes, across all of the skill areas on the test. These gaps are clearly visible in Figure 9 which summarises the average scores for both groups on six of the eight subtests. The score gaps are all strongly significant (1% level), except for the missing numbers subtest where the gap is the smallest at six percentage points (significant at 5% level). For the rest of the subtests, the absolute gaps are large, ranging from nine to 16 percentage points. Similarly large and significant gaps are also present for the more difficult level 2 addition and subtraction questions (not shown below).

Figure 9 Mean test scores on different mathematics skill areas by home language

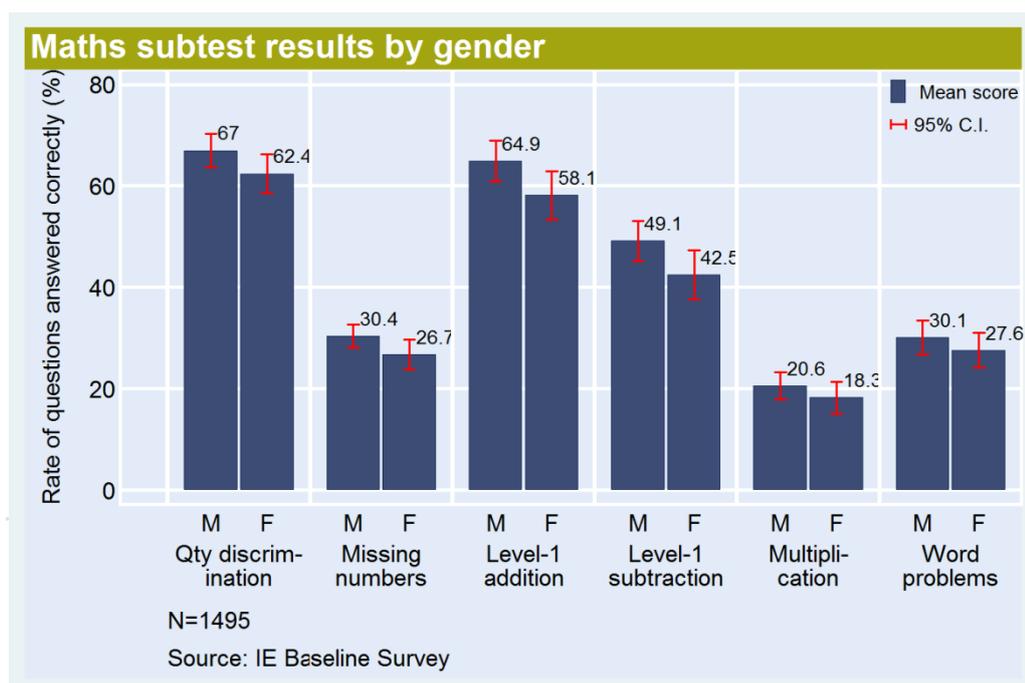


²⁷ Their research uses data from South Africa to show that learning gaps between overage pupils and their appropriately aged peers grew from about one-third of a year's worth of learning in grade 3 to three years' worth of learning by grade 9.

Notes: (1) Weighted estimates

Boys significantly outperform girls on only some elements of the mathematics test and, in these cases, the absolute difference in average scores is far smaller than the gaps related to home language background. As Figure 10 demonstrates, gaps in average scores for the boys and girls range between three and seven percentage points, with the largest differences in performance for level 1 addition and subtraction questions. For the more difficult addition and subtraction questions (level 2, not shown below), and for multiplication questions, the differences in scores by gender are no longer significant. In fact, it is only the first four subtests, containing most of the standard one level items, shown in Figure 10 that the gender gap in scores is strongly significant.

Figure 10 Mean test scores on different mathematics skill areas by gender



Notes: (1) Weighted estimates

2.3 Perceptions of pupil learning

The qualitative research probed the perceptions of various stakeholders in regard to the comparative academic performance of boys and girls. Teachers, head teachers and parents often perceived that girls perform better than boys. They base these judgements on the positive behaviours displayed by girls, including not being involved in negative peer groups and sitting at the front of the classroom. Education officials, school leaders, teachers and parents do not currently assess what happens in the classroom through a gender lens, but rather focus on attendance and completion. By contrast, EQUIP-T programme staff identified gender-based behaviour within the school as a factor affecting the performance of girls. Examples cited by EQUIP-T programme staff included teachers failing to answer girls' questions in class, and teacher and head teacher perceptions of the capacity of female students to understand specific subjects (such as maths and science).

While stakeholders tended to perceive girls as performing better than boys, teachers, head teachers and parents were asked to provide explanations for why many girls in a scenario had failed the national examinations. This was done as a means of probing attitudes towards girls and the constraints girls face.

Overwhelmingly, teachers and head teachers blamed communities and girls for the low performance levels of girls in the scenario, citing ‘love affairs with boys’, the greater domestic load carried by girls, and family influences. Family influences cited by teachers included real and tangible constraints for girls, such as expectations of early marriage. However, many school-level stakeholders’ dialogue also reflected beliefs that female children and the children of ‘illiterates’ inherently lack the capacity to learn. Explanations for the outcome in the scenario given by both parents and teachers included girls not having the capacity to deal with anxiety and stress in examination situations.

The accuracy of overall perceptions of pupil learning levels varies by stakeholder group. The qualitative study found that education official’s judgements regarding the performance of their regions, districts and schools are generally accurate when compared with available quantitative data. Similarly, WECs were aware of the comparative performance of the schools included in the case studies. By contrast, community judgements were less accurate because communities assess the quality of education provided by schools on the basis of teacher attendance (including classroom attendance) and the level of school facilities compared to neighbouring schools.

3 Context: Characteristics of pupils, schools and teachers

This chapter explores the context in which the programme interventions will take place.²⁸ It presents baseline values of programme context indicators from the quantitative survey. These indicators matter for the programme design, but most are highly unlikely to be directly affected by the programme.

3.1 Background characteristics and daily routine of pupils

To capture the background characteristics of pupils, the IE baseline survey used two approaches. First, it asked the sampled standard three pupils about their age and which language they speak at home, while enumerators also recorded pupil gender. Second, parents of the sampled pupils were asked to come to school to answer a set of questions on household characteristics. Their responses were then used to create an estimate of where pupils’ households are predicted to fall in relation to the national poverty line (see Volume II, Annex F for details).²⁹

Among the standard three pupils in the programme treatment districts, 52% are girls and the mean age is 10 years. 57% of pupils were of the correct age for standard three (nine or 10 years),³⁰ 6.5% were under-age (eight or younger) and 37% were over-age (11 years or older) (Figure 11). The reasons for this are that many pupils do not start primary school at the appropriate school-entry age and there is also some grade repetition.

It should be noted that over-age pupil enrolment is an outcome that the EQUIP-T programme will attempt to reduce in the long term by supporting schools and communities to encourage parents to

²⁸ A further reason for measuring some of the pupil background variables is that they potentially affect learning achievement but are unlikely to be influenced by the programme interventions, and so need to be tracked for the analysis of impact in subsequent rounds.

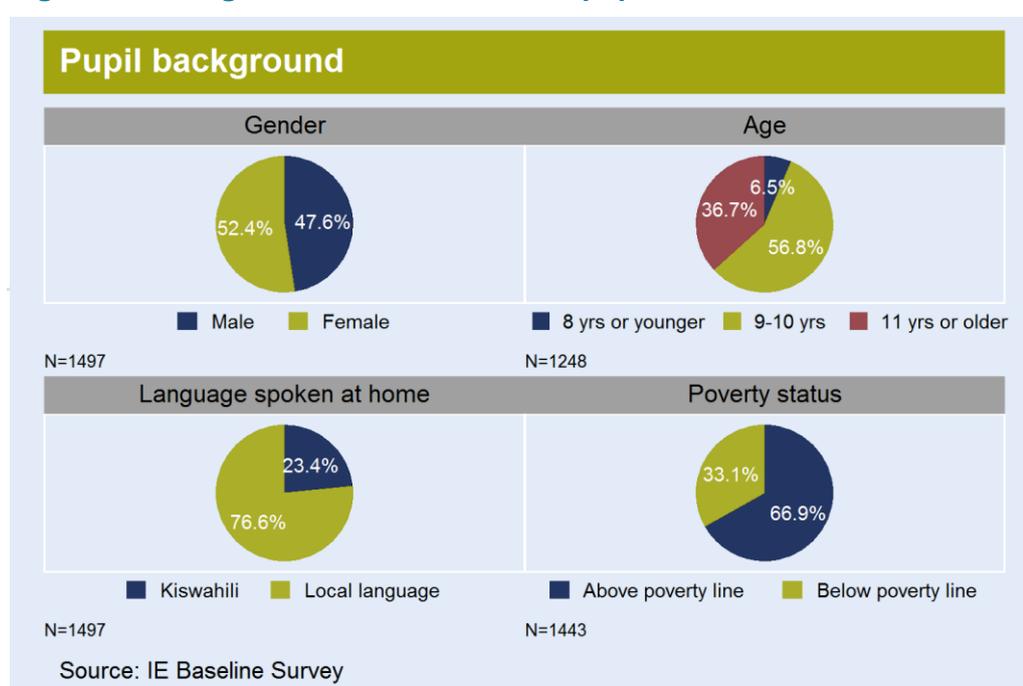
²⁹ A pupil is considered ‘poor’ if he/she comes from a household that has a greater than 50% chance of being below the Tanzania national poverty line, and ‘rich’ otherwise.

³⁰ The age of entry to primary school (standard one) is seven years, but the exact rules were not available from the MoEVT, MA or official documents. If children have to be age seven before they enter primary school then we would expect to find nine and 10 year olds in standard three if there was no repetition. On the other hand, if children can turn seven during standard one then we would expect to find eight and nine year olds in standard three, assuming no repetition.

enrol their children in school at the correct age (Cambridge Education 2014b, OPM 2014a). The extent of over-age enrolment is greater among boys than girls: 30% of female pupils in standard three were over-age compared to 44% of male pupils.³¹ The higher rates of over-age enrolment among boys imply that there may be particular challenges related to enrolling boys at correct school-entry age.

One-third (33%) of pupils belong to a household predicted to fall below the national poverty line. This clearly shows that a large proportion of the pupils come from low socioeconomic backgrounds, which would be expected given that EQUIP-T targets educationally disadvantaged districts that tend to be more remote and poorer (Cambridge Education 2014b). Among the pupils, the majority (77%) speak a local language other than Kiswahili at home. Many aspects of disadvantage tend to go hand in hand and this is the case among pupils in the programme treatment districts. Of pupils who speak a language other than Kiswahili at home, 38% are predicted to be from a household below the national poverty line, compared to 18% of pupils who speak Kiswahili at home.

Figure 11 Background characteristics of pupils in standard three



Note: (1) Weighted estimates. (2) There is relatively large non-response on pupil age (approximately 16%), which may affect the age estimates. Moreover, age reporting by young children tends to be unreliable.

The qualitative research in nine schools in three case study districts (Kishapu in Shinyanga, Mpwapwa in Dodoma, and Uyui in Tabora) provides more insights into the life of some pupils in the programme treatment districts. These patterns are not necessarily typical of pupils across the districts, but give a flavour of the constraints that some groups of pupils face.

Getting up and going to bed

The timing of pupils' daily routine varies across the nine research sites, but some patterns emerged within each district. In Uyui, girls wake up earlier and have longer days than boys, whereas in Kishapu boys tend to wake up earlier or at the same time as girls and have longer days. Both boys and girls in Mpwapwa have long days, but girls' days tend to start earlier. In

³¹ This difference is statistically significant.

Mpwapwa and Uyui, boys talked about their bedtimes as flexible and open to change if they decide to go to town or watch a football match. By contrast, girls in these districts all had a precise time that they go to bed. Conversely, in Kishapu boys who had been tending cattle all morning before school and all afternoon after school went to bed at a precise time whereas the girls had more flexible bedtimes.

Pupil activities outside of school

Boys and girls in all three districts discuss their responsibility to dress themselves and get themselves ready for school. In Kishapu, boys work outside in the morning and afternoon tending cattle and in the afternoon wash their uniforms – they did not mention having time to play after school. In Mpwapwa and Uyui, boys have some duties in the home and collect water in the morning and afternoon and talked about having time to play after school. Girls in Kishapu and Uyui fetch water before and after school, but in Mpwapwa only after school. Girls tend to have different chores compared to boys, including cooking, washing clothes and doing dishes. In some cases in Uyui, girls also help out at the family farm. Similar to boys in Kishapu, girls in this district did not talk about play as part of their day.

Pupil chores in school and studying outside of school

When children were asked what they do at school in addition to attending classes, all of them talked about their responsibility for cleaning the school. Boys clean the toilets and girls clean the school building, toilets and/or fetch water for the teachers.

Boys in Kishapu and Uyui mentioned trying to study or read at home when there is time, whereas boys in Mpwapwa did not talk about reading or studying outside of school. In contrast, girls in all three districts talked about reading and studying at home.

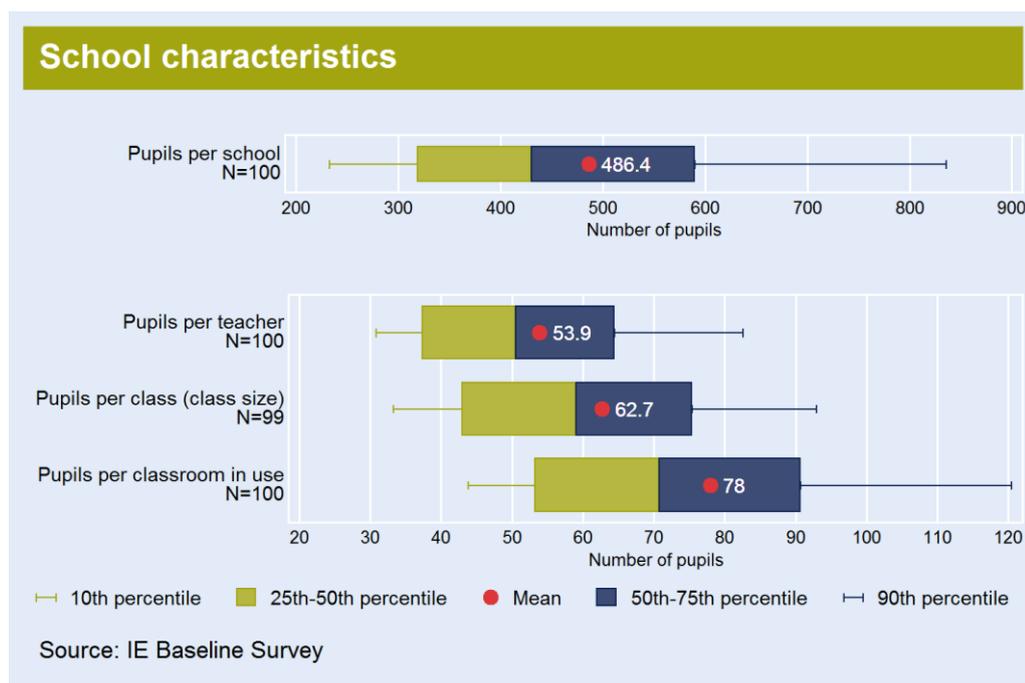
3.2 School characteristics

To provide contextual information on pupils' learning environment in schools, the IE baseline survey collected data through pupil head counts, school records checks, observations by enumerators and head teacher interviews.

The schools in the programme treatment districts vary in size, with a mean enrolment size of 486 pupils per school and with those in the first decile of schools having 233 or fewer pupils and those in the top decile of schools having 836 or more pupils (Figure 12).

Along with rising pupil enrolment nationally, the PTR in Tanzanian primary schools has been rising, with a SACMEQ survey of primary schools reporting an increase from 47 to 63 pupils per teacher between 2000 and 2007 (Ponera et al. 2011). The average PTR in the programme treatment schools was 54 pupils per teacher, which is higher than the recommended national benchmark of 40 pupils per teacher (MoEVT 2009a). There was substantial variation from 31 pupils per teacher (below the national benchmark) to 83 pupils per teacher (more than double the national benchmark). Per class (defined as a group of pupils taught together), there was an average of 63 pupils, with schools in the first decile having 33 or fewer pupils per class and schools in the ninth decile of schools having 93 or more pupils per class. On average, there were 78 pupils per classroom in use, in comparison to the recommended national benchmark of 45 pupils per classroom (MoEVT 2009a).

Figure 12 Background characteristics of schools



Note: (1) Weighted estimates.

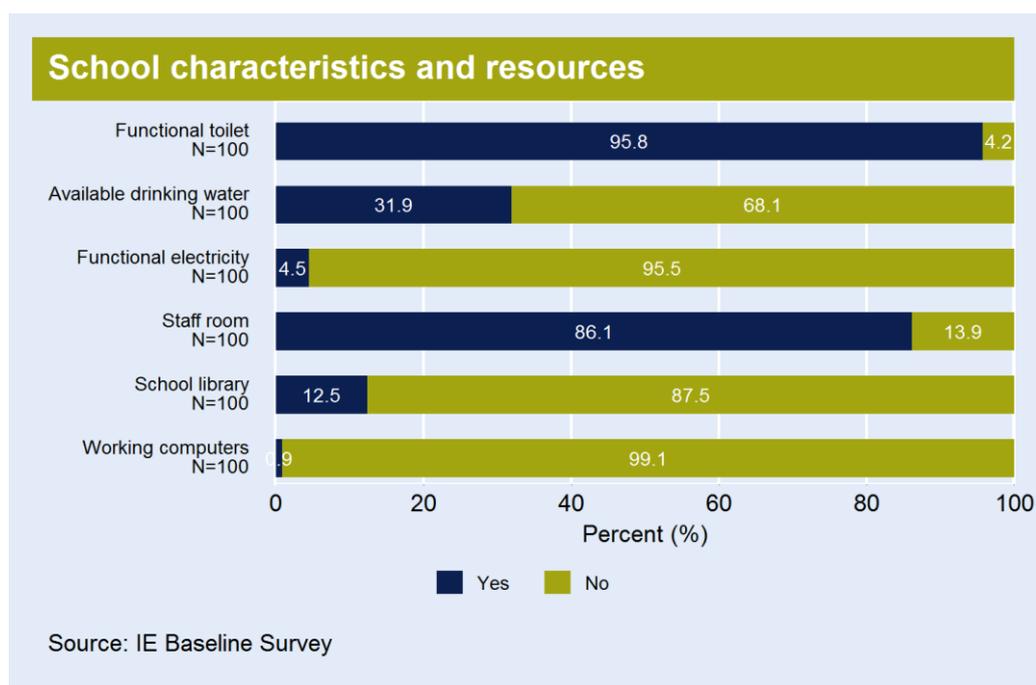
Another important aspect of education in the programme treatment districts is the use of shifts. Multiple shifts (sometimes referred to as double-shift schools), where a school caters to separate groups of pupils during the school day, are relatively common in the programme treatment districts, with 48% of schools having at least one class coming in for a second shift. From the IE baseline data it appears that the key constraint is the lack of classrooms, although subject specialisation means that in some cases there may also be insufficient numbers of teachers. Nevertheless, there are no significant differences in school size, PTRs, pupils per class or pupils per classroom for schools that use shifts and those that do not (results not shown).

Examining school infrastructure, the vast majority of schools (96%) had a functional toilet on the day of survey (Figure 13). However, there was on average 74 pupils per functioning toilet, and the number of pupils per functional toilet varied markedly across schools from 27 pupils or fewer per toilet for schools in the first decile and 131 pupils or more per toilet for schools in the ninth decile (not shown).³² This falls short of the government's recommended basic standard of at least one toilet per 20 girls and one toilet per 25 boys (MoEVT 2009a). Furthermore, only 32% of schools had drinking water available on the day of the survey compared to the recommended basic standard of at least 5 litres per pupil per day for each school. Although 86% of schools had a separate room for teachers to work in, other school resources were scarcer: only 12.5% of schools had a school library, an even smaller proportion (4.5%) had a functioning source of electricity on the day of the survey, and less than 1% had a working computer. Aside from staff rooms, these resources were found less frequently in the programme treatment districts compared to national averages (Ponera et al. 2011).³³

³² The IE baseline survey did not assess the cleanliness of toilets.

³³ In that survey, 34% of primary schools in mainland Tanzania had a class/school library, 45.1% had a water source, 17.2% had electricity and 4.0% had at least one computer. 86% had a separate staff room.

Figure 13 Overview of school infrastructure



Note: (1) Weighted estimates.

Thus, the primary schools in the programme treatment districts are on average far from meeting the government's recommended basic standards for primary education (MoEVT 2009a), reflecting the fact that the programme is targeting some of the most educationally disadvantaged regions in Tanzania.

3.3 Head teacher and teacher characteristics

The IE baseline survey also provides an overview of the characteristics of head teachers and standard one to three teachers in the programme treatment districts. A summary of these characteristics is presented in Table 6.

Table 6 Background characteristics of head teachers and teachers

Characteristic	Teachers		Head teachers	
	Mean estimate	N	Mean estimate	N
Personal characteristics				
Female (%)	55.2	329	15.6	100
Mean age (years)	39.6	329	43.4	100
Approaching retirement age 60 (%)	18.5	329	14.9	100
Work experience				
Time spent working as a teacher (years)	15.8	329	-	-
Time spent working as teacher less than five years (%)	32.1	329	-	-
Tenure				
Time spent working in current school (years)	8.3	329	3.9	100
Time in current school is one year or less (%)	18.7	329	37.5	100
Highest academic qualification				
Certificate or higher (%)	7.3	329	8.6	100

Form 6 (%)	3.1	329	19.6	100
Form 4 (%)	75.8	329	71.8	100
Primary schooling (%)	13.6	329	0.0	100
Other academic qualification (%)	0.2	329	0.0	100
Highest professional education qualification				
Diploma in education or higher (%)	1.9	328	9.3	100
Certificate in education (%)	94.2	328	89.6	100
Other professional qualification in education (%)	3.3	328	1.1	100
No professional qualification in education (%)	0.5	328	0.0	100
Sources: IE baseline survey (teacher interview). Notes: (1) Weighted estimates; (2) One teacher reported having a professional qualification but did not specify which one. Head teachers were not asked about their total years' experience, only their tenure in the current school; (3) Teachers/head teachers were identified as approaching retirement age if they were at least 56 years old at the time of the survey.				

The teacher estimates show that the average age of all who teach standards one to three in the treatment districts is 40 years compared with an average age of 43 years among head teachers. Just over half of teachers (55%) are female. By contrast, the gender distribution across head teachers in treatment districts is considerably less even – only 16% are female. This suggests that some gender bias in favour of men exists in regard to the promotion of teachers into head posts despite there being no significant differences in the qualifications or year of experience of female and male teachers.³⁴

The vast majority of teachers – over three-quarters (76%) – hold a Form 4 as their highest academic qualification, and roughly 10% have completed Form 6 or higher. However, there is a group of teachers, close to 14%, who have only completed primary schooling, i.e. the same level of schooling they are teaching. In addition, most teachers hold a professional education qualification: 94% of teachers have a Certificate in Education, 2% have a Diploma or Bachelor's degree in education, and about 3% have another professional education qualification.

Head teachers appear to be slightly more qualified than standard one to three teachers, both in terms of academic and professional educational qualifications. Among head teachers, 72% have a Form 4 qualification and roughly 28% have a Form 6 or higher. About 9% of head teachers hold a Diploma or Bachelor's degree in education but the large majority, at 90%, have a Certificate in Education. There are no head teachers who hold an academic qualification lower than Form 4.

The average teacher has spent about 16 years working as a teacher and eight years working in his/her current school but there are a few teachers with very long tenure and experience, so that these averages disguise low levels of experience and tenure among many teachers. The median level of experience is considerably lower at five years than the mean at nine years.

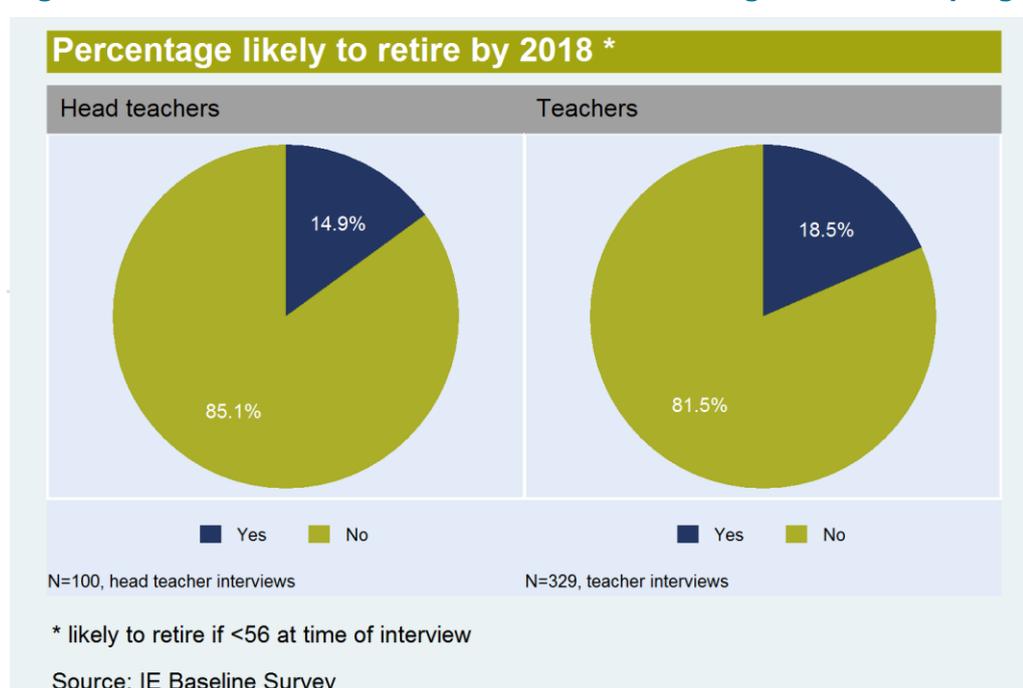
Considerable turnover of teachers in treatment districts is a notable finding of the IE baseline survey and this has potential implications for the EQUIP-T programme. Explained in the context of the programme, 50% of standard one to three teachers have only been in their current school for one year or more. In addition, about 19% of teachers are reported as having been teaching in their

³⁴ Mean estimates for indicators in Table 6 were generated for female and male teachers and tested for significant differences using a 5% confidence interval. All indicators by gender groupings, except the percentage retiring in the next four years, are not statistically significantly different across female and male teachers. The sample of female head teachers is too small to warrant reliable comparisons of characteristics of female versus male head teachers.

current school for one year or less.³⁵ A potential implication of the generally short tenure of teachers in programme treatment schools is that a significant proportion of teachers who are involved in the programme at the beginning may not be present by the end, potentially eroding the capacity built by the programme.³⁶

Figure 14 shows that roughly 19% of standard one to three teachers in treatment districts will reach the retirement age for Tanzanian civil servants of 60 years by 2018 (Ministry of Labour 2003). Among head teachers, 15% are likely to retire in the next four years; however, the recent turnover of head teachers in treatment schools far exceeds an amount that could be explained by natural attrition, i.e. retirement. As many as 38% of head teachers in treatment districts report that they have been in their current school for one year or less. As a result of this, average head teacher tenure in treatment districts is very low at only four years compared with eight years among teachers.

Figure 14 Head teacher and teacher retirement during the EQUIP-T programme



Notes: (1) Weighted estimates.

In briefings with a group of enumerators who collected baseline data, it was noted that a number of head teachers had been repositioned in schools from other districts and were not surprised by this finding. This anecdotal evidence suggests that, in some districts, the turnover of head teachers was due to 'policies' favouring the introduction of younger leaders into schools.

³⁵ Of the 66 of 329 teachers in treatment schools who were appointed to their current school in the past year, 44 have one year or less experience as a teacher, suggesting that the majority of new teachers appointed in schools are typically novice teachers.

³⁶ Some of this turnover may be attributable to retirement or natural attrition as well as the entry of novice teachers into the system. For example, almost one-third of standard one to three teachers in treatment districts have less than five years of teaching experience.

4 EQUIP-T Component 2: School leadership and management

One of the key five outputs of the EQUIP-T programme is enhanced SLM skills across programme treatment schools. This chapter presents baseline results in areas of SLM. These results are presented before the other programme components because effective SLM is a critical foundation for improving school performance. Studies for improving educational quality have identified that a school's baseline capacity or school functionality is critical to whether the school environment is able to absorb programme interventions (Hopkins, Harris and Jackson 1997).

The following discussion describes baseline management capacities across treatment schools in reference to the design of the EQUIP-T programme to enhance SLM through:

1. Implementing **leadership performance management systems**, guided by a newly developed leadership competency framework, to improve head teacher competency and performance accountability and inform professional development training;
2. Developing the **capacities of head teachers** through professional development training and peer support networks;
3. Strengthening **Whole School Planning and financial management** systems; and
4. Developing **school information systems and management processes**, guided by newly developed school quality frameworks, in order to measure quality and identify areas for improvement.

4.1 Head teacher leadership: Roles and responsibilities

As a starting point, it is important to understand at baseline how school leaders understand and enact their roles and responsibilities. Table 7 maps out findings from the qualitative study against five generic areas of school leadership responsibility.³⁷ Each area is briefly explained below, and referenced to the four competencies set out in the NSLCF that has been developed under the EQUIP-T programme (see Annex D2).³⁸

- *Shaping a vision of academic success for all pupils*: Establishing a school-wide vision of commitment to high standards and the success of all pupils. A fundamental part of this includes having high expectations and clear learning goals, as well as seeking to close the achievement gap between advantaged and less advantaged pupils. This is reflected in both the 'improved learning outcomes' and 'leading school improvement' NSLCF competencies.
- *Creating a climate hospitable to education*: Ensuring that the school is safe, orderly and supportive, and building a sense of school community. These roles are found in the 'school-based management' and 'leading the school community' NSLCF competencies.
- *Cultivating leadership in others*: Enabling stakeholders other than the head teacher to participate in decision-making processes, seeking out knowledge embedded within communities and within the school itself and providing leadership opportunities to others. This is related to roles under the 'leading the school community' NSLCF competency, but goes further to confer decision-making power upon others.

³⁷ These are drawn from Bush and Glover (2003).

³⁸ These are: (1) leading the school community; (2) improving learning outcomes; (3) leading school improvement; and (4) school-based management.

- *Improving instruction*: Includes encouraging continual professional learning and initiating discussions about instructional approaches within teams and to individual teachers. This is covered under the 'improving learning outcomes' NSLCF competency.
- *Managing people, data and processes*: Making good use of the available physical, human and financial resources, and using data as a means to not only pinpoint problems but to understand their nature and causes. This entails mapping out targets, implementing actions to meet targets, supporting others in their implementing roles, communicating plans to stakeholders, and monitoring results. These responsibilities are in the 'school-based management' NSLCF competency.

Within the case study schools, head teachers typically acknowledged their roles and responsibilities across two or three of the key areas of school leadership but reported carrying out only basic functions to fulfil these responsibilities.

The least acknowledged area of responsibility was *shaping a vision of academic success for all pupils*. There were only two head teachers who recognised academic leadership as part of their role and reported enacting these responsibilities. A belief that children within the school are not capable of achievement is a major reason why this responsibility is not commonly felt. Teachers within one such school discussed children from the school community as inherently lacking in the capacities and capabilities required to achieve educational success, and all educational failures were attributed to the pupils' families.

Creating a climate hospitable to education was in most cases reported only partially, through the head teacher's relationship with the community and ensuring the school had enough drinking water and other necessary supplies. *Cultivating leadership in others* was rarely mentioned. Only one school discussed a relationship with the SC and village committee that included participation in decision-making processes. Similarly, *improving instruction* was rarely considered part of the role of the head teacher.

There was some reported enactment of the role of *managing people, data and processes*, but this was most frequently related to monitoring teacher school attendance and teacher classroom attendance. All head teachers acknowledged teacher management as part of their responsibilities, but some admitted they do not fulfil these tasks either out of a lack of empowerment to enforce rules or a lack of willingness. Head teachers predominantly used examination results to assess the quality of teaching within their schools, but only one head teacher showed the ability to map out targets and implement actions to meet targets on the basis of data.

There were no apparent patterns of understanding or reported enactment of head teacher leadership responsibilities within districts, suggesting that district activities or personnel perhaps have little influence on school leadership behaviours currently.

Table 7 School leadership responsibilities: Qualitative findings from case study schools

Case study school	Shaping a vision of academic success for all pupils	Creating a climate hospitable to education	Cultivating leadership in others	Improving instruction	Managing people, data and processes
Uyui Upper	School culture is that some pupils are simply not capable of learning.	Ensures the school has the necessary basic infrastructure to operate.	Delegates leadership upwards to the WEC but discusses the school as needing 'control' over the community.	Improving instruction is viewed only in the basic sense of ensuring instruction is taking place.	Sees management as only involving managing teacher classroom attendance.
Uyui Typical	No vision for academic success.	Views the education environment as impossible and does not seek to effect change.	Delegates leadership upwards to the WEC.	Does not attempt to even ensure instruction is taking place. Does not feel empowered to do so.	Perceives him/herself as unable to manage as her salary is not higher than a teacher's.
Uyui Lower	Conceptualises his/her role as maintaining the standard of academic performance.	Eases the tensions between teachers and parents working as a conduit.	Has established ward-level INSET to share and learn from each other.	Conceptualises him/herself as an academic leader. Sets up a system for teachers to meet together to solve instructional issues.	Checks teachers' preparation schemes, lesson plans and observes teaching.
Kishapu Upper	No vision of academic success.	Has a good relationship with community but not with teachers.	Does not cultivate leadership in others.	Conceptualises the role mainly as managing teacher attendance.	Sees management as only involving managing teacher classroom attendance.
Kishapu Typical	No vision of academic success.	Has a good relationship with the community and recognises the importance of the SC.	Does not cultivate leadership in others.	Conceptualises the role mainly as managing teacher attendance.	Base level of managing the attendance book but does not check up on it and hold teachers accountable.
Kishapu Lower	No vision for academic success.	Perceives the school to be in opposition to the community. Perceives the pupils to be trying to 'destroy your aims'.	Does not cultivate leadership in others.	Focuses efforts on ensuring pupils attend, rather than quality of instruction.	Sees management as only involving managing teacher classroom attendance.
Mpwapwa Upper	No vision for academic success.	Does not enact leadership.	Does not enact leadership and therefore cannot cultivate it in others.	Does not feel that he/she can manage teacher attendance.	Does not manage, use data or monitor progress.
Mpwapwa Typical	Introduced remedial classes so that <i>all children</i> can read and write.	Respected by the community, is known as a hard worker. Described as a role model in the community.	Teachers talk about power being delegated.	Teachers talk about being encouraged to meet together and discuss practice.	Base level of managing the attendance book but does not check up on it and hold teachers accountable.
Mpwapwa Lower	No vision for academic success. The community fail to understand why the children are not learning and cannot find any explanations.	Conflict between the community and school over corporal punishment. The head teacher prefers it because it is easier for him than supervising a pupil given another punishment.	Expects parents to initiate communication with the school.	Does not feel able to manage teacher attendance.	Does not manage, use data or monitor progress.

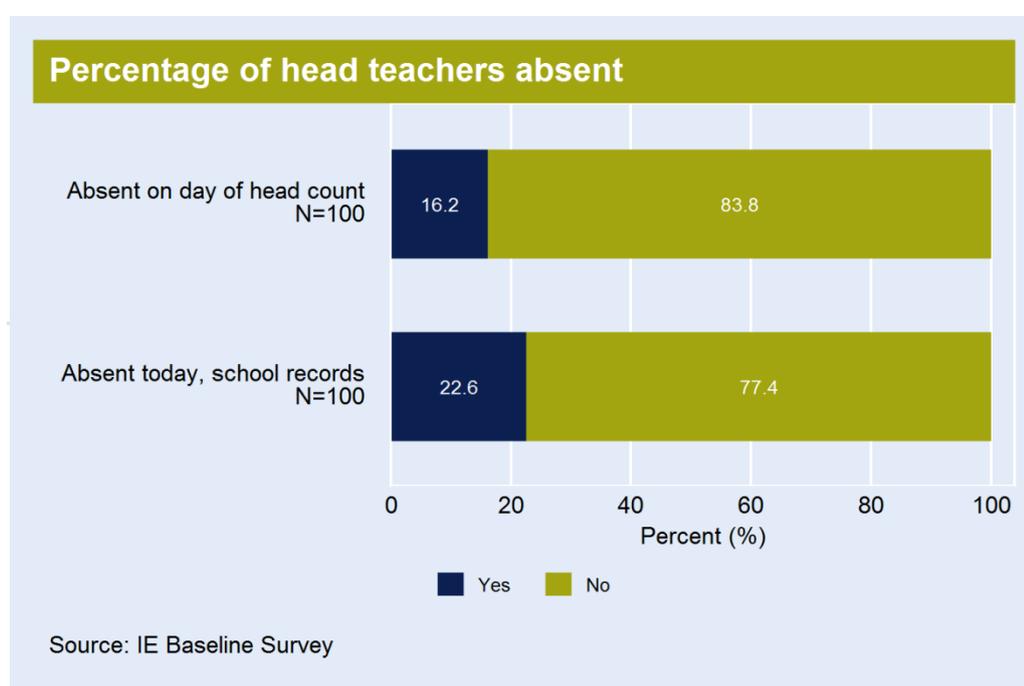
Source: IE qualitative study.

4.2 Head teacher absenteeism

A necessary condition for improving the efficacy with which head teachers lead and manage a school is to ensure that they are present at school. Absenteeism of head teachers was investigated in the IE baseline survey through a head count of all teachers including the head teacher, on the day of the survey, as well as checking school attendance records.

Figure 15 shows that head teacher absenteeism is very high. The head count on the day of the survey indicates that 16% of head teachers in the programme treatment schools were absent, while the school record data for the same day gives a figure of 23%.³⁹ The mismatch between the head count and school record figures is likely attributable to timing delays in filling in the school records and incomplete record keeping.⁴⁰

Figure 15 Absenteeism of head teacher



Notes: (1) Weighted estimates.

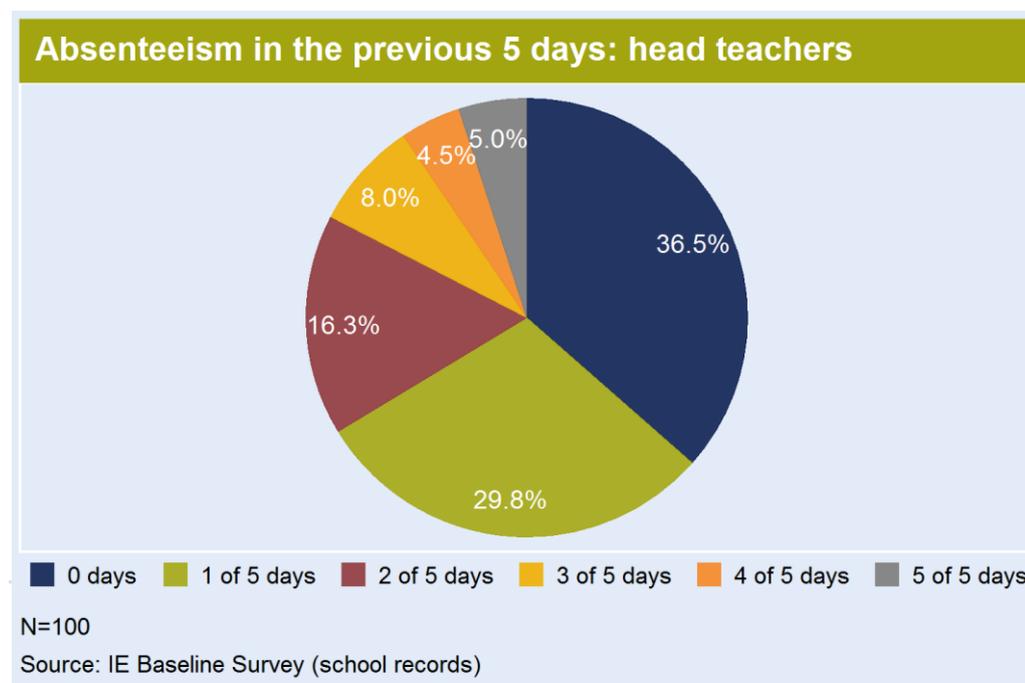
The absenteeism data from school records gives a picture of the frequency of absenteeism over the previous five days. On average, head teachers were absent for just over one day (1.3 days) in the previous five days. Figure 16 shows that 64% of head teachers were recorded as absent one or more days during the previous five days. About 30% of head teachers were absent for exactly one of the previous five days, 16% for two of the previous five days, and 18% for three or more days. Some of these absences may of course be 'beneficial' (for instance, during the qualitative research head teachers discussed the need to attend meetings at the municipal or district level as a reason for their absence) or 'non-beneficial', such as doing non-school work during school hours.

³⁹ This figure of 23% is consistent with the absenteeism figure found for teachers during announced visits for the 2010 School Delivery Indicators Survey in Tanzania (World Bank 2012).

⁴⁰ The head count was taken in the morning, possibly before the attendance record was completed in some cases. Incomplete attendance records may be a more general pattern in some schools, because school records of absenteeism from five days prior to the survey are consistently higher than the head count rate.

It might be expected that higher levels of head teacher absenteeism would be associated with higher levels of teacher absenteeism, lateness and absence from the classroom. However, the IE baseline survey found that this is not the case in the programme treatment schools.⁴¹ This may be partly explained by the qualitative research, which found that many head teachers feel powerless to manage their teachers or discipline them effectively. Indeed, many head teachers in the case study schools expressed the view that these types of responsibilities lie with the WEC.

Figure 16 Absenteeism of head teacher in previous five days



Notes: (1) Weighted estimates.

4.3 Head teacher capacity development: in-service teaching training and support networks

Few head teachers have been exposed to INSET on school leadership or management in recent years, as Figure 17 reveals. Only 11% of head teachers reported that they had attended any INSET on SLM in the last two school years (2012 and 2013).

In terms of support networks, head teachers appear to be reasonably satisfied with current arrangements from the WEC and, to a lesser extent, from the SC. Heads overall are least satisfied with the support they and their school receives from the community. Given a 1 to 5 scale with 1 for 'very poor' and 5 for 'very good', the vast majority of head teachers (80%) rated WEC support as good or very good, while over just over half (52%) rated SC support equally positively. However, only 20% felt the same way about community support (see chapters 6 (Section 6.3.1) and 7 (Section 7.4.1) for more details).

⁴¹ There are no significant differences in teacher absenteeism, lateness or classroom absenteeism when disaggregating the sample by levels of head teacher absenteeism (measured using the headcount and school records) using a 90% confidence interval (results not shown).

Figure 17 INSET on development planning and management for head teachers

Notes: (1) Weighted estimates.

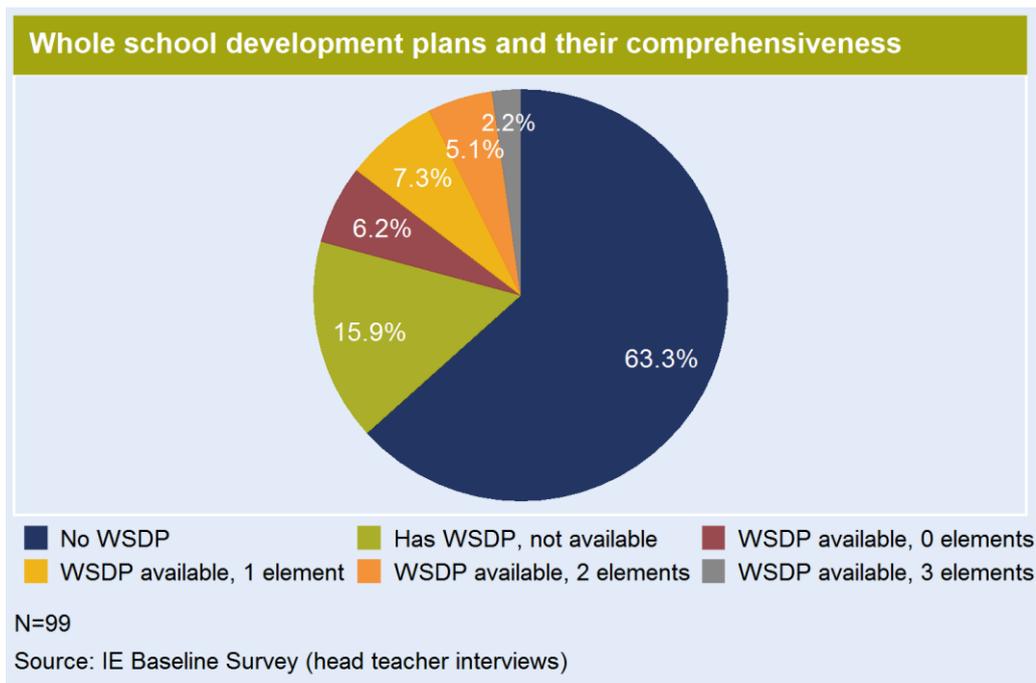
4.4 School management: Whole school planning and financial management systems

At the national level, resources, processes and guidelines are firmly in place for the development and implementation of WSDPs.⁴² Despite the importance placed on WSDPs, however, only 37% of head teachers in programme treatment schools reported that the school had a WSDP and only 21% of head teachers were able to show our team the physical document (Figure 18).

Where WSDPs were available, their comprehensiveness was very limited. Specifically, three elements – a) a budget; b) teaching and learning objectives; and c) baseline data and targets – are considered important features of the plans, but only 2% of the schools had WSDPs that included all three elements. Some 12% of available WSDPs included one or two of the elements and 6% of WSDPs contained none of these three core elements. This clearly suggests that head teachers need training on the importance of and on how to develop WSDPs.

⁴² Cambridge Education (2014b) (Annex XVI: Component 2: Enhancing School Leadership and Management), p. 5.

Figure 18 WSDPs

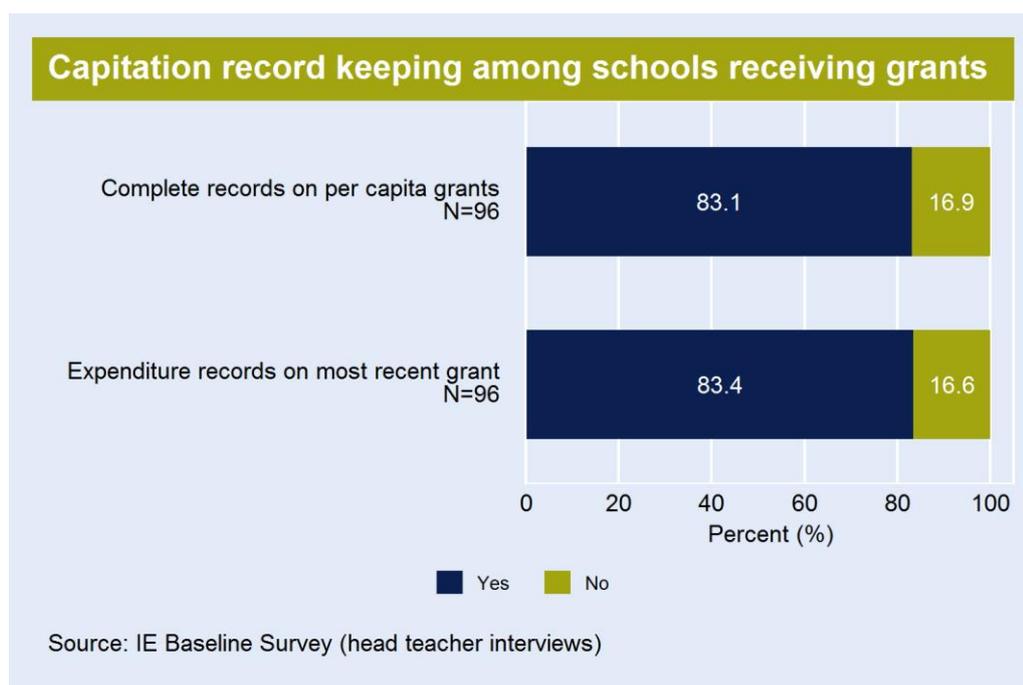


Notes: (1) Weighted estimates.

In contrast to the low proportion of schools that had WSDPs, and especially comprehensive ones, the majority of head teachers were able to show evidence that they are managing their capitation grant records to some extent. Over 83% of head teachers in schools receiving grants were identified by the enumerator as having complete records on per capita grants in 2012 and 2013. Similarly, 83% of head teachers were able to provide expenditure records related to the most recent capita grant received (Figure 19).⁴³

However, record keeping alone is a crude indicator of financial management and in particular good financial management. A possible proxy for whether school financial management systems are comprehensive is to identify the proportion of schools that have a budget in their WSDP, as good financial management requires spending in line with established budgets. Among the programme treatment schools only 9% had a WSDP with a budget. It should be acknowledged that the ability to implement sound financial systems and spend in line with budgets may be compromised in an environment where capitation funds are disbursed irregularly. Schools are subject to considerable financial uncertainty in terms of both the amounts of funds that will be received each year and when they are received, as noted by other studies (e.g. UWEZO 2013).

⁴³ However, these results should be treated with some caution as different enumerators may have different views of what constitutes complete capitation records or expenditure records despite receiving training.

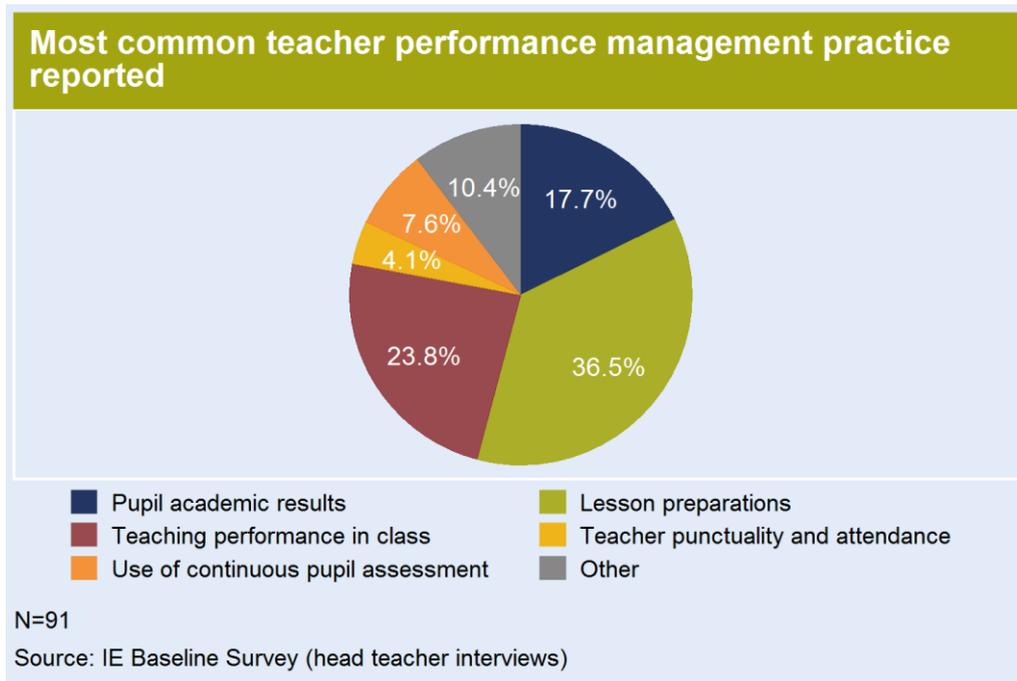
Figure 19 Capitation grant record keeping

Notes: (1) Weighted estimates.

4.5 School management and information systems: Teacher performance management

A core element of any school quality standards or framework, as well as an associated school information system, is teacher performance, since it is so closely related to pupil performance. The IE baseline research provides insights into current processes of monitoring and feedback of teacher performance

In the IE baseline survey, head teachers were asked about what was the most important method they used to assess teacher performance. Figure 20 summarises their responses. Observing teachers' lesson preparations is the most commonly cited assessment practice – 37% of head teachers selected this option. Observing teaching performance in class and using pupil academic results were the next most common assessment practices as reported by 24% and 18% of head teachers respectively. The least common factor chosen was teacher punctuality and attendance. Only 4% of head teachers chose this assessment practice despite the high levels of teacher absenteeism reported in schools and particularly the very high levels of absence from the classroom (see chapter five).

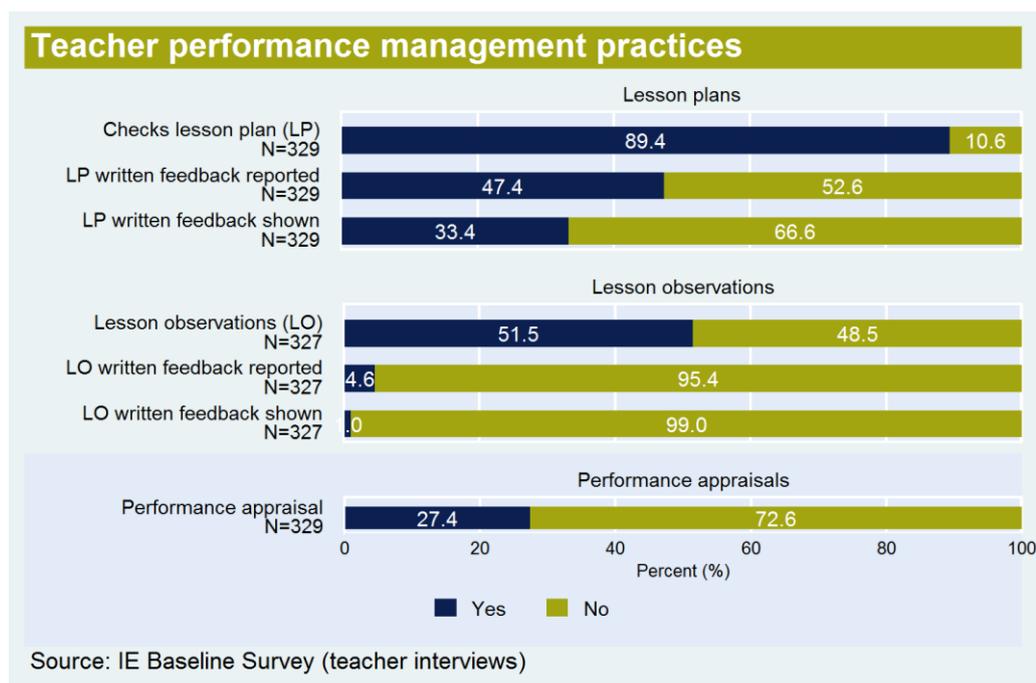
Figure 20 Most common teacher performance management practice reported by head teachers

Notes: (1) Weighted estimates.

In addition to asking head teachers about their teacher management performance practices, teachers of standards one to three were asked about the teacher assessment practices of their head teachers. These results, summarised in Figure 21 are particularly interesting since teachers' responses about school managers' behaviour are generally considered more reliable than reporting by school managers on their own behaviour (Hallinger and Heck 1996).

Consistent with lesson preparation being identified as the most common factor reported by head teachers in assessing teacher performance, almost 90% of standard one to three teachers responded that the head teacher or academic master checked their lesson plans at least once in the last 30 days. However, feedback in the form of written comments on lesson plans was not common. Only 47% of interviewed teachers reported that their head teachers had checked their lesson plans and provided written feedback in the last 30 days. Furthermore, only one-third of interviewed teachers were able to actually show this written feedback on lesson plans.

Observing teaching in the classroom is a less common practice than checking lesson plans. Just over half of the teachers reported that the head teacher had observed at least one of their lessons in the past 30 days and written feedback on these observations was scarce. Only 5% of the interviewed teachers reported that lesson observations were conducted *and* they had received written feedback on the lesson.

Figure 21 Teacher performance management practices reported by teachers

Notes: (1) Weighted estimates.

There is a clear mismatch between the extent to which head teachers engage in teacher assessment through lesson observations or checking lesson plans and providing written feedback on these assessments. In addition to the lack of formal written feedback, formal verbal feedback on assessment appears to be limited in most schools too. Nearly three-quarters of teachers said that they had not had a one-on-one meeting with the head teacher, assistant head teacher or academic master to discuss their performance or professional development needs in the previous school year (2013).⁴⁴

Staff meetings also present opportunities for providing regular feedback to teachers. In the majority of schools, these meetings occur once every month or two, according to both head teachers and teachers (Figure 22).

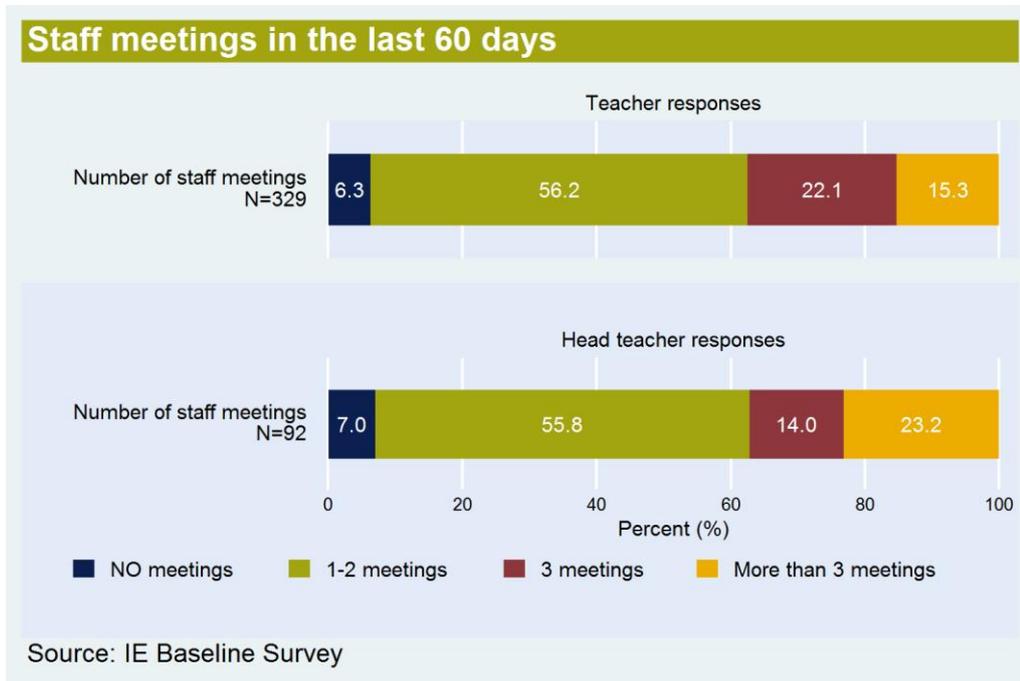
As well as the lack of concrete feedback for teachers, there is limited evidence of head teachers affirming teachers' good teaching performance. Nearly two-thirds of head teachers said that they did not reward teachers who performed well, either verbally or materially (Figure 23).⁴⁵ Where they were given, rewards were mostly financial or in the form of verbal recognition. Overall, just 15% of head teachers in treatment districts are verbally recognising good teaching performance. Given this situation, it is not surprising that when teachers were asked to rate on a scale of one (very poor) to five (very good) the level of support provided by the school to improve their teaching, only 12% chose a level five rating of 'very good', as is shown in Figure 23. There is some evidence that

⁴⁴ The mean results of Figure 21 have also been disaggregated by the gender of the teacher to identify if teacher performance management practices differ depending on whether the teachers responding are male or female. These results are not reported here as there appear to be no statistically significant differences in teacher management practices used depending on the gender of the teacher in question. A logical question that follows is whether teacher management practices differ depending on the gender of head teachers in relation to that of their teachers. Unfortunately, there are too few female head teachers, and especially male teachers in female-led schools, to allow for reliable tests of significant differences by head teacher–teacher gender combinations.

⁴⁵ There are 10 treatment schools where the head teacher did not respond to the question on teacher rewards, reducing the available treatment sample for the calculation to 90 head teachers.

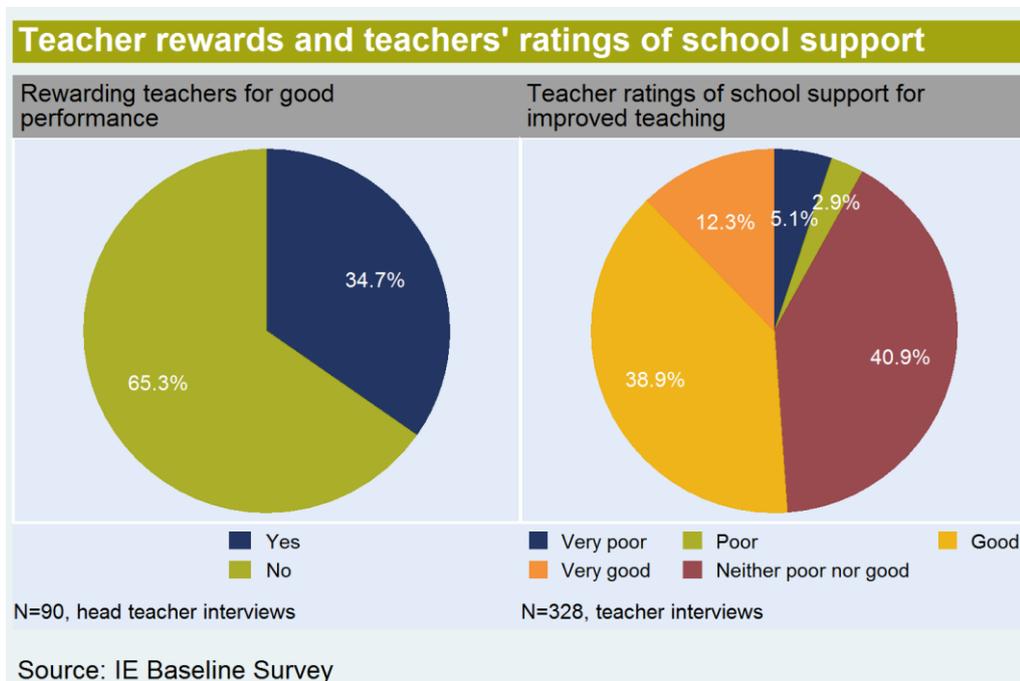
teachers' perceived levels of support is higher where head teachers provide written feedback on lesson plans or have staff meetings more frequently.

Figure 22 Staff meetings held in the last 60 days



Notes: (1) Weighted estimates.

Figure 23 Rewards for good teacher performance and teacher ratings of school support



Notes: (1) Weighted estimates.

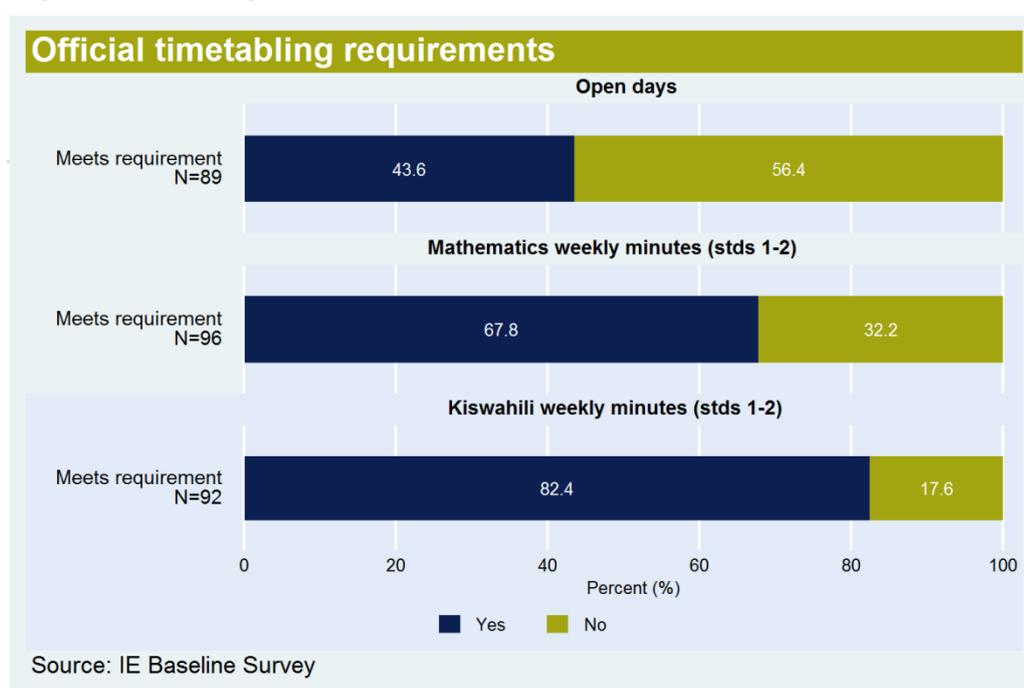
4.6 School management and information systems: Protecting instructional time

Another core element of school quality standards is the amount of instructional time that pupils experience, which directly influences the amount of learning that can take place. This means that, among other things, a key priority for instructional leaders is protecting an adequate amount of available teaching time (Robinson, Lloyd and Rowe 2008).

At a most basic level, protecting instructional time requires ensuring that schools are open for a minimum number of days and that timetables are structured to allow for enough teaching hours. In Tanzania, schools are required to be open for at least 194 days in a year.⁴⁶

The school records checked as part of the IE baseline survey suggest that the majority of schools are coming close to the official requirement. Although only 44% of schools met the requirement exactly (see the top bar in Figure 24), 72% of schools were open for at least 188 days, representing one school week less than the official 194-day requirement. Some schools exceed the requirement by some margin, and the average number of days that schools were open was 193 days, only one day short of the official requirement.

Figure 24 Meeting instructional time requirements



Notes: (1) Weighted estimates.

In addition to stipulating the number of school days per week, requirements are set for the amount of instructional time per week in specific subjects. To investigate whether head teachers are currently following these guidelines and timetabling sufficient instructional time, the mathematics and Kiswahili timetables for standards one and two pupils are used as an example.

⁴⁶ The official requirements were provided by the Director of Primary Education, as per the Education Circular of 2001. The number of days does not vary per year or region.

Each school week, standards one to two pupils are required to be exposed to at least 210 minutes of mathematics and 180 minutes of Kiswahili.⁴⁷ Schools in programme treatment districts currently exceed this on average, according to the timetable information collected in the survey. Average timetabled minutes per week for mathematics and Kiswahili are 219 and 207 minutes respectively.⁴⁸ While the vast majority of schools are meeting or exceeding the timetabling requirement, Figure 24 indicates that about one-third of schools are not providing sufficient mathematics periods and nearly 20% of schools do not meet Kiswahili timetabling requirements for standards one and two. The qualitative research finds explanations for this in terms of infrastructure and teacher number constraints. Some head teachers reported that when two classes have to share a classroom, instructional time is reduced, while in other cases the two classes are taught together resulting in large class sizes. The baseline survey results find that classrooms are a particularly acute constraint. As reported in Chapter 3 (Section 3.2), the pupil-to-classroom ratio is almost double the norm for class size prescribed by official guidelines, and nearly half of schools have at least one class that comes in a different shift.

While instructional time may be formally timetabled and schools are open for instruction, protecting instructional time ultimately depends on whether teaching is actually taking place in the classroom. This requires teachers to be in school and to attend their timetabled lessons. It is these factors that turn out to have a very large effect in reducing instructional time for pupils in treatment schools. These results are discussed in the next chapter (see Chapter 5, Section 5.3.1).

5 EQUIP-T Component 1: Teacher professional capacity, performance, motivation and morale

To achieve the EQUIP-T programme intermediate outcome of enhanced professional capacity and performance of teachers, the programme will support four core activities:

1. Development of a TCF expected to result in ‘clear, transparent criteria that links and can be used for accreditation, INSET and performance management’ (Cambridge Education 2014a, p. 2). The TCF will cover five areas: i) professionalism, ii) subject knowledge, iii) planning skills, iv) classroom teaching skills, and v) assessment skills.
2. Teacher professional development through school-based INSET intended to strengthen teacher subject knowledge and pedagogy.⁴⁹
3. Development of a teacher performance management system that will be ‘linked to concrete pathways for career development and other incentives to reinforce and encourage outstanding performance’ (Cambridge Education 2014a, p. 15).
4. Improvement of teachers’ morale through a Teacher Morale Toolkit anticipated to improve teachers’ ‘attitude and behaviour in school’ (Cambridge Education 2014a, p. 2).

The IE includes quantitative indicators for teacher subject knowledge, classroom teaching skills and assessment skills to allow measurement of (any) programme impact over time in these areas. This is supported by qualitative research that explores the reasons for some of the key quantitative

⁴⁷ See Government of Tanzania (2007), *Mtaala wa elimu ya msingi*, p. 19, section 5.2.

⁴⁸ There were some major outliers for minutes spent on Kiswahili per week. Two treatment schools for which 600 minutes are recorded are excluded.

⁴⁹ This will start with teachers of lower primary school mathematics and Kiswahili. Later the programme will also include upper primary school teachers and cover more subjects and more advanced content (Correspondence with EQUIP-T MA, March 2014).

baseline results. Selected issues in the areas of teacher professionalism are explored through the qualitative research. The mixed methods baseline results for these areas are presented and discussed in the remainder of this chapter.

To set the context in terms of INSET, in the programme treatment area, the levels of INSET received by teachers are currently very low – just 8.4% of teachers reported having attended any form of in-service training in 2012–2013.⁵⁰ This is roughly consistent with the EQUIP-T programme’s finding for the five programme regions,⁵¹ according to which only 4% of teachers reported having received any INSET during the last three years (Cambridge Education 2014a).

5.1 Teacher professional capacity: subject knowledge

The IE baseline survey provides baseline estimates of teachers’ subject knowledge in mathematics and Kiswahili. These are based on a TDNA⁵² instrument, which contains questions linked to the primary school curriculum.⁵³ The TDNA is in the form of a mock pupil test that teachers mark, indirectly providing information on their subject knowledge, which can be used to help design teacher support programmes. The TDNA was developed by OPM with a national team of experts and this is the first time a TDNA has been used in Tanzania.⁵⁴ For details on instrument development see Volume II, Section 3.3.

Before discussing the results, it is important to mention that the teachers assessed are those that were present at school on the day of the survey. The exclusion of absent teachers may affect the results slightly, but is unlikely to change the main conclusions drawn given the scale of teacher absence (reported later in this chapter).

5.1.1 Teacher subject knowledge in Kiswahili

The content of the Kiswahili TDNA is relatively evenly balanced between questions from lower and upper-primary levels in the curriculum: 21 standard one to four level questions and 22 standard five to seven level questions. It covers a range of core curriculum topics as follows:⁵⁵ reading comprehension (10 questions), grammar (five questions), synonyms (five questions), proverbs (five questions), direct and indirect speech (three questions), tenses (five questions) and punctuation (nine questions). The TDNA Kiswahili score is defined as the number of questions answered correctly as a percentage of the total number of questions for each teacher (person score) (for the full definition see Volume II, Annex F).

⁵⁰ Among this small group of teachers, the average days spent on in-service training over the period was five days. This includes donor-funded programmes such as BRN and the Student Teacher Enrichment Programme and government-funded programmes such as the MWAKEM/MoEVT in-service training in English, mathematics and pedagogy, and Mafunzo ya Ualimu Kazini.

⁵¹ This excludes Lindi and Mara regions, which will join the programme at a later stage.

⁵² The IE baseline TDNA instrument is based on the format used by a TDNA instrument developed by Dr David Johnson for DFID Nigeria.

⁵³ If teachers left questions blank, this ‘non-response’ was treated as incorrect in the teacher score results (person scores). Since non-response rates were fairly high for some of the questions appearing later in the TDNA, a table of scores for each question (item scores) can be found in Volume II, Annex H. The analysis presented here excludes non-responses if they appear after the last question attempted.

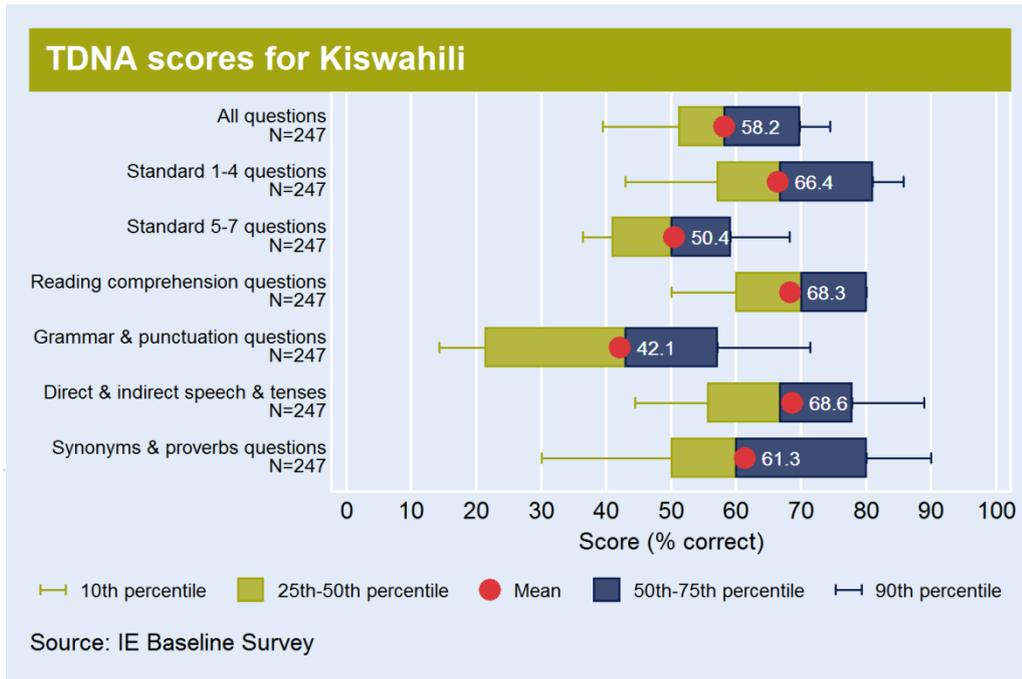
⁵⁴ The instrument used for the SDIS was in the form of a regular test administered to teachers (World Bank 2012).

⁵⁵ The Kiswahili TDNA was developed by a team of Tanzanian educators, led by a Professor of Kiswahili from the University of Dar es Salam. The team selected topics from the primary curriculum based on their expert view of the core topics. The instrument was pre-tested twice, partly to try to avoid floor and ceiling effects.

Teachers scored 58% on average in the Kiswahili TDNA, with a wide spread of performance between the bottom and top performing 10% of teachers of almost 35 percentage points. Figure 25 shows these overall results together with results grouped by curriculum level and topic.⁵⁶

Teachers performed considerably better on the questions drawn from the standard one to four levels in the curriculum compared with those from standards five to seven. This is perhaps not surprising given that the sampled teachers are those who teach pupils from standards one to three and are thus likely to be more familiar with curriculum material from the lower curriculum levels from their everyday teaching.

Figure 25 TDNA results in Kiswahili



Notes: (1) Weighted estimates

Looking at the results by topic, teachers performed relatively poorly on grammar and punctuation questions, scoring 42% on average compared with more than 60% on the other topics. The punctuation questions were located at the end of the test and so it is relevant to ask if non-response (which has been treated as incorrect in the scores presented above) rather than question difficulty partly explains this poorer topic performance. In fact, non-response rates for the punctuation questions were close to 20%, which is far higher than for other questions (see Volume II, Annex H); however, even taking this into account the item scores are still very low for many of these questions, suggesting that the majority of teachers who attempted them found at least some of them very challenging.

5.1.2 Teacher subject knowledge in mathematics

The contents of the mathematics TDNA is focused on questions from the upper standards of the primary curriculum: three standard one to three level questions, eight standard four to five level questions and 25 standard six to seven level questions. This is because some of the questions

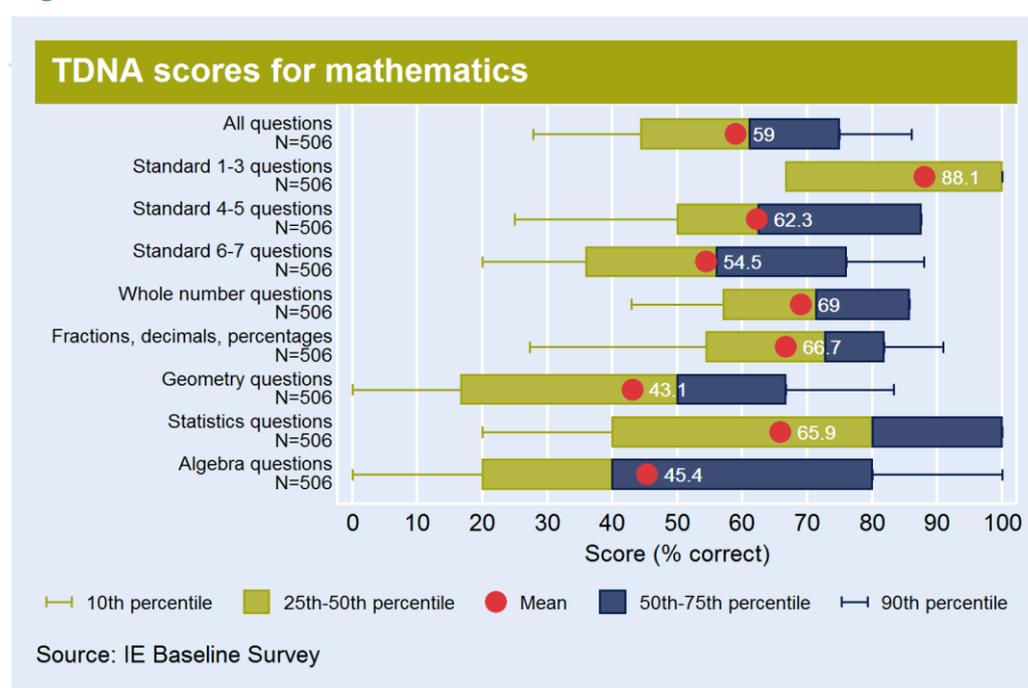
⁵⁶ Younger teachers appear to score higher on the TDNA Kiswahili than older teachers and this difference is statistically significantly different. However, breaking the sample into two age groups yields two very small sub-samples with implications for statistical inference.

based on the standard one to three curriculum levels were removed after pre-test results showed that the vast majority of pre-tested teachers found these questions very easy, which suggests that generally teacher subject knowledge with respect to the lower curriculum levels is good.⁵⁷ The TDNA mathematics covers eight core topics^{58, 59}: whole numbers (seven questions), fractions (four questions), decimals (four questions), percentages (three questions), measurement (two questions), geometry (six questions), statistics (five questions) and algebra (five questions). The TDNA mathematics score is defined similarly to the TDNA Kiswahili score: the number of questions answered correctly as a percentage of the total number of questions (for the full definition see Volume II, Annex F).

The mathematics TDNA was given to two groups of teachers: those who teach mathematics to lower primary (standards one to three) and those who teach mathematics to upper-primary (standards four to seven).⁶⁰

Teachers scored 59% on average in the mathematics TDNA, but this masks a wide range of teacher performance. The lowest performing 10% of teachers scored 28% or less, while the top performing 10% of teachers scored 86% or more. Figure 26 shows these overall results together with results grouped by curriculum level and topic.

Figure 26 TDNA results in mathematics



Notes: (1) Weighted estimates; (2) The two questions on the measurement topic were excluded.

⁵⁷ From the perspective of the IE, it is important to avoid ceiling effects or it will be difficult to detect change during IE follow-up survey rounds.

⁵⁸ The mathematics TDNA was developed by a team of Tanzanian educators led by a Professor of Mathematics from the University of Dar es Salam. The team selected topics from the primary curriculum based on their expert view of the core topics.

⁵⁹ Note that the estimates presented in this Baseline Report differ slightly from the preliminary results contained in the IE Baseline Summary Report (OPM 2014b). This is due to a partial remarking exercise to ensure consistency in the marking of questions that required mathematical units to be specified as part of the answer.

⁶⁰ About 5% of the teacher sample teach mathematics to both lower and upper primary standards.

Teachers performed well on the questions drawn from the lower standards in the curriculum, scoring an average of 88%, helped by (at least) the top performing 10% of teachers scoring full marks. The average score drops sharply to 62% for the middle standard questions, and drops again to 55% for standard six to seven level questions. Perhaps of even greater note is the vast range of the results on the middle and upper standard questions. The gap between the lowest and highest performing 10% of teachers is more than 60 percentage points in both cases. This implies that there are groups of teachers with very strong subject knowledge and also groups of teachers with very weak subject knowledge

Considering the results by topic, teachers performed strongly on whole numbers, fractions, decimals and percentages, and statistics, scoring two-thirds or more of questions correctly on average. The weakest teacher performance was in geometry and algebra, which were standard six and seven level questions. This is in line the EQUIP-T programme needs assessment in the five programme regions for which teachers stated that they found geometry and algebra particularly difficult subjects (Cambridge Education 2014b⁶¹). However, the difference in performance by topic is likely to be only partly related to how difficult teachers found the questions. One of the reasons for relatively poorer scores in algebra, for example, may be the position of these questions at the end of the TDNA instrument, so that some teachers may simply have run out of time for these questions. The average non-response rate for algebra questions is 17%, compared with less than 6% for other questions. Annex H in Volume II contains scores for each of the questions, adjusted for non-responses.

Teachers who teach mathematics to upper-primary school pupils demonstrated much stronger subject knowledge than their colleagues who teach lower primary school pupils.⁶² Table 8 reveals significant and large differences in scores between the two groups, with the standard four to seven group of teachers consistently outperforming the standard one to three group of teachers. For all the questions, the difference in average score between the two groups is 14 percentage points. The largest relative gap in performance by far is for geometry questions, where the average score for the upper-primary school group of teachers was 45% greater than the average score for the lower primary group of teachers.

Table 8 TDNA results in mathematics by standards taught by teachers¹

Indicator	Teachers of std 1–3	N (teachers)	Teachers of std 4–7	N (teachers)
	Mean score (%)	51.8***	246	65.6
Mean scores by curriculum level:				
Mean score std 1–3 qns (%)	84.1***	246	91.5	282
Mean score std 4–5 qns (%)	54.1***	246	70.1	282
Mean score std 6–7 qns (%)	47.2***	246	61.0	282
Questions correct by topic ² :				
Mean score whole number qns (%)	62.2***	246	75.3	282
Mean score fractions, decimals and percentages qns (%)	60.1***	246	73.1	282

⁶¹ See p1 EQUIP-Tanzania Component 1 Teacher Performance". Background Paper (part of Cambridge Education (2014b) EQUIP-Tanzania Programme Inception Report)

⁶² Younger teachers also appear to score higher on the TDNA mathematics than older teachers and this difference is statistically significantly different. However, breaking the sample into two age groups yields two very small sub-samples with implications for statistical inference

Mean score geometry qns (%)	35.2***	246	51.0	282
Mean score statistics qns (%)	59.3***	246	71.7	282
Mean score algebra qns (%)	37.8***	246	50.9	282

Sources: IE baseline survey (TDNA mathematics). Notes: 1) Weighted estimates 2) Statistically significant differences between groups are marked with asterisks on the estimate for the first category in the group: *significant at the 10% level; **significant at the 5% level; ***significant at the 1% level; 3) The two questions on the measurement topic were excluded.

The qualitative research found that teachers felt confident in their knowledge of the subject in which they specialised, but many teachers are now teaching subject areas in which they have not specialised.

In the case study schools, the biggest gap in ‘overall’ teacher knowledge, which comprises curriculum, subject and pedagogic knowledge, was in curriculum knowledge. This refers to knowing what should be taught to a group of pupils, knowledge of the national syllabus and examinations content, and understanding of the school and grade level planning documents. Due to the lack of teacher training, teachers were unaware of the content of the new syllabus, did not know what subject matter should be taught to which grade level pupil, and were unclear how various textbooks relate to the curriculum they are required to teach. Although teachers may have sufficient subject knowledge, without curriculum knowledge they will be unable to adequately prepare pupils for examinations. Many teachers also questioned the introduction of new curriculum subjects such as Information and Communication Technology (ICT), as they felt this type of subject did not take the rural context into account – indeed, many teachers found themselves covering topics like how to turn a computer on and off in the absence of a computer.

5.2 Teacher performance: Pedagogy

The IE baseline survey provides baseline estimates of teacher classroom practices. The estimates are based on standard two Kiswahili and mathematics lesson observations during which enumerators carried out two types of observation: (i) mapping of teacher–pupil interactions by gender and classroom space, and (ii) recording demonstration by teachers of a set of selected teaching behaviours.

5.2.1 Gender balance in teacher interaction with pupils

During the lesson observation, when the teacher interacted with a pupil individually, enumerators were asked to note the gender of the pupil.⁶³ From these observation records, teachers’ interactions with pupils could be categorised according to whether they are gender balanced, i.e. whether teachers interacted with pupils proportionally to their presence in the classroom (for the full definition see Volume II, Annex F).

Among all schools, just over half (54%) of the observed teacher interactions with pupils were gender balanced, while 30% of teachers interacted more with boys than girls and 16% of teachers interacted more with girls than with boys (see Figure 27). This suggests that gender balance in teaching practice is not an acute problem, although there is obviously scope for improvement, and girls are more disadvantaged than boys. There are some indications that among the 46% of

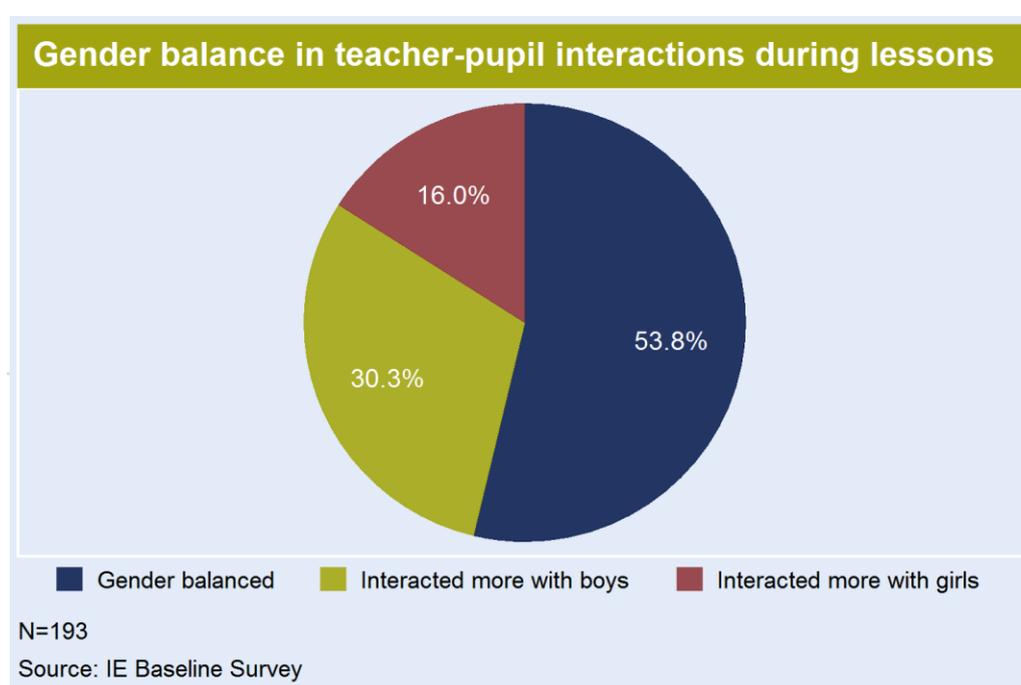
⁶³ Enumerators were asked not to include classroom management type interactions (e.g. disciplining or if a teacher asked a pupil to distribute materials). If a teacher interacted with the same pupil multiple times during the lesson, enumerators were asked to record each separate interaction.

teachers who did not demonstrate gender balance in their pupil interactions, male teachers were more likely to show bias towards boys. However, these differences were not found to be statistically significant.⁶⁴

5.2.2 Spatial balance in teacher interaction with pupils

Another measure of inclusive teaching is whether a teacher interacts with pupils across the whole classroom space or focuses on only certain areas, for instance only on pupils sitting at the front of the classroom. This information was collected using a classroom mapping instrument that divides the classroom into six approximately equally sized areas, with enumerators recording the number of teacher–pupil interactions in each of the six classroom areas (for the full definition see Volume II, Annex F).

Figure 27 Gender and spatial balance in teacher interaction with pupils

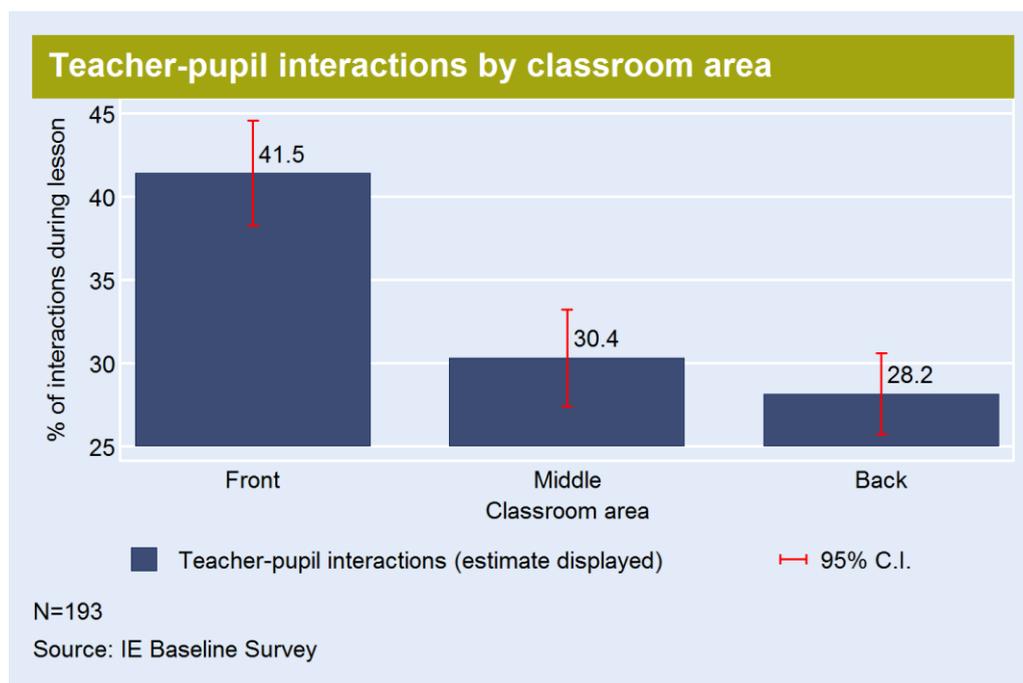


Notes: (1) Weighted estimates.

In 59% of the lessons observed, teachers were spatially inclusive, engaging with at least one pupil from all six areas of the classroom. Of all recorded teacher interactions with pupils, interactions were most common with pupils seated at the front of the classroom (42% of all interactions) compared to in the middle two areas (30%) and the back two areas (28%). The differences in the number of interactions by classroom area between the front two areas and middle two areas and between the front two areas and back two areas respectively are strongly significant (see Figure 28). This shows that teachers interact significantly less with pupils seated towards the middle and back of the classroom compared to pupils seated at the front.

In seven of the case study schools, the boys generally sat at the back of the classroom and girls at the front, but in two schools the opposite was true, with boys generally sitting at the front and girls at the back.

⁶⁴ Bivariate regressions were run to test bias toward boys and girls respectively by female and male teachers (results not shown here).

Figure 28 Distribution of teacher–pupil interactions across the classroom

Notes: (1) Weighted estimates.

5.2.3 Teacher performance: Teaching behaviours

In the case study schools, most teachers were unable to identify pedagogy as a knowledge required of a teacher and tended to focus on subject and curriculum knowledge during the FGDs. Only a few teachers displayed elementary knowledge of pedagogy and discussed child-centred teaching and participatory methods. This finding is consistent with the quantitative baseline results on pedagogy presented below. Some teachers expressed frustration that they could not understand why pupils in their classrooms do not understand particular topics or cannot display certain skills, not linking this to pedagogy and often concluding that the pupils simply lack the capacity to learn.

The IE baseline survey measures the use of different teaching behaviours in the classroom with an instrument based on tools used to evaluate a school-based INSET programme in Tanzania (Hardman and Dachi 2012).⁶⁵ The lesson observation instrument captures 14 effective teacher behaviours/skills, which may occur at the start of a lesson (the first 5 minutes), the middle of the lesson, or the end of a lesson (the final 5–10 minutes).

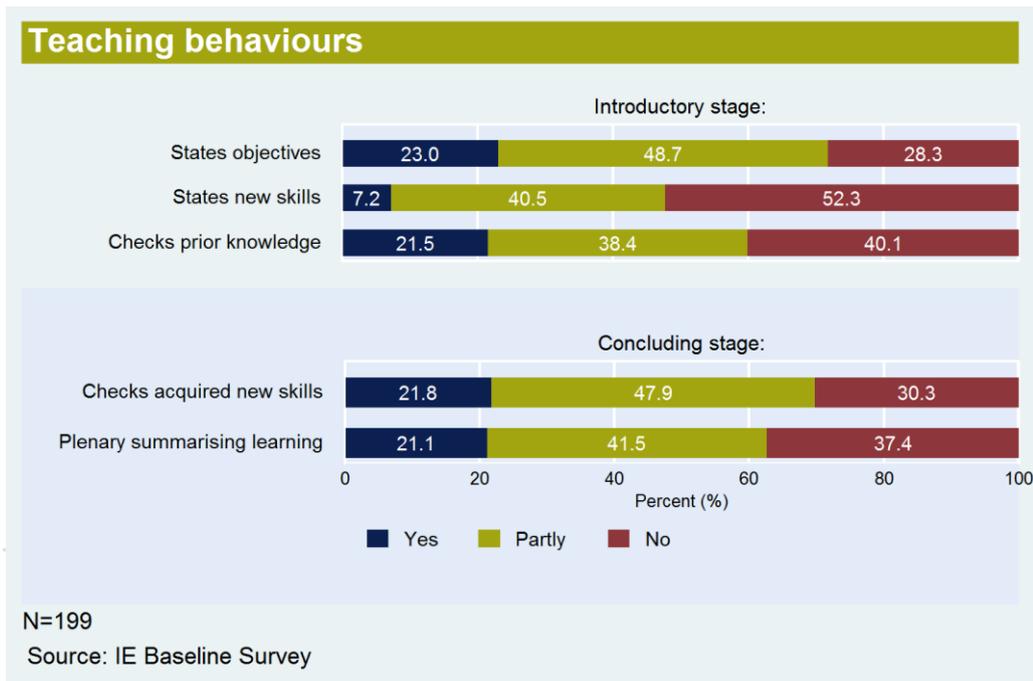
During the introductory stage of the lessons, teachers are expected to state the objectives of lessons and introduce what new skills or knowledge pupils are expected to acquire, linking these back to their prior knowledge. In observed lessons, 23% of teachers clearly stated learning objectives and 22% checked knowledge of prior work covered, while only 7% of teachers specifically stated what new skills or knowledge the pupils should have acquired by the end of the lesson (Figure 29).

⁶⁵ The set of 14 teaching behaviours/skills were mainly drawn from research carried out in many countries into effective teaching skills.

At the concluding stage of the observed lessons, only 22% of teachers checked if pupils had acquired the new skills or knowledge set out in the lesson introduction. Moreover, only 21% of teachers drew the whole class together to summarise what materials/topics had been covered and directed pupils to the next stage of the topic.

This means that the vast majority of teachers fail to display a core set of effective teaching behaviours, i.e. stating lesson objectives and new skills to be acquired, checking prior knowledge at the introductory stage and checking if new skills have been acquired and summarising learning at the concluding lesson stages.

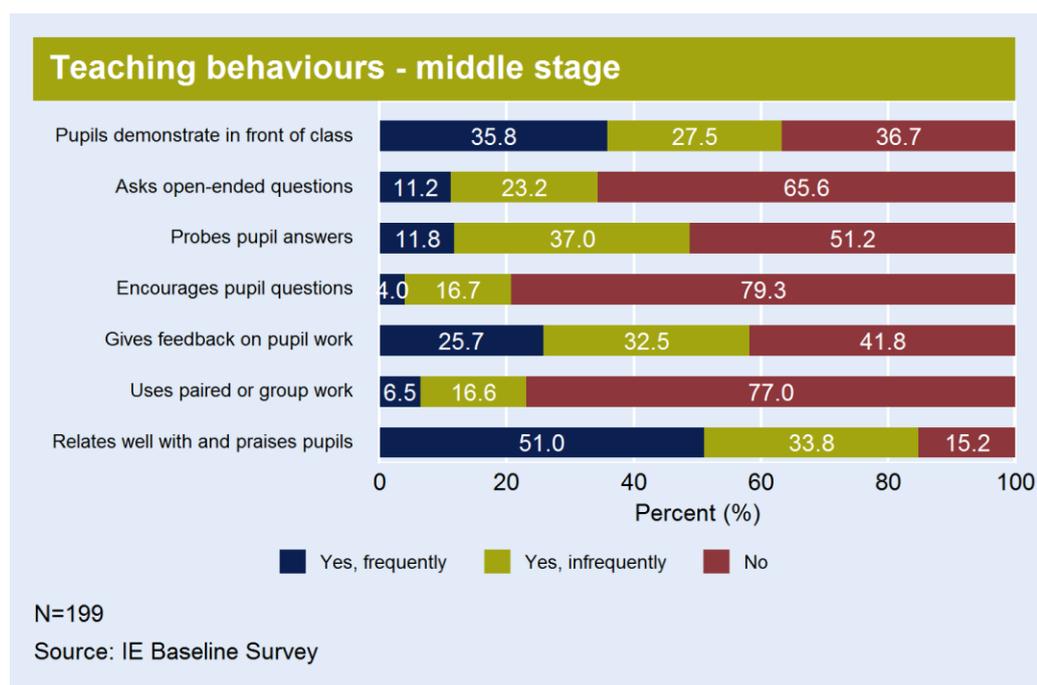
Figure 29 Observed teaching behaviours during introductory and concluding lesson stages



Notes: (1) Weighted estimates.

For the middle stages of lessons enumerators observed whether teachers displayed a given teaching behaviour frequently, infrequently or not at all. For clarity of presentation only the frequently observed and never observed categories are discussed below, whereas Figure 30 shows all three categories.

In the middle stage of lessons, only 4% of teachers frequently encouraged individual pupils to ask questions and explain ideas and 79% of teachers never did this (see Figure 30). 11% of teachers frequently asked pupils open-ended questions but 66% of teachers never did. 12% of teachers frequently commented on or probed pupil answers but again a large group of teachers (51%) did not do this. Just 6% of teachers frequently asked pupils to carry out activities in pairs or in groups and the vast majority (77%) of teachers did not do this at all. This suggests that there may be a small group of teachers that engages with individual pupils, encourages pupil interaction and uses paired or group work, but also a very large group of teachers that does not display these effective teaching behaviours.

Figure 30 Observed teaching behaviours during middle stages of lessons

Notes: (1) Weighted estimates. (2) Two of the middle-stage teaching behaviours are not shown in the figure: use of teaching aids, and use of the blackboard (results are in Volume II, Annex G2).

It was relatively more common for teachers to provide feedback on pupil work (26% of teachers did this frequently, while 42% never did) and to ask pupils to demonstrate in front of the class (36% of teachers did this frequently, 37% never did).

The vast majority of observed teachers (85%) were found to relate well to pupils, including conveying enthusiasm, using encouragement and praise to give positive feedback, and calling on pupils by name.

In summary, only a small proportion of the observed teachers demonstrated a range of effective teaching behaviours in the classroom. For the observed lessons, only 9% of teachers frequently demonstrated seven or more of the 14 effective teaching behaviours while 42% frequently displayed fewer than three of the 14 behaviours (results not shown). It is important to keep in mind that class size may affect which teaching behaviours are feasible: large classes may provide fewer opportunities for using interactive methods and may require whole class teaching methods. The average class size for the observed lessons was 60 pupils.⁶⁶ However, for one-third it was 68–156 pupils. This would indicate that class size may be restricting the use of some effective teaching behaviours, for example having pupils work in pairs, in some of the schools.

Finally, there was considerable overlap between the different teaching behaviours. Of the small group of teachers who demonstrated seven or more effective teaching skills, more than two-thirds also displayed gender and spatial balance in their interactions with pupils. This suggests that teachers who display effective skills may also be more likely to practice inclusive teaching.

⁶⁶ The official class size policy is 40 pupils (MoEVT 2009a).

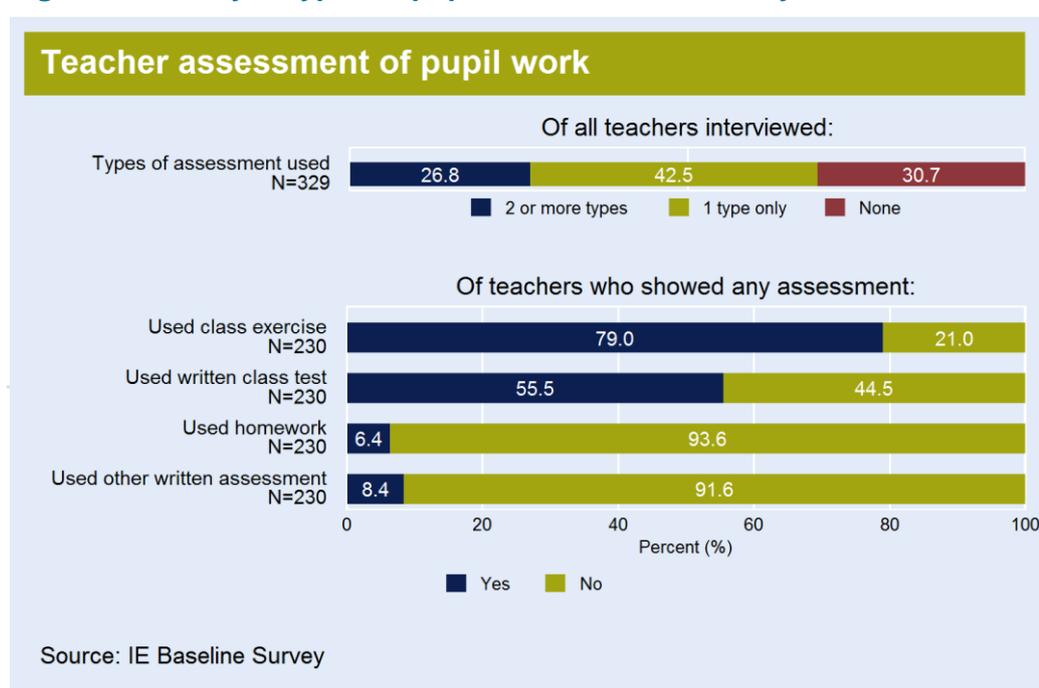
5.2.4 Teacher performance: Assessment of pupil work

During teacher interviews, teachers were asked what types of pupil assessments they used in the preceding five school days and to show a marked example of each type of assessment they mentioned (for the full definition see Volume II, Annex F).

Slightly more than one-quarter (27%) of teachers were able to show that they carried out at least two different types of pupil assessment in the previous five school days (see Figure 31). By contrast, 31% of teachers were unable to show that they had carried out *any* pupil assessment.

Teachers who reported using and could show examples of pupil assessment most commonly used class exercises and written class tests. More than half of these teachers (79%) were able to show a marked example of a class exercise and 56% were able to provide an example of a written class test. However, only 6% of teachers could show that they had assigned and marked homework.

Figure 31 Variety in types of pupil assessments used by teachers



Notes: (1) Weighted estimates.

One additional teacher behaviour was recorded during the lesson observations and in the teacher interviews: whether teachers switch between Kiswahili (official language of instruction) and a local language other than Kiswahili during teaching. Language switching was only observed in 4% of the observed lessons and in the interviews a similar proportion of teachers reported that they only teach in a language other than Kiswahili or switch languages.

Based on the quantitative results regarding the number of pupils speaking languages other than Kiswahili at home, the qualitative study aimed to gather data on teachers' language of instruction and if and how teachers address the fact that in many cases pupils speak a local language other than Kiswahili at home. The qualitative study found that two main approaches are used in the early years. Where someone who speaks the relevant local language is available, usually a teaching assistant, teachers employ a method most closely related to a grammar-translation method whereby instruction is translated from the mother tongue into the second language (Kiswahili) (Richards and Rodgers 2001). Where there is no one available who speaks the local language, teachers tend to use the direct (or 'natural') method of language teaching, whereby the learners'

native language is not used at all and the target language is used. The method relies on pupils memorising, mimicking and building vocabulary. Where the direct method of language teaching was used, teachers questioned the extent to which pupils are able to acquire subject knowledge, yet believe they must cover multiple subject areas despite this as this is required by the curriculum.

5.3 Teacher motivation and morale

The EQUIP-T programme will not only seek to improve teachers' subject knowledge and teaching behaviours but also to improve teacher motivation and morale in order to increase teacher attendance, time on task, level of effort and commitment to the job (Cambridge Education 2014a).

Teacher motivation can be defined as 'the willingness of employees to expend effort and perform certain desired patterns of work behaviour and teacher morale' and teacher morale as 'a state of mind determined by a teacher's level of work and personal well-being' (Molander and Winterton 1994; Evans 1992). Furthermore, 'Morale has a very close relationship to motivation, in that it acts to enhance and sustain any extrinsic or intrinsic motivation that might be present; however, both motivation and morale can also exist without the other' (Cambridge Education 2014a).

The IE collected data on selected higher-order measures of teacher motivation (absence and punctuality) and of teacher morale (job satisfaction and valuation by others), which are presented below together with findings on factors that influence teacher motivation and morale.

5.3.1 Teacher motivation and morale: Absence and punctuality

The IE baseline survey focuses on three aspects of teacher behaviour that provide indications of underlying levels of teacher motivation and morale: being absent from school (school absenteeism), being absent from classrooms although scheduled to teach (classroom absenteeism), and arriving at school on time (punctuality). For details of the measurement of absenteeism and punctuality see Volume II, Annex F.

Teacher presence at school and in classrooms is a prerequisite for teacher-focused interventions to have an effect on pupil learning because if teachers are absent the instructional time received by pupils is reduced, which has negative implications for pupil learning outcomes.⁶⁷

More than one in 10 (12%) teachers were absent from school on the day of the survey (see Figure 32). Furthermore, of those teachers who were present, almost two-thirds (63%) arrived late.

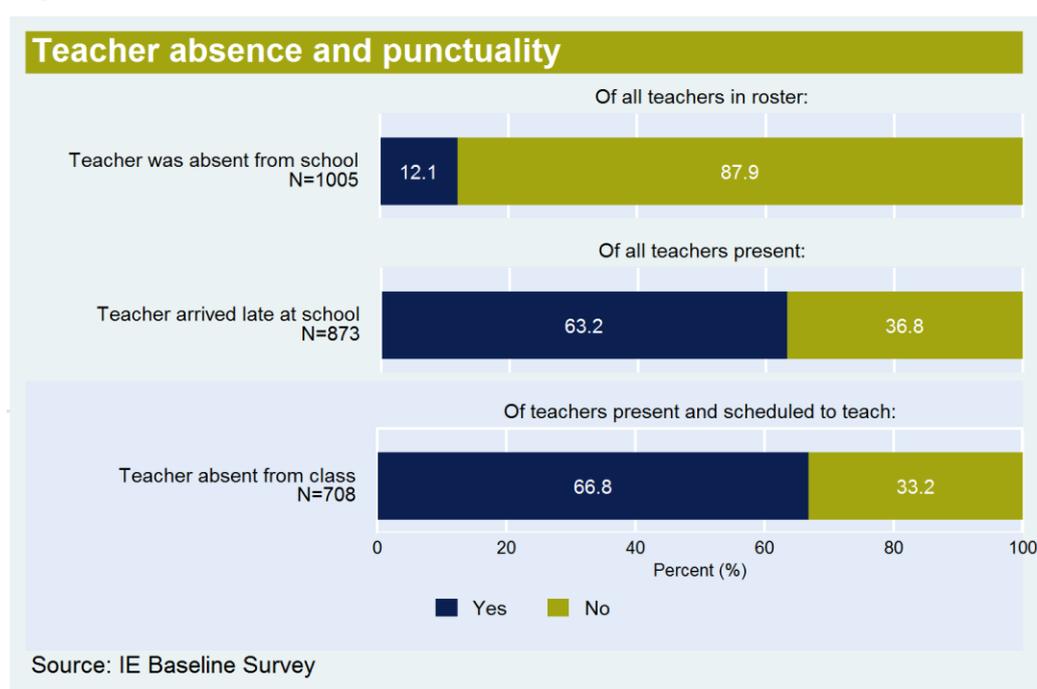
There is other evidence from Tanzania on school absenteeism. The Tanzania SDIS in 2010 estimated a school absenteeism rate of 20% for rural teachers (36% for urban teachers) (World Bank 2012), compared to the 12% for the IE baseline survey. There are differences in the two measures: the SDIS measured school absenteeism making one unannounced visit whereas for the IE baseline survey schools had some advance notice that the survey teams were coming. The two results are therefore not directly comparable, with advance notice for the visit possibly inducing lower absence rates in the schools visited for the IE baseline survey. Regardless, together these findings provide a strong indication that school absenteeism in Tanzania is high among teachers.

⁶⁷ Assuming that teachers have adequate subject knowledge and pedagogical skills.

Perhaps more surprisingly, of the teachers present at school and timetabled to teach a lesson before lunch, 67% were absent from their lesson, suggesting that classroom absenteeism is extremely high in the programme treatment area.

To enable comparison with the SDIS, which also measured classroom absenteeism, a second measure of classroom absenteeism is computed defined as 'teachers present in school, but not teaching' (note that this measure ignores whether a teacher was scheduled to teach or not). For the IE baseline survey, classroom absenteeism defined this way was again 67% compared to 50% for rural schools in the SDIS (68% for urban schools) (World Bank 2012).⁶⁸ This corroborates the IE baseline survey finding that classroom absenteeism is a serious problem in Tanzania.

Figure 32 Absence and punctuality of teachers



Note: (1) Weighted estimates; (2) Mean across all teachers (except for means based on lesson observation data).

To further compound classroom absenteeism, teachers timetabled to teach may be absent for parts of the lesson. The survey teams therefore monitored and recorded teachers who left during the lessons observed. For the 199 lessons observed, 15% of teachers left the classroom during the lesson,⁶⁹ spending an average of 4.5 minutes away from the classroom.⁷⁰ However, it should be underlined that there are some legitimate reasons for leaving the classroom during a lesson, for instance to fetch materials.

In the case study schools, multiple stakeholders discussed the scheduling of teachers to several classes at once due to a lack of teachers and a lack of classrooms, which can lead to teachers being absent from classrooms (although that this would not affect the measure of classroom absence presented above). In the FGDs with teachers, many said that developing their teaching

⁶⁸ Classroom absenteeism defined to take into account timetabling is 66.8% compared to the corresponding measure that does not take timetabling into account (which is 67%).

⁶⁹ The reasons for leaving the classroom were not recorded.

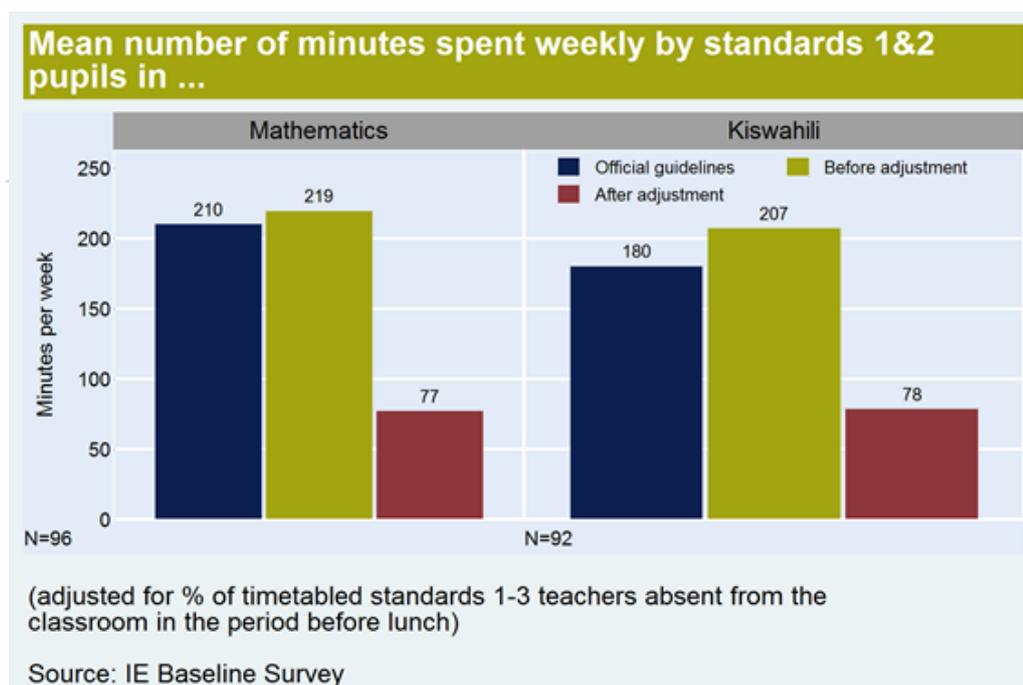
⁷⁰ As teachers were observed it is likely that these numbers underestimate the real proportion of teachers leaving the classroom and the time spent away.

schemes for the year is a task they find complex and time-consuming, but it is required by the WEC, and some teachers said this was one explanation for classroom absence.

While schools may be open for instruction and instructional time formally timetabled as 210 minutes for mathematics and 180 minutes for Kiswahili per week (see Chapter 4 section 4.6), the actual amount of instructional time received by pupils ultimately depends on whether teachers are in the classroom and if they are teaching.

The very high rates of teacher absence from school and classroom imply that there are substantial losses in instructional time. If an adjustment is made to account for the classroom absenteeism in the programme treatment area (see Volume II, Annex F for details of how this was done⁷¹), pupils only receive 77 minutes of mathematics instruction per week compared to the mean 219 minutes timetabled and 78 minutes of Kiswahili instruction per week compared to the mean 219 minutes timetabled (see Figure 33). If further adjustments were made to account for whether teachers who are present in the classroom use their lesson times effectively for teaching, instructional time estimates would likely be even lower. Such large losses of instructional time will have large, adverse consequences for pupil learning.

Figure 33 Official requirements, timetabled and estimated actual instructional time in mathematics and Kiswahili for standards 1 and 2



Notes: (1) Weighted estimates.

In the qualitative study, it was established that a key reason for loss of instructional time in one of the schools (Kishapu Better) was due to a lack of classrooms and teachers. Teachers described the lower standards teacher teaching pre-primary for 40 minutes, then sending those children away and teaching standard one for 40 minutes and then standard two for forty minutes.

⁷¹ In Volume II, Annex F, see description and notes beside indicator 'actual weekly timetabled minutes, after adjustment'

5.3.2 Teacher motivation and morale: Job satisfaction and valuation by others

The IE baseline survey also asked teachers how satisfied they felt with their job and how valued they feel for their work by the community and head teacher respectively. They were asked to report where they placed themselves on the day of the survey on a 10-point scale/ladder.

Teachers reported a mean score of 7.7 out of 10 for job satisfaction, where 1 is completely unsatisfied and 10 is extremely satisfied. On the perceived valuation by others scale, where again 1 is low and 10 is high, teachers' mean community appreciation rating was 6.5 out of 10 and the mean head teacher appreciation rating was 8.5 out of 10. These scores suggest that teachers tend to feel less valued by the community than by their head teachers⁷² and that there is a lot of variation in how valued different teachers feel by their community (with a range from 2 from 10). However, it is important to note that these values for the teacher job satisfaction and community and head teacher valuation indicators should not be interpreted on their own at the baseline but rather will be used to measure any changes to these teacher ratings attributable to the EQUIP-T programme after the 2016 follow-up survey.⁷³

5.3.3 Factors that influence teacher motivation and morale

If teacher motivation and morale are low this tends to affect teacher performance, consequently manifesting in teacher absence, tardiness, and reduced efforts to prepare for lessons, engage with pupils and so on. Therefore, it is vital to identify factors that influence teacher motivation and morale and that can be targeted to raise motivation and morale, ultimately improving teacher performance.

Starting with the perceptions of others, the qualitative study found that the majority of national, regional and district respondents believe there to be only a small percentage of teachers within the education system who work diligently despite environmental constraints due to their passion to teach children and intrinsic motivation, i.e. the desire to do something for its own sake (Prendergast 2008). It was agreed by these stakeholders that these teachers are in the minority, with estimates of between 5% and 30% of the current teaching force. These respondents also discussed the existence of a second group of teachers whose intrinsic motivation is undermined by environmental constraints, including by a lack of professional development opportunities.

The vast majority of regional and district education officials indicated that the low levels of societal respect for the teaching profession combined with the low requirements to join the profession have resulted in a large number of teachers in the workforce who are either unlikely to respond to efforts to increase teacher motivation and morale or who are only likely to respond to tangible and extrinsic incentive mechanisms. One respondent proposed that this group of teachers currently makes up 25% of the teaching force.

Teacher housing

Teacher housing was the most common problem identified by teachers, head teachers, community leaders and parents during the FGDs and KIIs, and teacher absenteeism and lateness was most often attributed to a lack of teacher housing near the school.

⁷² Statistical tests found that these three scores are statically significantly different from each other.

⁷³ Teachers may be reluctant to report truthfully about job satisfaction and head teacher support, for instance, if they believe what they say may get back to their head teacher.

In the majority of case study schools teacher housing presented, what is referred to in a seminal paper on this topic (Getzels and Guba 1957), as a *role conflict* because there was no housing available for teachers and, therefore, they had no choice but to rent houses in nearby towns and travel long distances to and from work, often on transport that is unpredictable. This is a role conflict because, regardless of teacher motivation and morale levels, the expectation that teachers will arrive at school on time conflicts with the requirement to rent houses far away from the school, which stems from the lack of (adequate) teacher housing close to the school. However, in two case study schools (Kishapu Typical and Kishapu Lower) the issue of teacher housing was instead a *role-personality conflict*. In these schools teachers were provided with housing near the school, but some of them chose to rent houses in town due to their personal preference to live in more comfortable housing whereas some chose to live in the housing provided by the community.

Within some case study locations, communities and school stakeholders agreed that the community is responsible for funding and building teacher houses, yet the standard or type of housing expected differed between the groups. In these cases communities built mud houses, but these houses did not meet the expectations of teachers.

Whether caused by role conflict or by role-personality conflict, the issue of teacher housing is a major source of the low levels of teacher motivation and morale in the case study schools.

Salaries

Low and late salary payments were often cited as reducing teacher motivation and morale and teachers across all the case study schools were engaging in other IGAs, including entrepreneurship and agricultural activities. In the schools where teachers engage in agricultural activities, they are absent for long periods at particular times during the year.

Community members and head teachers recognise the role conflict with respect to salaries, acknowledging that teachers cannot be expected to live on the salaries provided. Salaries were also considered an issue of respect; teachers and head teachers often discussed feeling 'demoralised' or 'disrespected' by the government, citing their low salaries as evidence. One head teacher's comment reflected this sentiment: 'If the government doesn't respect him, why should teachers?' In all but one of the case study schools all stakeholders agreed that, until salaries are raised, teachers will not be effective in their roles.

If the phenomenon had been observed only among a few teachers, it could have been argued to be a matter of personal preference. However, the fact that all stakeholders agree suggests that this is a role conflict as all teachers fall short of the expectations of their roles in order to meet the obligations of another role – providing for their families.

5.3.4 What schools do to increase teacher motivation and morale

About half of the case study schools had developed informal and formal systems to incentivise teachers:

- Teachers at many of the schools were provided with financial incentives when their pupils performed well in examinations.
- Several schools offered teachers incentive funds to teach remedial classes to pupils soon to undertake examinations.

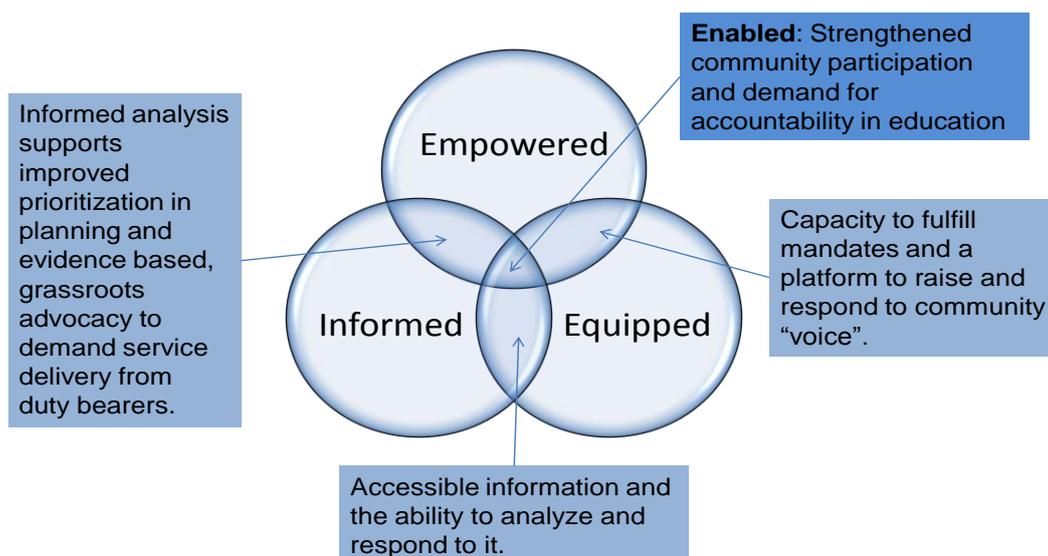
- In one school, teachers had devised a system to support each other through a financial scheme wherein all teachers contribute to a shared pot and the funds are distributed as agreed by the teachers.
- Teachers at several schools had informally created support networks to assist each other in learning in the absence of INSET, and also organised regular open discussions on the challenges faced within their roles. In one school, these initiatives were instituted by the head teacher.
- Two SCs discussed an initiative they use to increase teacher morale, i.e. the organisation of an awards ceremony at the end of each year to acknowledge teacher efforts. However, the parents and teachers in these two schools did not discuss this initiative, indicating that it is unlikely to be an effective mechanism for increasing teacher morale.

6 EQUIP-T Component 4: Community participation and accountability

This chapter presents baseline indicators from the quantitative survey and findings from the qualitative research in the area of community participation and accountability.

Component 4 of the EQUIP-T programme seeks to **empower** communities ‘to take an active role in improving school outcomes and services by **equipping** them with tools, resources, and the provision of appropriate, accessible **information**. The theory underpinning this component is that the greater engagement by the community in school operations and outcomes will provide much needed support and resources, that will enhance the relevance of education (Cambridge Education, 2014b).’

Figure 34 Key features of EQUIP-T Component 4, Community participation and accountability



Source: Cambridge Education, 2014b

In order to provide a picture of the situation at baseline, the following sections will report on the current state against these three areas in which the programme seeks to affect change.

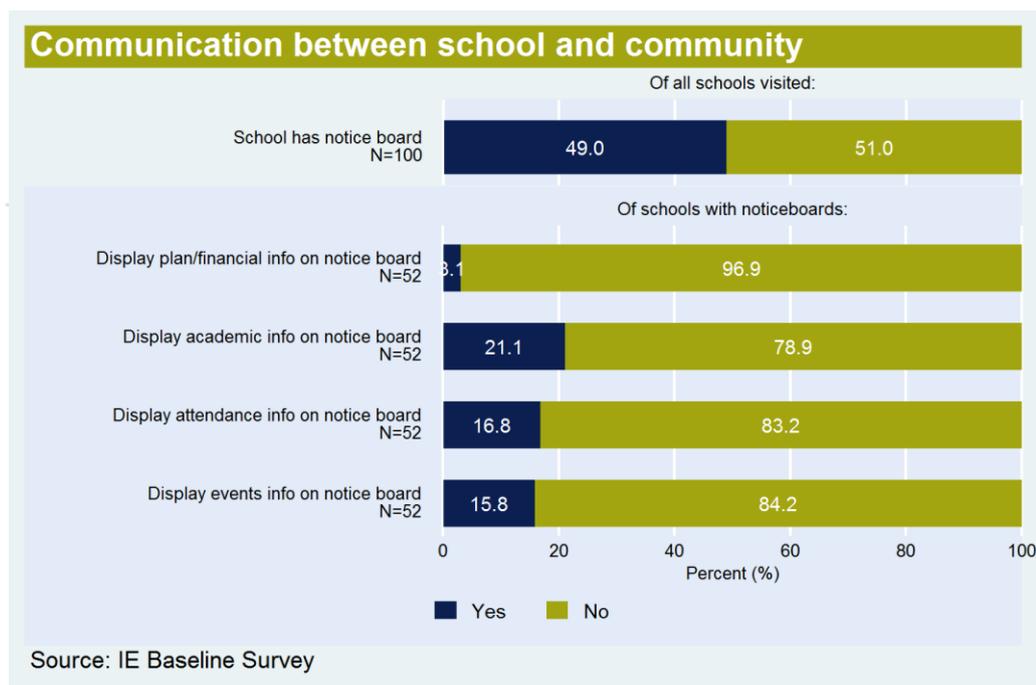
6.1 Informed communities – communication between schools and communities

One of the EQUIP-T core activities on community participation and accountability is to improve communication mechanisms between local government, SCs, parents and the wider community. The aim of this activity is for communities to be ‘adequately informed about school operations, needs, plans and performance’ (Cambridge Education 2014a).

6.1.1 Mechanisms of communication

The IE baseline survey found that just over half of schools in treatment districts (51%) did not have a notice board (Figure 35). Furthermore, only a small proportion of the schools that do have notice boards actually used them to communicate the types of relevant education information that the EQUIP-T programme seeks to target. Only 8.1% displayed information on the WSDP, the school budget or school capitation grants on their notice boards. Around one in five (21%) displayed academic results or other information related to teaching and learning, while 17% of schools with notice boards displayed pupil or teacher attendance information and 16% displayed information on school or community events or meetings on their boards.

Figure 35 Use of school notice boards to communicate with communities



Notes: (1) Weighted estimates.

The qualitative research found that the main form of communication between schools and parents is through letters given to pupils to take home to their parents. Notice boards were not mentioned as a source of information by parents during FGDs. Head teachers explained that letters would be written in Kiswahili and that pupils were expected to deliver the letters to their parents and if the parents are either not able to read, or not able to read Kiswahili, pupils would read out the letters and translate them for their parents. Head teachers, teachers and SC members expressed frustration at the fact that parents frequently do not respond to these letters and fail to attend meetings. While parents acknowledged that some parents fail to respond to communication from the school, they largely felt that parents reacted to communication from schools. A scenario

discussion with children during the FGDs with standard three pupils found that pupils carefully screen letters before delivering them to their parents:

I will read it first before taking it to him in order to know if it has to do with me. [FGD with standard three pupils in Kishapu Lower Performing School]

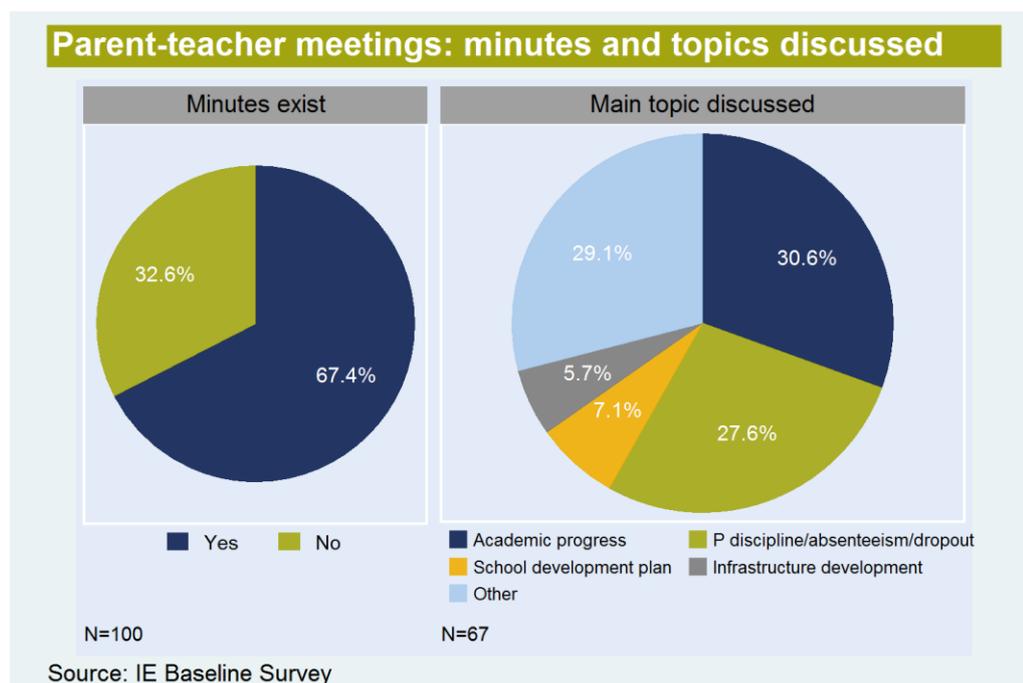
Pupils reported that letters containing information for parents about mistakes they had made or their absence from school would not be delivered by them for fear of being punished:

I will throw it away. [FGD with standard three pupils in Uyui Better Performing School]

However, letters asking parents to come in for meetings were typically delivered. It is likely that some letters are lost by children who often walk several kilometres to get home without bags for their exercise books and that some pupils are unable to read Kiswahili effectively themselves. This method of communication between the school and parents therefore seems unreliable, even in cases where parents would be supportive of the school if informed.

It happened one day that I noticed that my children had not gone to school for almost one week... From then on I decided to come to school and ask the headmaster if my children were coming to school and he said that they are not and moreover he had sent me a letter and I hadn't shown up. After discovering this I decided to take my children to school myself. So I think the best way is to use mobile phones as a means of communication between the school and parents instead of letters. [FGD with SC in Uyui Lower Performing School]

Other mechanisms used by schools to communicate with parents that were frequently mentioned were village meetings, convened with the help of the village council and the SC, school meetings organised in conjunction with the SC, and in some cases visits to parents' houses by head teachers or SC members. The IE baseline survey found that whole-school parent-teacher meetings are taking place in the majority of schools (see Figure 36). Two-thirds of schools were able to produce minutes from a meeting, and these showed that the main issues being discussed are academic progress, pupil discipline, absenteeism or dropout, and an unspecified 'other' category.

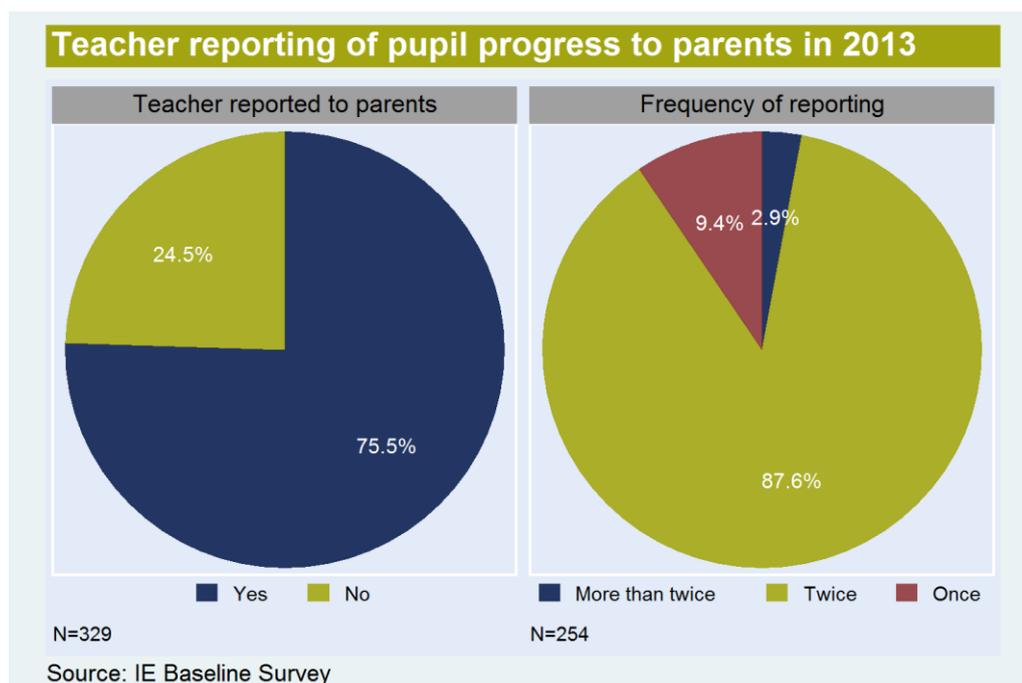
Figure 36 Availability of minutes from parent-teacher meetings and topics discussed

Notes: (1) Weighted estimates. (2) The 'P' in the label 'P discipline/absenteeism/dropout' means pupil.

If communication channels are perceived to have broken down between the school and the community, both seem to rely on the village council for a resolution. However, both teachers and parents viewed the involvement of the village council to constitute an unfair escalation, although simultaneously they reported that the involvement of the village council frequently led to better outcomes. Head teachers explained that the involvement of the village council allowed for fines to be issued to parents who failed to address disciplinary issues such as the attendance of pupils, and parents reported that the involvement of the village council was beneficial in addressing the disciplinary behaviour of teachers. Both the parents and schools cited a lack of formal mechanisms of authority as reasons for involving the village council.

6.1.2 Reporting on academic progress

As part of the IE baseline survey, teachers were asked about their reporting of individual pupils' academic progress to parents. A quarter of all teachers said that they did not report pupil progress to parents at all in 2013 (see Figure 37). Of those who said that they did, 9% reported doing so once and 88% twice. These can be assumed to coincide with end-of-year and end-of-term reporting to parents, a finding supported by the qualitative research. Only 3% of teachers who said that they did any individual reporting of pupil progress to parents in 2013 did so more than twice. This suggests that only a very small proportion of teachers (about 2%) report individual pupil progress to parents prior to presenting their end-of-term results, meaning that very few take ad hoc opportunities to highlight areas where parental support might particularly be needed.

Figure 37 Teacher reporting of individual pupil progress to parents

Notes: (1) Weighted estimates.

The qualitative research found that checking on the academic progress of children was largely perceived as being the responsibility of parents. This was to be achieved by checking exercise books when the child returns home. Some stakeholders acknowledged that this poses a problem for parents who themselves only enjoyed limited education or who are not proficient in Kiswahili. All stakeholders agreed that only a small portion of parents come to school and ask about the progress of their child. While head teachers and teachers across all schools reported addressing parents' questions, the FGDs with parents showed that parents feel that teachers do not – at times – take them and their concerns seriously. Responses by parents differed greatly between schools, dependent on whether the relationship between the school, SC and parents was good or not.

6.1.3 Relationship between schools and communities

Effective communication was widely viewed as the key factor leading to a good relationship between the school and community. However, it was striking that opinions on who should ensure that this takes place and that the relationship is good differ, leading to scenarios where all stakeholders agree that communication has broken down but the responsibility for fixing it is attributed in turn to parents, the SC, the village council, teachers and the head teacher. In addition, the qualitative research also showed that the perceptions of the state of communication and the relationship between the school and community differs between stakeholders at the same school, with – in one instance – the head teacher and SC reporting a good relationship with the community while parents perceived the relationship as bad and communication as lacking.

A lack of adequate information, especially in relation to child work at school, was also found to negatively affect the relationship between the school, parents and the wider community. Parents are aware that their children work on school farms and participate in IGAs. As long as this takes place after school hours, parents typically support this. However, parents frequently questioned whether working on school farms really only takes place after hours, and are unhappy with the fact that, in addition to this task, children also clean the school, fetch water and work in teacher houses. Parents stated that that they free up time for their children to attend school not to conduct

household chores for teachers. In one case, parents who complained about the practise were told by the school that if they were unhappy about this, they should no longer expect to receive permission for their children to help them at home during harvest season.

In many schools, IGAs and the income generated from them seems to be an area of conflict. Parents frequently accuse schools of misuse of the proceeds from IGAs and of using the money for the benefit of teachers, while SCs and head teachers explain that IGAs do not lead to substantial yields. The use of money generated from IGAs to incentivise teachers or buy teaching aids rather than food for the children is at times perceived by parents as a breach of promise. Parents largely want to see the schools provide food for their children from the maize cultivated on school farms.

Partly related to child work at school, but also to harassment and violence, the relationship between some teachers and pupils is a source of serious concern for many parents. Some parents felt very strongly that their children, especially girls, should not be made to work in teachers' houses because of the potential for sexual harassment. Parents frequently complained about sexual relationships between teachers and pupils and felt that teachers were not held to account for their behaviour. Young, male, single teachers were perceived as posing a risk to girl pupils.

They convince our girls to sleep with them and get them pregnant and then they blame it on other villagers [FGD with parents at Kishapu Better Performing Primary School]

There is a teacher who is still in this school who has made pregnant many girls. [FGD with parents at Kishapu Better Performing School]

Communities had very different views on ways of decreasing the risk of sexual harassment of pupils by teachers. In particular the issue of where teachers ought to live, arose repeatedly. Some communities feel that having teachers living close to the school enables them to monitor teachers more effectively. Others feel that it is much better if teachers live far from communities reducing the opportunity for abuse.

Another area of conflict between schools and communities is around corporal punishment, the frequency with which it is applied, and the severity of the punishments administered. In particular, revised government regulations on the severity with which children can be punished have led to a situation wherein some parents believe corporal punishment to have been abolished while teachers generally favour it as a disciplinary measure. The opposition by parents to corporal punishment differs both between the different case study schools and within schools. Children, however, universally reported 'being beaten with sticks' as the primary negative experience at school.

As head teacher, I favour corporal punishment because it saves time and children feel the pain so they won't repeat the mistake. If a teacher punishes a child in a different way with a task, then I punish myself because I have to supervise it. Also, giving children tasks to do might injure them and there is no clear policy that describes a good or a useful punishment of pupils. [KII with head teacher at Mpwapwa Lower Performing School]

6.2 Equipped communities – capacity of communities and school committees

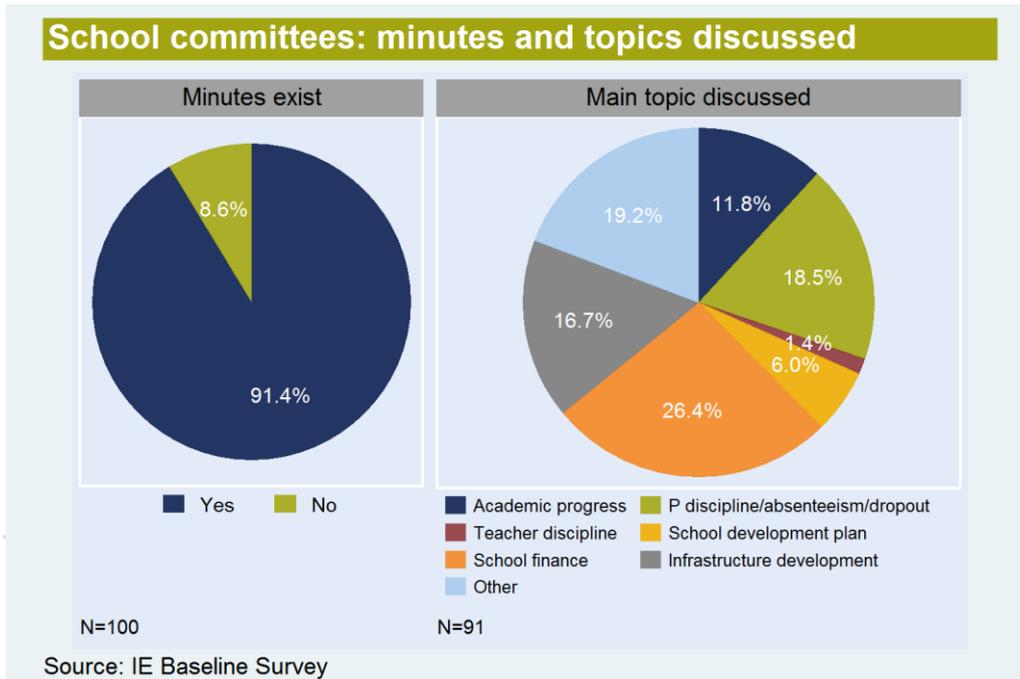
The EQUIP-T programme seeks to equip parents, the wider communities and SCs through capacity training and an enhanced understanding of their and other stakeholders' roles and

responsibilities. The programme will support communities to develop and express priorities for the school and advocate for their delivery.

6.2.1 Understanding of the roles and responsibilities of the school committee

The EQUIP-T programme activities aim to improve the capacity of communities and SCs to function effectively. These include capacity-building workshops for school communities and SCs, working with WECs to increase their understanding on SCs’ mandate and operations, and the formation of parent–teacher partnership groups under the guidance of SCs.

Figure 38 Availability of SC meeting minutes and topics discussed



Notes: (1) Weighted estimates. (2) The ‘P’ in the label ‘P discipline/absenteeism/dropout’ means pupil.

The IE baseline survey asked head teachers whether SCs already existed in their schools and, if so, if they could show minutes from their last meeting. Nearly all schools (91%) could show minutes from a SC meeting (see Figure 38), which suggests that SCs are functioning to some extent in most schools. A review of these minutes revealed that the three main topics discussed were school finance, pupil discipline, absenteeism or dropout, and infrastructure development. Given the scale of the problem of teacher absenteeism described in the previous chapter, it is telling that in only 1% of SC minutes was teacher discipline the main item under discussion.

The survey also asked about the presence of formal parent-teacher bodies (not simply whether the school holds whole-school parent-teacher meetings). Only 14% of schools reporting having some type of formal parent-teacher group in place.

The qualitative research found that SCs have a very broad understanding of their roles and responsibilities. Typically SCs describe their role as being that of a problem solver at the school, mediating between teachers, parents and pupils and engaging with the village council. SCs are very aware that their power is limited and that they can only call meetings with parents. In order to call meetings with the community, they have to liaise with the village council, which then initiates these meetings. If SCs do not have a good relationship with the village council this can lead to long delays and is a cause of frustration. The relationship between SC and village council differs greatly between schools.

SCs also see the creation of a good relationship between teachers and parents as one of their core functions. Dealing with pupil and teacher absenteeism, ensuring that teaching takes place, improving academic performance, maintaining infrastructure, finding additional forms of revenue for the school through IGAs, ensuring the proper use of financial resources, persuading parents to contribute financially and managing their in-kind and financial contributions, and increasing the awareness of the importance of education were also mentioned frequently as constituting key responsibilities of SCs. Only in one case did a SC explicitly mention representing parents' views as a core responsibility.

6.2.2 Strengths and weaknesses of school committees

The qualitative research found that SCs generally felt that they were doing a good job at fulfilling their responsibilities. Notably, even in cases where parents reported that they were unhappy with the performance of the SC, the committee itself seemed unaware of any tensions:

We only know the school has a SC because we were neither invited to select them nor introduced to them. We are not told what they do. Maybe it's because they think we are a problem. [FGD with parents at Kishapu Typically Performing School]

Teachers and parents perceived the SC as being at a disadvantage when dealing with questions of financial resourcing and education matters:

A SC can ask what a stamp pin is and a teacher tells them it costs TZS 200,000 or that a flip chart costs TZS 400,000. They approve the budget thinking that is how much these items are and that they are important for teachers to have. And when the list reaches the district level, there is a person there who approves it, believing that if the money is given he will also have a share. [FGD with teachers at Kishapu Typically Performing School]

We are not happy with how the SC sells school grasses to the cattle keepers and does not direct the income to the pupils' needs. I think it is because SC members are not skilled enough to deal with this kind of issue, as with the income it would have been possible for the school to provide even porridge to the children but they do not know how to spend the income generated from the school farms, rented piece of farm and crops sold. Also the SC members should be educated so they can handle school issues – some of them stopped in standard 7 so they are not qualified to deal with the school problems in the manner we expect. [Parent FGD at Kishapu Typically Performing School]

SCs themselves reported facing a number of problems in conducting their job, including in certain cases a lack of support from the village council and district. In about half of the case study sites, the relationship between the village council and the SC was described as difficult and as leading to delays to school projects. In one case, a SC reported waiting for seven years for the government to release funds for the renovation of a classroom. SC members reported that parents often blamed this lack of responsiveness and action on the SC. Other problems mentioned were a lack of funds for transport to attend meetings at the district and ward level, difficulties when talking to parents who are unaware of the importance of education, a lack of support from parents for IGA ideas, non-responsiveness when asking for financial contributions from parents, late payments by parents, a general lack of funds, lack of attendance at meetings, and the long distances pupils and teachers travel to school, which leads to unhappiness on the part of teachers and parents. Interestingly, only one SC felt like the school was not adequately involving it in decision-making processes. However, another SC reported working in a school environment where the relationship between teachers and the head teacher, teachers and parents, and the SC and the village council had broken down.

In two schools in Tabora, SCs stated that parents were reluctant to join the SC and refused to stand for election. SC members explained that parents were either scared of what the role would entail or preferred to pursue IGAs rather than volunteer at school. The average duration of service on a SC by SC members who participated in the FGDs was just over four years. However, when individuals' time of service on SCs is examined, just under 50% of committee members have served for periods in excess of the stipulated three-year period, with one member sitting on a SC for 14 years, another two for 10 years, one for nine years, five for eight years, and seven for six or six and-a-half years. The process of selecting and rotating members on SCs was therefore not functioning as described in the regulations in the qualitative research sites.

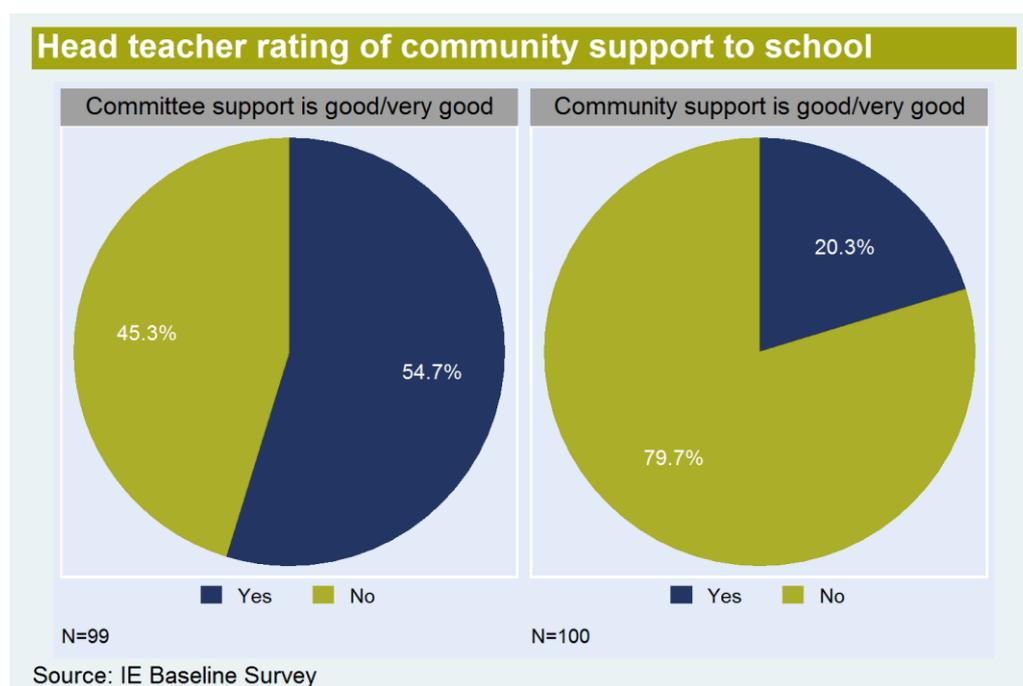
6.3 Empowered communities – engagement and accountability

Another EQUIP-T programme activity seeks to deliver increased community participation in problem analysis, planning and resource mapping. This in turn will 'mobilize, motivate and activate communities to address education needs and access relevant support. Communities will be motivated and able to coordinate actions to deliver some of their planned objectives themselves, as well as identify and seek support from external actors and demand service delivery from duty bearers' (Cambridge Education 2014b).

6.3.1 Supporting schools

Head teachers were asked to rate the support of the SC and of the wider community to their school on a 1 to 5 scale, with 1 for 'very poor' and 5 for 'very good' (Figure 39). While just over half of all head teachers in treatment districts (55%) rated SC support to their school as either good or very good, only 20% rated wider community support equally positively. Also related to this trend is the earlier finding that teachers gave a lower rating for how they perceived they are valued by the community compared to how they are valued by head teachers or in regard to their own job satisfaction (see chapter five, Section 5.3.2).

Figure 39 Head teacher rating of SC and community support to school



Notes: (i) Weighted estimates.

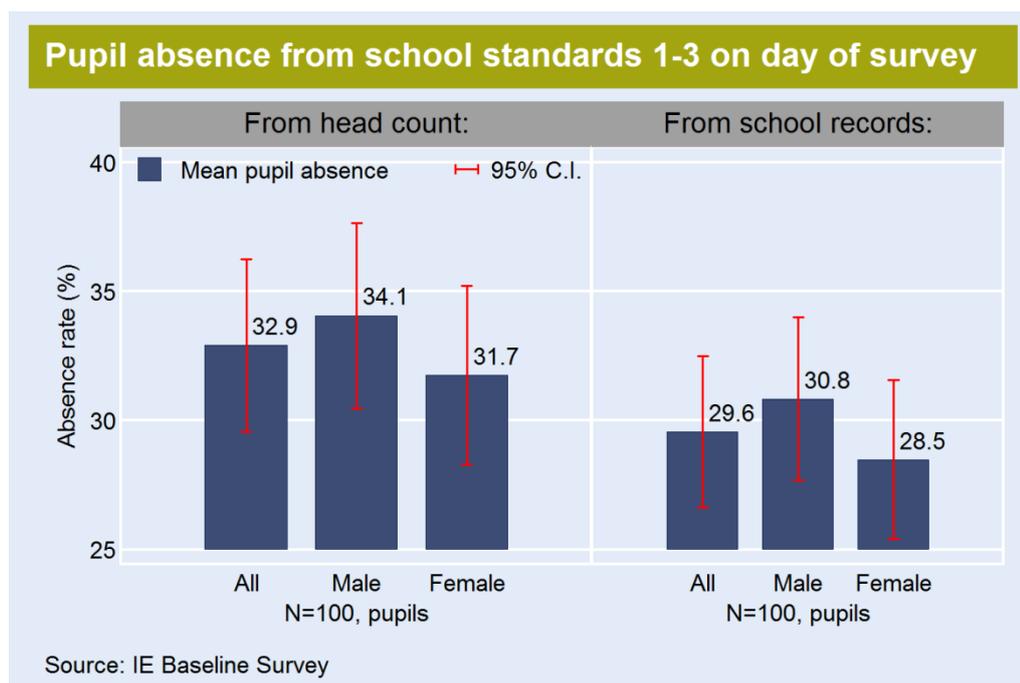
The qualitative research found that the perception that community support to schools is limited might be due to what the different stakeholder groups perceive as constituting support. One way of estimating the current level of support to schools by communities and parents is by looking at financial and in-kind support provided to schools. Across all nine schools, parents and village chairpersons reported providing some form of support to schools. This support could either be financially for specific tasks such as school feeding or payment of the school security guard or cook or it could be in kind. Typical in-kind contributions were either maize for school feeding programmes, materials for constructing teacher housing or renovations at school, or man power for school-related tasks. Notably, head teachers, teachers and at times also SC members seemed to only perceive financial contributions as proper support to schools. FGDs with parents discussed the socioeconomic status of the communities making financial contributions and – at times – even in-kind contributions in the form of maize difficult. Parents across all nine schools felt that they were contributing to the school, but did not always know what happened with their contributions, which in some cases had led to unwillingness to contribute further. However, parents also acknowledged that recent droughts had diminished the contributions that they were able to make to schools and that they failed in supporting schools in the upkeep of facilities and the building of teacher housing. However, parents largely attributed this failure not to a lack of willingness but rather to poverty and recent economic shocks.

6.3.2 School attendance

The rates of pupil absence can serve as an additional indicator of community engagement and support of schools. High rates of absence among early grades could indicate a lack of parental valuation of education. However, it should be noted that this is an imperfect indicator of community engagement, as it does not consider other reasons that might make parents reluctant to send their children to school, such as doubts about the quality of education, indirect or direct costs associated with sending their children to school, or perceptions about the return on education. All of these factors were mentioned during the qualitative research as affecting pupil attendance.

On the days that their schools were surveyed, a head count found that one-third (33%) of pupils in standards one to three were absent (Figure 40) and that boys had a slightly higher absence rate than girls, although this difference is only weakly statistically significant.⁷⁴

⁷⁴ On both the figures obtained through the head count and the school records, the differences between the weighted estimates of the absence rate of boys and girls were statistically significant, with a $p < 0.10$. Full results not shown here.

Figure 40 Standards 1–3 pupil absence rates from school on day of survey

Notes: (1) Weighted estimates.

The data on pupil absence was collected during the rainy season (late March to mid-May), and the high rates observed may be partly seasonal. On the other hand, the schools were aware in advance that the survey team were coming, so the absence observed may be an underestimate compared with a typical day. The qualitative research found that during the rainy season pupil attendance in farming communities could drop significantly; for example, in Uyui Better Performing School, estimates of attendance during the rainy season were 50% to 60%, reaching around 80% to 85% during the rest of the year. All stakeholders across all nine schools agreed that feeding programmes at schools increased attendance and enrolment in school. In addition to improving attendance, feeding programmes were also reported to improve educational outcomes.

Reasons for dropping out cited across all schools and communities were tending to livestock, pressure to get married for girls, involvement in bad peer groups, lack of uniforms and exercise books, bad performance at school leading to discouragement of pupils, and a fear of punishment for bad behaviour or academic performance. Girls themselves were often blamed for dropping out of school for falling pregnant and described as 'easily tempted'. The fault of dropout due to relationships, even if with older men, was placed on the girl child. In some instances teachers expressed frustration at the lack of support provided by communities in finding the men who got the children pregnant. These teachers and head teachers claimed that communities at times prevented them from holding these men to account and preferred to come to an agreement with the men and either receive payment or marry their daughters in return for a dowry. Head teachers, teachers and some community members attributed this behaviour to the perception held by parts of the community that the education of girls continues to be viewed as less important than that of boys, especially in regard to secondary education.

Teachers generally blamed parents for pupil absenteeism and said it was parents' responsibility to ensure that children attended school. In some cases, children were also held directly responsible. While parents generally agreed that it was their responsibility to ensure that children attend school, they also reported having to rely on information from the school about non-attendance of their children. A lack of awareness of the importance of education was widely cited as a reason for low attendance, while parents felt that this was only true for certain parts of the community. All

stakeholders agreed that more should be done to inform the community about the importance of education, especially for girls.

Some teachers demonstrated an unwillingness or inability to adequately deal with the question of absenteeism, blaming it on uneducated parents and an inability to educate a child whose parents themselves are uneducated:

There was a certain parent who came to complain that his standard four child does not know how to write but after we made some follow ups we found that even that parent does not know how to read while he has completed standard seven, so it is the family problem.
[FGD with teachers at Uyui Better Performing School]

6.3.3 Perceived ability of the community/parents to hold school management and teachers to account for the delivery of education

Teachers and head teachers across all case study schools felt that parents frequently came to school to complain if they were unhappy with the state of schooling. Parents were described as furious, aggressive and angry and teachers and head teachers were weary of the effect this had on the wider relationship with the community and the perception of their school.

One pupil came to me and said head teacher can you please look at my exam and the overall mark. I took the paper and added up the overall mark and found that some of the marks had been reduced. He said that he was supposed to be the first and how come he came second. I was pained by this. If he had gone to the parent first we [school] would have suffered as this would have been preached to the whole village. They would then say, can you see what they [teachers] do? These kinds of issues have affected us. [KII with head teacher at Uyui Lower Performing School]

At the same time, parents in several schools discussed the repercussions of trying to hold teachers and the head teacher to account. Parents felt that teachers would indirectly punish the children of parents who demanded better education, branding them as trouble makers, and no longer providing them with the support they needed. These parents were therefore sceptical of their ability to hold schools to account.

It is interesting that while parents generally felt that their ability to hold teachers to account was limited, teachers and head teachers reported fearing parental attempts to hold them to account, especially if the village council and other spheres of government were involved. One possible explanation might be that parents tended to focus on the resolution of their complaint, which they typically reported as lacking or taking a significant amount of time. This delay or lack of action was perceived by parents as an inability to hold schools to account. Teachers and head teachers, on the other hand, mostly perceived the power of parents in terms of causing problems for the involved individuals and necessitating justification of behaviour to the head teacher or WEC and less in terms of a hard measure that could swiftly affect change.

The multiple sources of perceived conflicts between communities and schools which emerged from the qualitative analysis (see section 6.1.3) including child work at school, corporal punishment, sexual harassment of girls by teachers, and the use of community contributions and revenue from IGAs, also strongly suggests that accountability relationships are not working well in many cases.

7 EQUIP-T Component 3: District and regional management

The chapter presents baseline indicators from the quantitative survey and findings from qualitative research in areas of district and regional management where the programme expects to see change.

7.1 Current district and regional management of primary education

The objective of Component 3 of the EQUIP-T programme is to strengthen systems and human resource capacity to support sub-national management of education (Cambridge Education, 2014b). The programme seeks to do this by providing training and mentoring to district and regional officials on planning, budgeting, EMIS analysis, public financial management and M&E.

A capacity-development strategy relies on three key processes. The first is to define the desired future capacities, the second is to define the level of desired future capacities and the third process is to assess existing capacity levels. The third stage includes identifying existing gaps and what needs to be done in order to strengthen capacity in the short, medium and longer term.

The EQUIP-T Inception Report (Annex XVIII) outlines these gaps as including:

- Physical resources, including IT equipment, communications and vehicles within districts require upgrading in order for districts to provide an appropriate level of support to schools;
- Transparency and information on key financial and budgetary procedures;
- Support systems to permit effective coordination and monitoring of key activities such as teacher and head teacher training;
- Support for school improvement systems (including school–community participation mechanisms, classroom teaching, head teacher leadership, equity issues and transparency on education funding); and
- Key tasks across the policy cycle including budgeting, planning (including prioritisation of need, operational planning, budgeting and execution) monitoring (including data collection, management, analysis and interpretation), and evaluation.

The desired level of future capacity is expressed in targets in the EQUIP-T Inception Report (Annex XVIII) as:

- *'[Districts] manage all aspects of their work at the District level more effectively and efficiently, taking full responsibility for the planning, coordination, delivery and monitoring of all their activities and systems'.*
- *'...successful Districts will manage EQUIP-T funds'.*
- Districts will be able to provide support to SCs and schools with tools and skills to be able to develop an informed prioritised budget, and record and report on school financial matters.
- All capitation grants received at the district level will be disbursed in full and on time.

In order to provide a picture of the baseline capacities in which the EQUIP-T programme aims to affect change, the qualitative study reports on existing capacity levels against these key gaps, complemented by quantitative data from the survey in selected areas.

7.2 Physical resources

Overwhelmingly, REOs, DEOs and WECs⁷⁵ cited insufficient resources as the biggest challenge to fulfilling their responsibilities. Participants were probed several times within interviews to consider challenges in addition to resource constraints, but the challenges WECs and DEOs face were always cited as resource constraints, either implicitly or explicitly. Some DEOs and WECs focused on the resources required within their area of work and others focused on their own personal circumstances and the need for increased resources to improve their job performance (citing transport from home to work and funding for lunches, etc.). The resources REOs, DEOs and WECs perceive as being required in order to do their jobs include funds to support transportation and teaching and learning materials. For WECs, office space, housing, desks and stationery were also cited as necessary requirements to fulfil their responsibilities. In addition, WECs and DEOs perceived the sharing of offices between WECs and head teachers, the school office, village leaders and ward executive officers as key challenges.

The qualitative study found that school-level stakeholders rarely perceived the resourcing of education officials and offices at the district and ward level as having any tangible connection with education delivery, with the exception of WEC transport. While the IE survey found that almost all schools (99%) were visited by WECs in 2013 (see Section 7.4.1), head teachers, community leaders, WECs and DEOs all agreed that there are limited funds for WEC transport to schools. WECs stated that transportation to schools was covered through their own private funds, whereas several schools stated that WEC visits are funded from the school budget or IGAs and indicated a preference for this to come from the government.

School-level stakeholders indicated a better relationship with WECs who had the means to visit schools without relying on funding from the school budget. Additionally, in schools with weak school leadership, teachers viewed the WEC as their first line of accountability for attendance. This indicates regular WEC visits are important to manage teachers in the short term, but also to provide capacity building in school leadership to head teachers until such time as head teachers are able to manage teachers at the school level.

Education officials at the regional, district and ward levels perceived an increase in resources as the primary factor that would improve education delivery in the districts. However, school and community stakeholders tended to provide less tangible political motivations as explanations for poor support for education at the regional, district and ward levels. Such explanations included education officials seeking allowances, attending their own training and events, and maintaining the hierarchy rather than focusing on education service delivery. As the EQUIP-T Inception Report states, district offices sit at an interface between educational activity and political interest. Non-technical (and non-resource) dimensions of change such as attitude and motivation and cultural and political interests are forces that can facilitate or impede change. Therefore, the qualitative study confirms the EQUIP-T diagnosis that resources to support the regional and district offices are required, in addition to other non-material changes.

7.3 Transparency and information on key financial and budgetary procedures

Political or vested interests disrupting the use of resources and weakening the education planning cycle was cited by officials at the regional, district and ward level. Officials did not discuss the

⁷⁵ Since the baseline research took place, WECs have been retitled as Ward Education Officers.

releasing of capitation grants in full and on time as a capacity problem; instead, the diverting of education funds to other areas, a lack of accountability, and vested interests were discussed. Examples included directing the funding of education activities to other areas as the ‘District Planning Officer doesn’t feel that there is a high need to execute something’ (REO) or the prioritisation of school visits over other problems in order to claim per diems. Several stakeholders highlighted the use of education funding for non-education activities, with one stakeholder emphasising the need to have ‘in place a system that ensures that all the money that is allocated to the education sector is used for education purposes only’.

Teachers and head teachers at a number of school discussed having to attend the District Office to request salary payments regularly, often filling in the same forms several times. Teachers’ explanations for late payment were typically linked to a lack of transparency at the district level. Head teachers tended to avoid providing explanations but noted the loss of time there could be at the school and the decrease in people’s effective salaries as a result of these inefficiencies. As has been discussed in the Teacher and School Leadership and Management component sections of this report, delays in payment and multiple visits to the District Office leads to low teacher morale, a loss of instruction time and a loss of head teacher management time at the school. Many teachers and head teachers also perceived such occurrences as a lack of respect for the profession of teaching.

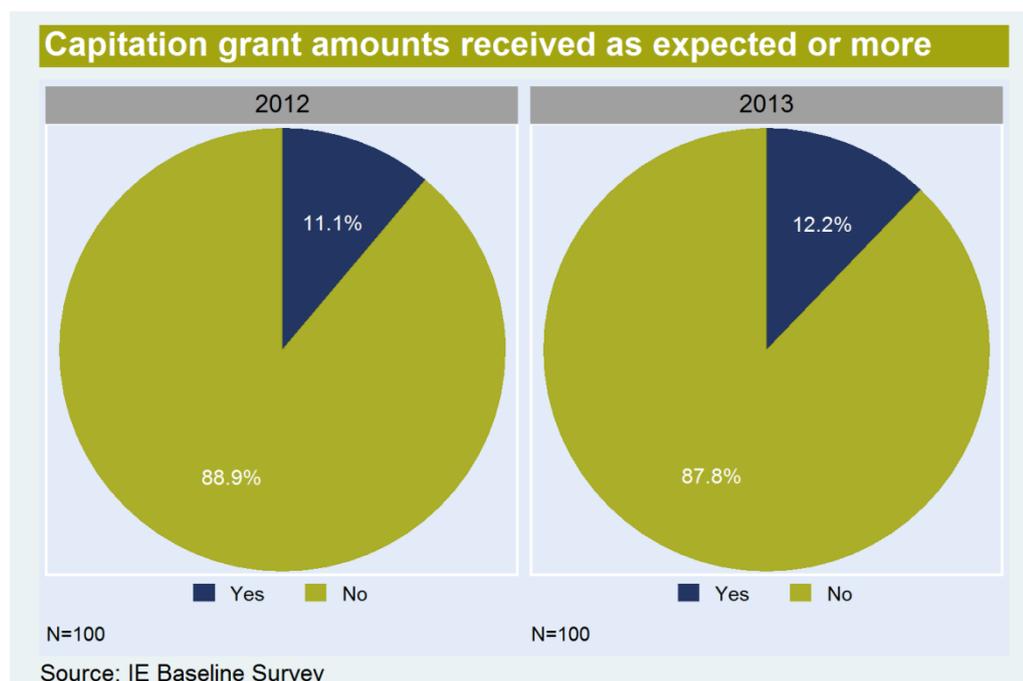
7.3.1 Capitation grants

A key constraint on district planning and management of education that was identified during the design of the EQUIP-T programme was the inconsistent and unpredictable flow of funds from central to district level to schools, resulting in a funding situation for schools that is also inconsistent and unpredictable. To examine the conditions prior to the start of the programme, the IE baseline survey collected data on capitation grants received by schools through head teacher interviews and the checking of school financial records.

Schools are officially supposed to receive TZS 10,000 per primary pupil in the form of capitation grants. However, on average head teachers in treatment districts expected to receive around TZS 8,400 per pupil annually – but only received around TZS 3,500 in 2012 and TZS 3,800 in 2013.⁷⁶ Moreover, just 11-12% of schools reported receiving capitation grant amounts that met (or exceeded) the expected amounts (see Figure 41). It should be noted that these estimates are indicative only, as records on capitation grant amounts, and in particular date received, were frequently incomplete.

Other studies also find that capitation grants received tend to be lower than intended: in 2007/08, primary schools were estimated to receive on average about TZS 4,200 per pupil (URT 2010) and in 2013 roughly TZS 6,000 (DataVision 2013) compared to an expected amount of TZS 10,000. These estimates are not directly comparable to those of the IE baseline survey, but together suggest that the capitation grants schools received tend to be substantially lower than intended.

⁷⁶ The estimates were obtained by asking for the total amounts of capitation grants schools received for 2012 and for 2013. This amount was then divided by the number of pupils enrolled. As enrolment data from 2012 and 2013 were not available, enrolment data from 2014 were used as a proxy for school size.

Figure 41 Amount of per pupil capitation grant payments received meets expectation

Note: (1) Weighted estimates; (2) Expected capitation grants reported by head teachers as per pupil amounts; (3) Total capitation grant amounts received reported by head teachers with supporting financial documentation verified by enumerators; (4) Capitation grant amounts received calculated as per pupil amounts using enrolment numbers from 2014.

Within schools and communities, frustrations with the capitation grant were predominately related to low and late release, but the qualitative study also found that other frustrations were also present. Several teachers and head teachers found the approval process for capitation purchases frustrating because they perceived the needs of the school to be best known by those present at the school. This was discussed in relation to having requested purchases for teaching and learning materials denied by DEOs. Other teachers expressed the same frustration about seeking SC approval for budgets in cases where the requests included purchases the community had not previously encountered, such as flip charts.

The lack of basic teaching and learning materials in schools was primarily linked by school and community stakeholders to the lack of capitation funding received by schools. The majority of schools felt that the extremely low capitation rates were not enough to buy basics such as chalk, pens, pencils and exercise books. This has led to conflicts with communities who believe the school should be supplying these materials from the capitation funds.

The late and reduced-level release of capitation funding also affected the ability of schools to raise funding through IGAs in several schools, as head teachers planned to engage the school in IGAs but missed investment opportunities (or seasonal opportunities for agricultural activities) due to the late release of funds.

Several head teachers discussed running the school with their own funds while awaiting the disbursement of late capitation.

7.4 Support systems and school improvement systems

Support systems include the systems that support effective coordination and monitoring of key activities such as teacher and head teacher training. School improvement systems include school–

community participation mechanisms, classroom teaching, head teacher leadership, equity issues and transparency on education funding.

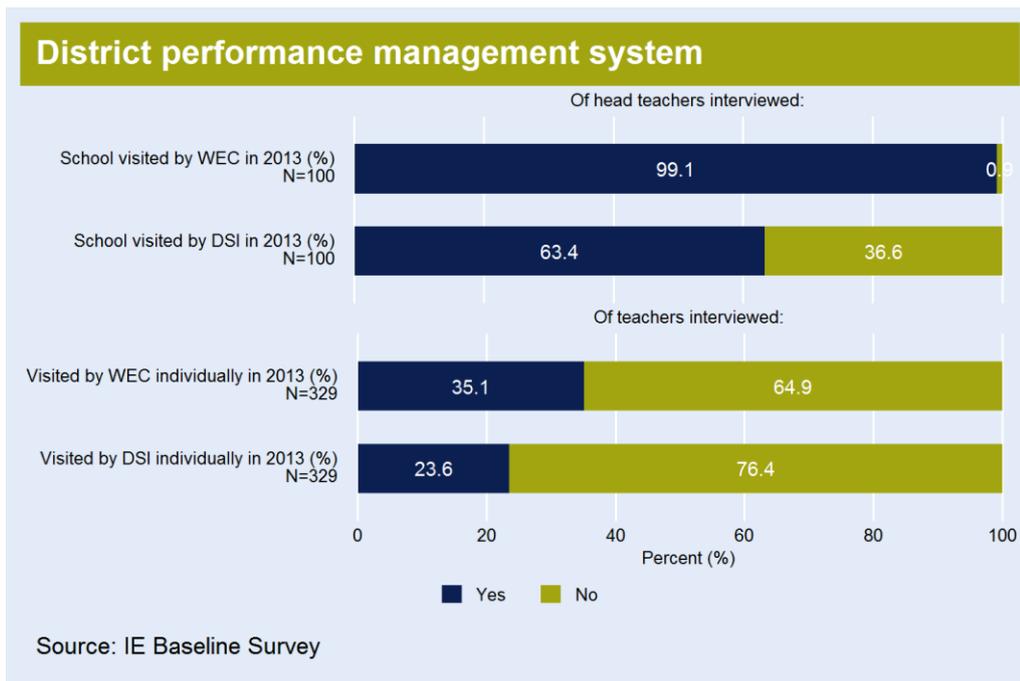
The extent to which district offices provide support to school improvement systems varied across the case study schools. There were no identifiable patterns within districts, indicating that the extent to which WECs provide support to school improvement systems is primarily driven by the individual. Several schools reported conflict with the District Office, perceiving that until the village committee requested action to be taken the district was slow to respond. One community in Uyui, Tabora expressed their concern regarding the extreme tensions in the school as a result of two head teachers being appointed to the school. The community stated that they had written to the DEO for two years regarding the tension, with no resolution.

7.4.1 Ward Education Coordinator and District School Inspector visits

School inspectors are expected to assess schools on curriculum implementation, teaching and learning resources, and all aspects that contribute to a child-friendly learning environment (MoEVT 2009b). Inspectors are required to observe teachers and pupils at work, assess whether learning is participatory, and identify issues and teacher training needs, among other things. The school inspector's handbook recommends that every school be inspected once per academic year, although 'the frequency is to be dictated by the availability of funds' (MoEVT 2009b, p. 22).

WECs, meanwhile, are expected to work closely with head teachers to provide continuous external supervision and coordinate the formulation of WSDPs. The aim of this supervisory relationship is to empower schools to be more responsible in managing administrative and pedagogical issues (MoEVT 2009c). School supervision is focused on overseeing the work of the school and providing professional guidance and advice to teachers. WECs are then expected to provide written reports to the District Education Office and DSIs.

The EQUIP-T IE survey found that almost all schools (99%) were visited by WECs in 2013 and just under one-third (35%) of teachers were visited individually by WECs to observe their teaching or hold a meeting or discussion with them. The comparable figures for DSIs are lower, with two-thirds of schools (63%) and less than one-quarter of teachers (24%) visited by them in 2013.

Figure 42 District visits to school and individual meetings with teachers

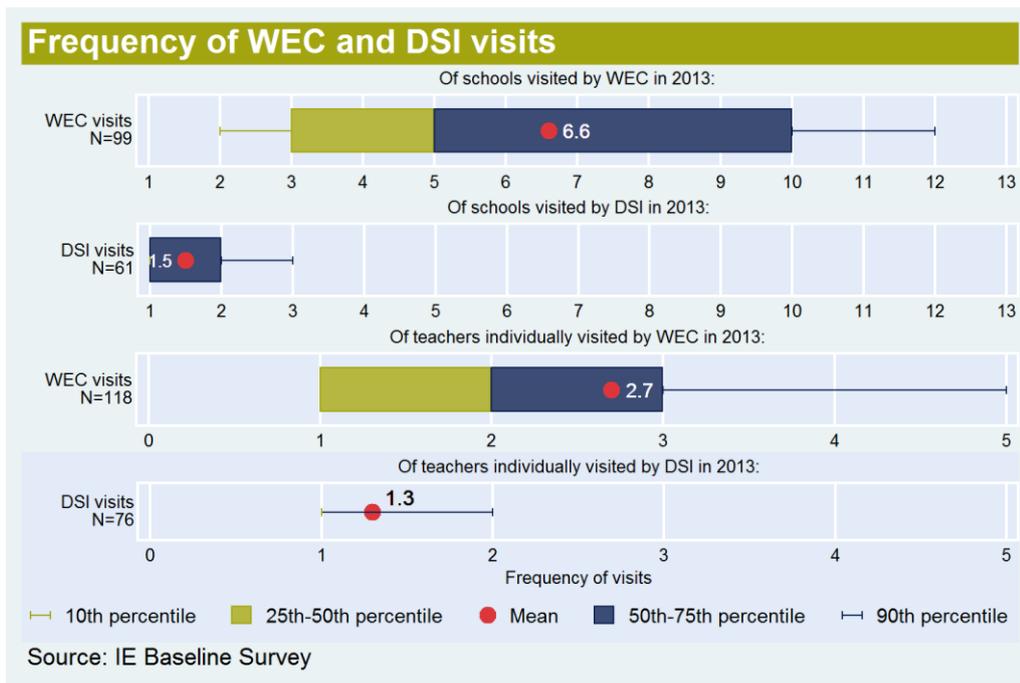
Notes: (1) Weighted estimates.

Schools that were visited by the WEC were visited an average of 6.6 times in 2013 (see Figure 43). The frequency of visits varies considerably between schools, with head teachers in the middle 50% of schools in terms of WEC visits reporting being visited three to 10 times. Teachers who were visited individually by WECs reported being visited an average of 2.7 times. Visits by the DSI averaged 1.5 times for those schools that were visited, and 1.3 times for those teachers that received individual visits.

The qualitative study found that while WECs perceived their roles to include providing academic support to teachers and support to school leaders, the majority of teachers and head teachers felt the lack of provision of academic and leadership support was a primary factor in decreasing levels of education service delivery. The main roles and responsibilities cited by school and community stakeholders regarding WEC support for school improvement referred to disciplinary actions against absent teachers, engagement with the community regarding pupil absences, and dealing with community reports of difficulties at the school.

Engagement with the WEC varied greatly between case study schools. In an extreme case, the WEC was reported to have not visited the school in over two years. Within this school the head teacher perceived the school to be in opposition to the community and saw the role of head teacher as being only to manage pupil attendance and classroom attendance. There was no focus on classroom teaching quality, equity or community participation in the school. However, in other schools the WEC visited regularly there was no evidence that the head teacher and teachers necessarily functioned more effectively in their roles. This indicates that while transport for WEC visits is extremely important, the enacting of the roles and responsibilities of the WEC and the corresponding enacting of the roles and responsibilities of a head teacher do not necessarily take place purely on the basis of frequent WEC visits.

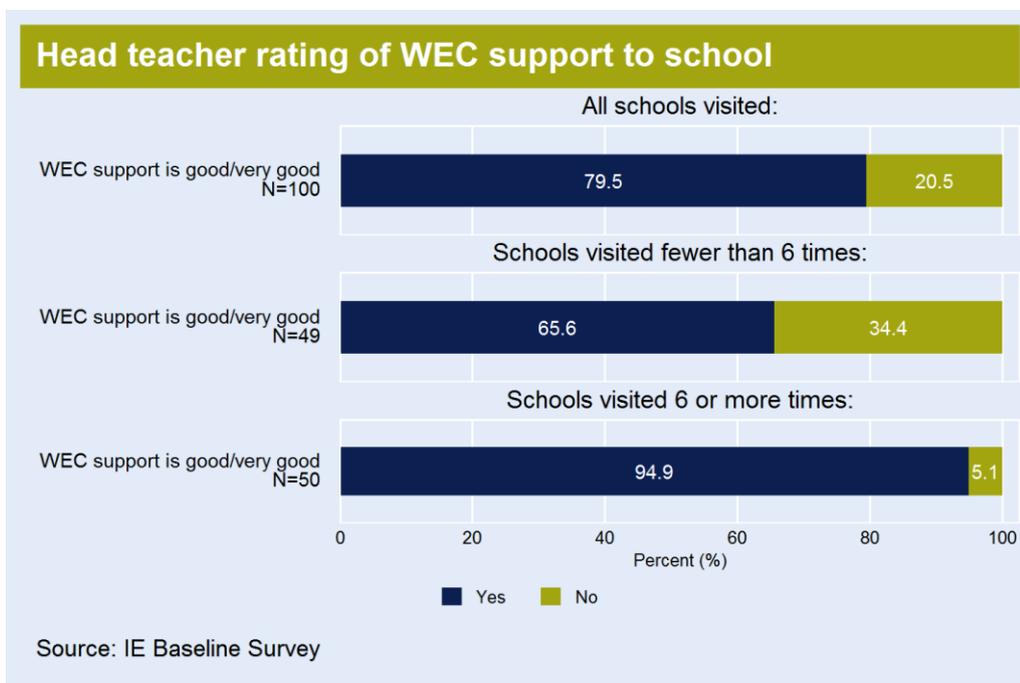
Figure 43 Frequency of district visits to school and meetings with teachers



Notes: (1) Weighted estimates.

As part of the EQUIP-T IE survey head teachers were asked how they rate the support of WEC to their school, from very poor to very good. Overall, 80% of head teachers in the programme treatment districts rated WEC support as either good or very good (see Figure 44).

Figure 44 Head teacher rating of WEC support to school, by frequency of visit



Notes: (1) Weighted estimates.

The EQUIP-T IE survey found that head teachers' perceptions vary by how often their school was visited by the WEC the previous year. Among schools that were visited fewer than six times in

2012,⁷⁷ 66% of head teachers rated WEC support as good or very good. Meanwhile, almost all (95%) head teachers in schools that were visited six times or more gave positive ratings. Conceptually, this was supported through the qualitative study, whereby head teachers who reported positive support from the WEC were from schools visited more frequently by the WEC. However, while the IE survey presents a predominantly positive picture of head teachers' ratings of WEC support, further probing through the qualitative study finds that head teachers are able to identify significant gaps in WEC support. In some schools these gaps were directly attributed to gaps in WEC support, but in the majority of schools head teachers were able to identify gaps but did not directly attribute these gaps as being related to the role of the WEC.

Both the IE survey and the qualitative study find that WECs are clearly active; however, the extent to which the number of WEC visits meets head teacher expectations, the funding sources of per diems for WEC visits, and the extent to which WECs fulfil the full spectrum of tasks associated with their roles is unclear.

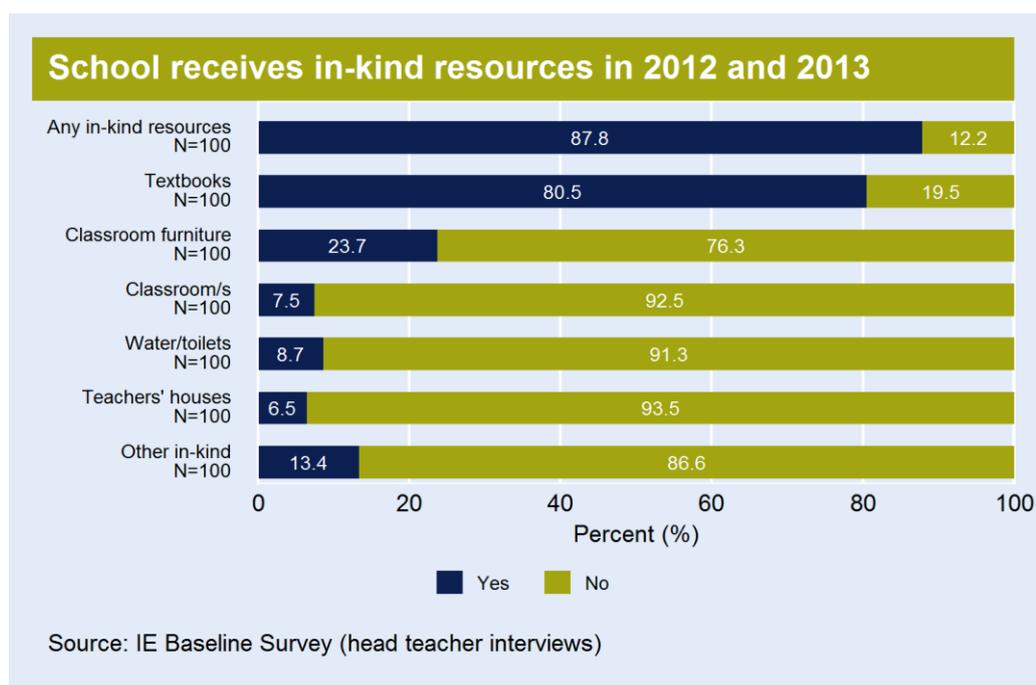
7.4.2 In-kind resources received by schools

In the survey interview with head teachers, they were asked whether the schools received any in-kind resources in the calendar years 2012 and 2013, such as textbooks, classroom furniture or new infrastructure. Where head teachers reported that in-kind resources were received, they were asked what these were and which agency/organisation provided them. It must be noted that the in-kind resources identified are unlikely to have been purchased by schools with capitation grant funds received as the head teachers were specifically asked not to include in-kind resources purchased with per capita grants in their response.

As shown in Figure 45 the majority of schools are receiving in-kind resources, which are typically textbooks. About 88% of schools in treatment districts were identified as having received some form of in-kind resource in 2012 and 2013. With respect to textbooks specifically, over 80% of schools in treatment districts identified that these were received in 2012 or 2013. Among schools receiving textbooks, 90% of head teachers reported that they had been received from the Tanzanian government at either local district/ward level or regional/central government.⁷⁸

⁷⁷ The median number of WEC visits in 2013, i.e. six visits, was used as a cut-off.

⁷⁸ During mid-2013-2014 there was a mass distribution of primary textbooks across the whole country financed from a large one-off refund which Government of Tanzania received. This may partly account for the high receipt rate of textbooks reported by schools in 2014.

Figure 45 Proportion of schools receiving different types of in-kind resources

(1) Weighted estimates.

The district/ward level of government, specifically, plays a considerable role in the provision of textbooks to schools. Among the 82 treatment schools that reported receiving textbooks, 48% of these schools received them from their district or ward offices.⁷⁹ Classroom furniture was the next most common in-kind resource received (as reported by 24% of head teachers) and was most commonly provided by district or ward offices.⁸⁰ The receipt of infrastructural resources such as classrooms, water/toilets or teachers' houses was limited. Less than 8% of schools received classrooms, 7% of schools received water/toilets and 7% received teachers' houses. When major infrastructural resources (i.e. classrooms or teachers' houses) were provided, these were typically from the Tanzanian government.

The qualitative study found that major impediments to school improvement included a lack of promised funds for improvement projects. Two communities indicated that they had started major projects through parental contributions, with the promise of future resourcing to complete the projects from the District Office. However, in both cases the community stated that the funds were not forthcoming and the projects remained incomplete.

One school acknowledged in-kind support from the District Office and the additional supply of teachers when requested, but most school- and community-level stakeholders did not perceive the Region or District Office as providing support to school improvement.

7.5 Key tasks across the policy cycle

Most stakeholders identified weakness across all stages of the education planning cycle. Stakeholders discussed the difficulties experienced in planning education due to very low levels of

⁷⁹ While the majority of schools are receiving textbooks, the survey did not probe head teachers on what number of textbooks were received in relation to enrolment numbers or about timings and delays associated with their receipt. The SDIS report prepared on Tanzania in 2012, for example, indicates that primary schools in urban and rural areas had just less than one textbook per pupil (World Bank 2012).

⁸⁰ 61% of the 24 schools receiving classroom furniture indicated that it came from local government.

data quality in the education sector. Implementation effectiveness was perceived to be low due to incomplete and late disbursement of funds, misuse of resources, disruption of funding allocated to education due to political interests and a lack of personnel and material resources, especially in newly divided districts. Monitoring effectiveness was perceived to be low due to a lack of funds to undertake school visits.

Strengths identified by officials tended to be focused on specific events or people, rather than systematic strengths. Examples include working with a supportive Regional Commissioner on education planning, planning district sporting events, and providing teachers for unregistered schools when there is availability.

The limited availability of data to inform education planning was perceived as a major constraint. The reliability of data was questioned by many stakeholder groups, and the resources and infrastructure available to gather and verify data was identified as the main constraint to education planning. An EQUIP-T staff member stated: 'The wards, the districts are looking down at all of the people below, trying to understand the reporting. Which schools need more focus? At each level the information is needed for proper planning and allocation of resources'. This was verified by REOs, DEOs and WECs. An REO for a poorer performing region stated that, 'If the EMIS could be in place strongly... sometimes if you ask for the same data today and tomorrow, from a different person, you get different data, but for the same region'.

These responses indicate that REOs, DEOs and WECs are familiar with the process of using EMIS data for planning purposes, but do not currently trust the available data and find it difficult to verify due to limited resources. In addition, limited infrastructure frustrates the process, with WECs unable to access mobile reception at many schools.

PART C: Expanding and assessing the EQUIP-T theory of change

8 The programme theory of change: relevance, causal mechanisms and underlying assumptions

The EQUIP-T programme TOC, (see Annex B) identifies six groups of constraints acting on pupils' capability to learn to their full potential. The programme's overarching theory is that by reducing or removing these constraints the quality of education and pupil learning in Tanzania will improve. This part of the report assesses the EQUIP-T programme TOC drawing on the qualitative and quantitative IE baseline results as well as Tanzanian and international evidence.

The starting point for this analysis was to expand the programme TOC by explicitly mapping out the key causal pathways through which the programme expects to see change. This process makes the numerous links between different causal pathways explicit, which is a useful process in itself as it can help to think about the different interventions more holistically. The next step was to assess the strength of the key assumptions underpinning each link. This was done by reviewing the wider literature base to see if there is evidence that the particular assumption about anticipated change holds in principle (i.e. evidence suggests that interventions of a particular type can lead to the change anticipated, under various context and design conditions), if the contextual conditions for anticipated change are present in the treatment areas and assessing if the change seems to be needed to improve learning in the EQUIP-T treatment areas.

One of the aims of this analysis, as described in the inception report, is to provide information to the programme on areas of the TOC where further investigation could be useful. There are, of course, clear limitations to this analysis, including that it has necessarily been done at a high level with respect to the intervention type/design. This means that the detailed design of the interventions may already mitigate some of the factors that appear to suggest that a particular link may be comparatively weak, for example. Moreover, in drawing on a wide evidence base, only some of the research shows causal effects based on a randomised control trial or quasi-experimental design; more commonly the studies are based on other standard research techniques and rely on associations or patterns of evidence to draw conclusions.

This chapter has five analytical sections, covering the key programme impact (improved pupil learning) and each of the four programme components covered by the IE.⁸¹ Each section summarises the causal mechanisms from the expanded TOC, and then assesses the assumptions underpinning the causal links which make up the overall mechanism. The final section presents the whole expanded TOC, and highlights the key findings from this chapter. It also notes that the budget for some components of the programme is far higher than for others. When there is a large planned investment, and the TOC analysis raises some concern about the strength of the assumptions underpinning the related causal pathways, this clearly merits further reflection on the design.

⁸¹ Component 1: Teacher professional capacity, performance, motivation and morale; Component 2: School leadership and management; Component 3: Community participation and accountability; and Component 4: District and regional education management.

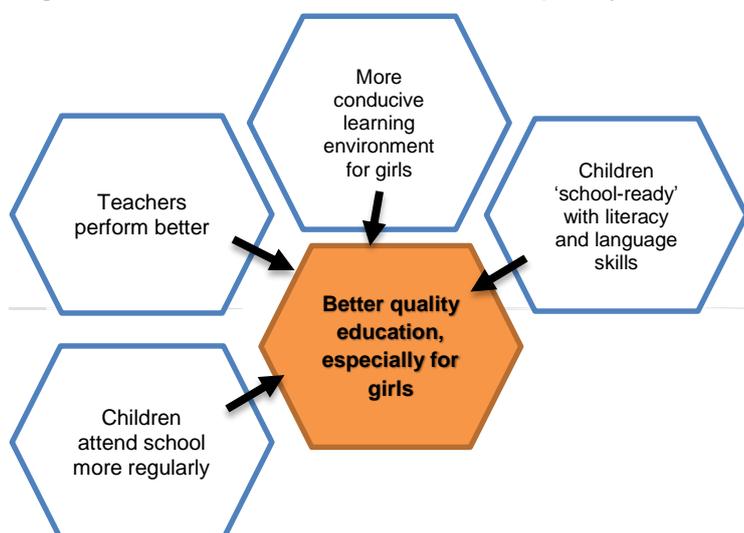
8.1 EQUIP-T Impact: Improving learning outcomes and education quality, especially for girls

The ultimate impact of the EQUIP-T programme is improved learning outcomes and education quality, especially for girls.

Causal mechanism

The central causal mechanism of the expanded TOC can be summarised as follows: There will be better quality education in Tanzania⁸², especially for girls, because children will be school-ready when they begin school, there will be a more conducive learning environment for girls, teachers will perform better, and children will attend school more regularly (Figure 46). This mechanism is made up of four direct links. As with all of the causal pathways, each link is reliant on the success of other links further up the causal pathway.

Figure 46 Causal mechanism: Better quality education, especially for girls



Assessing the assumptions underpinning the causal pathways

The links that contribute to this mechanism are represented as arrows in Figure 46. Each arrow is underpinned by a range of implicit and explicit assumptions. For example, for 'school-ready' children with literacy and language skills to lead to better quality education, especially for girls, it is assumed that pupils (especially girls) are not currently school-ready, and that school-readiness programmes can improve education quality in principle.⁸³ As an overview of results, Table 9 summarises the causal links and the strength of the assumptions underpinning them based on the analysis which follows.

⁸² In this report, the term 'better quality education' is synonymous with improving learning outcomes for pupils.

⁸³ Under certain contextual and intervention design assumptions.

Table 9 Summary of strength of assumptions underpinning the causal links

Better quality education, especially for girls	Strength of assumptions underpinning causal links		
	Strong	Weak	Mixed/Contextual
Links			
1a: 'School-ready' children with literacy and language skills will lead to better quality education, especially for girls	X		
1b: A more conducive learning environment for girls will lead to better quality education, especially for girls	X		
1c: Better performing teachers will lead to better quality education, especially for girls	X		
1d: Children attending school more regularly will lead to better quality education, especially for girls	X		
<p>Strong: The wider literature base and contextual data from the IE baseline survey provide substantial evidence that the main assumptions underpinning the link are likely to hold.</p> <p>Weak: The wider literature and contextual data from the IE baseline survey provide little to no evidence that the main assumptions underpinning the link are likely to hold.</p> <p>Mixed/contextual: There is some evidence from the wider literature and the contextual data from the IE baseline survey that the main assumptions underpinning the link are likely to hold. However, the assumptions may only hold under certain conditions and may therefore lead to heterogeneous results across treatment districts.</p>			

Mechanism 1: There will be better quality education in Tanzania, especially for girls because children will be school-ready when they begin school, there will be a more conducive learning environment for girls, teachers will perform better and children will attend school more regularly

Link 1a: 'School-ready' children with literacy and language skills will lead to better quality education, especially for girls

There is substantial evidence globally that being school-ready is one of the strongest predictors of later academic achievement (Duncan et al. 2008⁸⁴). Reading and literacy skills at the time of school entry have been found to be the second strongest predictor of later achievement (after mathematics skills at the time of school entry), and have been found to predict later success to a greater extent than attention skills and socio-emotional behaviours (Romano et al. 2010⁸⁵). These patterns are observed similarly for boys and girls and for children from high and low socioeconomic backgrounds. While some African countries have undertaken educational quality assessments through early years assessments, longitudinal studies tracking educational achievement at the pupil level are not available. However, given the strong evidence base from studies such as TIMSS, PIRLS and SACMEQ regarding the strong associations between exposure to reading materials and academic success across a range of country contexts (including Africa), it seems reasonable to assume that these findings will hold across locations.

The IE baseline survey found that 77% of the pupils in the treatment schools speak a language other than Kiswahili at home. The overwhelming majority of the children in the EQUIP-T programme intervention areas are learning in a language they are unfamiliar with prior to entering the schooling environment. The IE baseline data indicates that speaking a language other than Kiswahili at home is strongly negatively correlated with pupil performance in Kiswahili literacy and mathematics.

Both the IE baseline survey results and the international evidence, therefore, provide strong evidence that school-readiness improves learning outcomes, and it is likely that this holds in

⁸⁴ Synthesis of evidence including experimental and quasi experimental studies and other studies

⁸⁵ Psychological study using replication and extension methods

EQUIP-T districts. However, there is no available evidence regarding differences in school readiness for boys and girls.

Link 1b: A more conducive learning environment for girls will lead to better quality education, especially for girls

A conducive learning environment for girls is characterised by access to adequate sanitation facilities, safety and security in and on the way to school and being able to participate effectively in class (FAWE 1998, IDP 2010). A conducive learning environment not only increases girls' access to school, but positively affects girls' experiences of schooling, which is likely to lead to more frequent attendance and greater engagement in the schooling environment (Dunne, M., & Leach, F. 2005).⁸⁶

The IE qualitative research indicates that one of the fundamental issues limiting girls' educational futures in EQUIP-T districts is pervasive sexual harassment experienced within the school or on the way to school. Community members in several case study schools raised this concern.

Community members also reported that some teachers use their daughters to undertake domestic chores at teachers' homes either during class or after school. Head teachers acknowledged that this can take place and standard three pupils in FGDs described 'fetching water for the teachers' as one of the activities that takes place at school for girls. In the most immediate sense, this reduces the time that girls are in class, however more broadly such actions reinforce social norms that view domestic activities as women's work, while boys gain access to learning opportunities within the classroom.

Across the case study schools, pupils and community members repeatedly discussed the role of corporal punishment in schools. All standard three pupils that participated in all nine schools agreed that 'being punished with sticks' was one of the aspects of school that they did not like. Studies show that discipline at school through either physical or emotional humiliation hinders a child's ability to learn, leads to higher incidence of truancy, and is linked to higher drop-out rates.⁸⁷

Within the classroom, the IE baseline survey found that a small majority of teachers practice gender-balanced interactions with pupils, but that 30% of teachers interact more with boys than girls, and 16% interact more with girls than boys.

Overall, the evidence from the IE research suggests that there is significant scope to improve the learning environment for girls in EQUIP-T districts. International evidence suggests that in general a more conducive learning environment for girls can lead to better quality education, as long as the school and the schooling system is performing at a level where if pupils attend school and participate in class, they will learn. In EQUIP-T schools, as we will see, this will require improvements in the school-readiness of children (pathway 1a) the performance of teachers (pathway 1c). If school-readiness and teacher performance improve, there is reasonably strong evidence that a more conducive learning environment for girls in EQUIP-T schools will lead to better education for girls.

Link 1c: Better performing teachers will lead to better quality education, especially for girls

International studies indicate that what teachers know, what they do and how much they care accounts for more variance in pupil achievement than any other policy amenable variable (Hattie

⁸⁶ See the work of the Independent Development Fund, the United Nations Girls Education Initiative for example

⁸⁷ See more at: <http://acei.org/corporal-punishment.html#sthash.Meu8X16k.dpuf>

2003⁸⁸). The IE survey finds that teachers in treatment schools vary considerably in their current levels of professional capacity and performance in the classroom, but that low teacher motivation and morale is pervasive. There is a clear need for major change in teacher behaviour, especially with respect to improving classroom attendance. On classroom teaching, overall girls are currently more disadvantaged than boys in interactions with the teacher, although this only applies to a minority of teachers.⁸⁹

On balance then, the EQUIP-T baseline study shows that there is a clear need for change as teacher performance is generally poor, and the international research suggests that in this type of context there is strong evidence that this assumption will hold.

Link 1d: Children attending school more regularly will lead to better quality education, especially for girls

The key assumption underpinning this causal link is that when children are in school, they are learning. Hence, similar to link 1b above this will rely on the success of the two other links directly linked to improving quality in the classroom (more school-ready pupils and better performing teachers).

The IE survey found pupil absenteeism is a serious problem. One-third of pupils (both boys and girls) in standards one to three absent on the day of the survey. Given this baseline situation, and that the evidence on the assumptions underpinning the two key links 1a and 1c was deemed to be strong, it is reasonable to infer that the evidence is strong that if children attend school more regularly, and if they are school ready and well taught, they will learn more.

8.2 EQUIP-T component 1: Improving teacher professional capacity, performance, motivation and morale

This section examines the causal links and assumptions which support the first component of EQUIP-T, which aims to effect change in three of the four areas discussed above: improving the performance of teachers in the classroom; ensuring that children are school-ready with adequate literacy and language skills; and providing a more conducive learning environment for girls.

Causal mechanisms

Three sets of causal mechanisms link EQUIP-T activities (in blue boxes) to the desired outcomes of i) improved school readiness, ii) more conducive learning environment for girls, and iii) improved teacher performance, and then to the eventual impact on learning (the orange box in Figure 47). These mechanisms all rely on sets of assumptions denoted by arrows in the figure. In EQUIP-T's management, these activities all fall within component 1 on improving teacher professional capacity, performance, motivation and morale.

The three causal mechanisms can be summarised as:

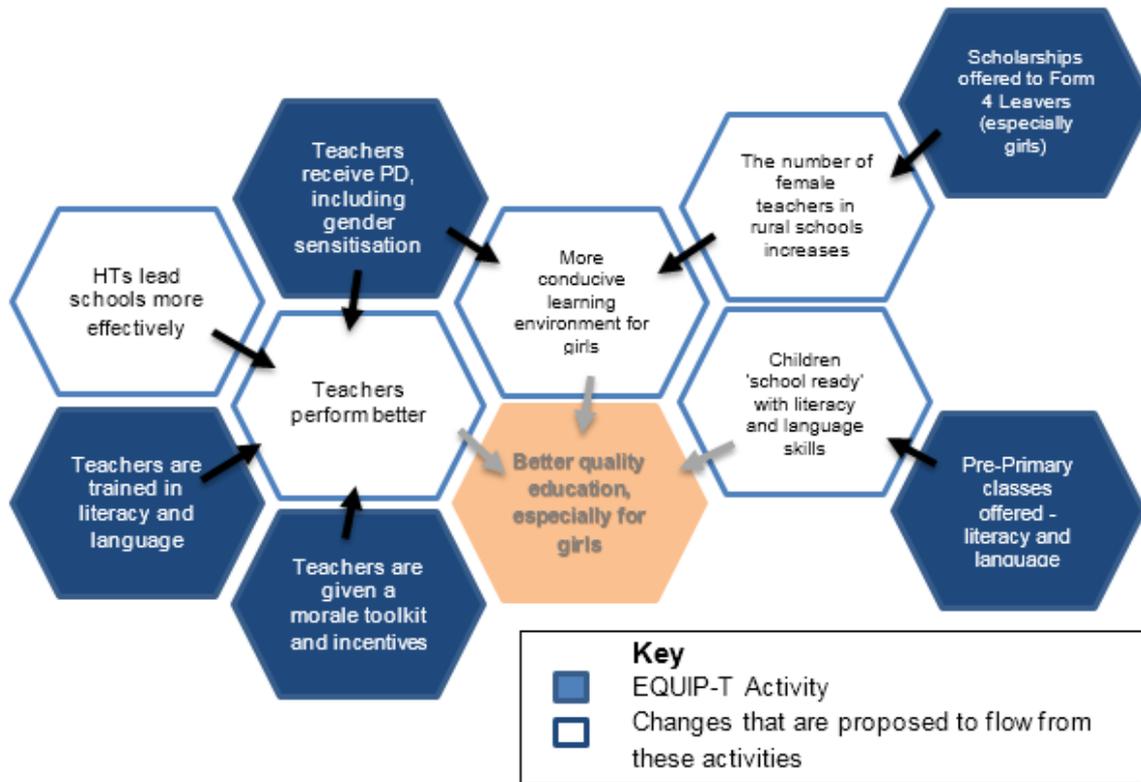
- Children will be 'school-ready' with literacy and language skills because pre-primary classes focussing on language skills will be provided;

⁸⁸ This study uses international pupil assessment data from multiple international assessment programmes to identify the predictors of academic success across locations.

⁸⁹ More details on finding from the IE baseline research on current weaknesses in professional capacity, performance and motivation and morale are given in the next section in relation to causal pathways underpinning EQUIP-T component 1.

- There will be a more conducive learning environment for girls because teachers will receive gender sensitisation and the number of female teachers in rural schools will increase after scholarships are offered to Form 4 female leavers; and
- Teachers will perform better because they will receive training (including gender sensitisation), a teacher morale toolkit will be introduced and incentives linked to a performance management system, they will be trained in literacy and language and head teachers will lead schools more effectively.

Figure 47 Causal mechanism: Improving school readiness, the learning environment for girls and teacher performance



Assessing the assumptions underpinning the causal links

The causal links that contribute to the three mechanisms expected to achieve the desired changes shown in Figure 47 are underpinned by a range of implicit and explicit assumptions as indicated by the arrows. A summary of the causal links and strength of the assumptions underpinning these are shown in Table 10 and are discussed in the following sections.

Table 10 Summary of causal links and strength of underpinning assumptions for EQUIP-T component one

Teachers will perform better	Strength of assumptions underpinning causal links		
	Strong	Weak	Mixed/Contextual
Causal links			
1a: Offering pre-primary classes in literacy and language will improve the 'school readiness' of pupils		X	
2a: The provision of gender sensitisation training to teachers will lead to a more conducive learning environment for girls			X
2b: Offering scholarship to Form 4 leavers, especially girls, will lead to an increase in the number of female teachers in rural school*		X	
2c: An increase in the number of female teachers in rural schools will lead to a more conducive learning environment for girls		X	
3a: Teachers will perform better because they will receive training	X		
3b: Teachers will perform better because they receive a teachers morale toolkit and incentives linked to a PMS			X
3c: Teachers will perform better because head teachers will lead schools more effectively.	X		
<p>Strong: The wider literature base and contextual data from the IE baseline survey provide substantial evidence that the main assumptions underpinning the link are likely to hold.</p> <p>Weak: The wider literature and contextual data from the IE baseline survey provide little to no evidence that the main assumptions underpinning the link are likely to hold.</p> <p>Mixed/contextual: There is some evidence from the wider literature and the contextual data from the IE baseline survey that the main assumptions underpinning the link are likely to hold. However, the assumptions may only hold under certain conditions and may therefore lead to heterogeneous results across treatment districts.</p>			

Mechanism 1: Children will be 'school-ready' with literacy and language skills because pre-primary classes focussing on language skills will be provided.

Link 1a: Offering pre-primary classes in literacy and language, children will come to school 'school-ready'.

This link is underpinned by the assumptions that pre-primary education increases school readiness and that households whose children are not currently school-ready will send their children to pre-school classes.

There is substantial evidence to support the assumption that pre-primary education increases school readiness in Africa and elsewhere (See UNICEF 2012 for a summary of studies and Nores & Barnett 2010⁹⁰). 'School readiness' is a contextual term, with multiple meanings depending on location, education jurisdiction and educational approach. Definitions tend to include cognitive skills, problem-solving abilities, motor skills and fine motor skills, physical readiness including nutrition and growth, and socio-emotional skills, such as resilience and behavioural skills (UNICEF 2012). In some locations, school readiness includes having acquired the required level of proficiency in the language of instruction, as is the case in the design of the EQUIP-T programme intervention.

A recent study in Mozambique (Martinez et al 2012⁹¹) found that children who participated in an early learning class experienced improvements along a number of child development outcomes

⁹⁰ The authors include a total of 56 studies assessing the effectiveness of 30 different interventions in 24 countries across Africa, Latin America and the Caribbean, Europe, Asia and the Pacific. The review includes randomised trials and quasi-experimental designs using difference-in-difference, propensity score matching, instrumental variables or other simultaneous equations estimation techniques.

⁹¹ This study used a RCT.

including in cognitive and problem-solving abilities, fine motor skills, socio-emotional and behavioural outcomes. However the evaluators did not observe significantly different gains in language and communication between the control and treatment groups, with levels remaining alarmingly low for both groups. It is also worth noting that in this study, a lot of technical support was offered by supporting NGOs to ensure the quality of the pre-primary classes. There is limited evidence that this level of technical support is available for pre-school teachers within the EQUIP-T programme.

A methodologically robust longitudinal RCT in Mexico (Fernald, Gertler and Neufeld 2009) explored the effects of early enrolment on behaviour, child growth, cognition and language. The study found that early enrolment reduced behavioural problems for all children enrolled early but did not have any detectable effects in child growth, cognition or language.

A review of recent DFID Education Rigorous Literature Reviews (Westbrook et al, 2013, and Nag et al, 2014) found some examples of successful early years programmes. These include a ten week summer preschool intervention offered to Kurdish-Turkish bilingual children from low income homes, which showed positive effects of school readiness including cognitive skills, oral language, socio-emotional development and physical competencies. Children showed gains on measures of emergent literacy, syntactic knowledge (grammar) and story comprehension, as compared with control children. The interpretation of the positive findings from these studies is, however, not straightforward. Nag et al (2014, Section 8) highlighted the probability of a ‘Hawthorne effect’ influencing pupil performance⁹², and also warned that positive effects are often not able to be observed over the long term. They summarise that ‘collectively, the interventions that have been evaluated provide limited evidence of how the impact of the intervention mediates (causes to change) children’s attainments’.

Westbrook et al (2013) also conclude that there is little evidence regarding the optimal intensity and duration of early literacy and language interventions, given the ‘extant research’ available. However the key contextual and programme characteristics of successful programmes, and the likely barriers have been identified through several DFID literature reviews. In order for sustainable long-term improvements in literacy and language the evidence suggests the following are required (Westbrook et al, 2013):

- A cadre of reading specialists to provide pre-service and in-service training. Such specialists should be familiar with second language methods of reading, techniques in intensive reading, familiarity with teaching reading in local schools and an understanding of the application of reading theory to local conditions.
- Teachers having a positive attitude toward their training and their students.
- Teachers being communicative with students.
- Teachers giving feedback and paying sustained and inclusive attention to all students.
- Teachers drawing on student’s backgrounds in their pedagogic discourse.
- Code switching⁹³ and the use of local language.

Barriers to literacy and language gains from two DFID Education Research papers (DFID, 1998 and Westbrook et al, 2013) included:

⁹² The ‘Hawthorne effect’ is also known as the ‘observer effect’ and is the tendency of some people to perform better when they are being observed, as part of an experiment or study. In this context, Nag et al (2014) note: ‘that if an intervention provides extra attention or a new direction for teaching – and if this is not controlled for in the design of the RCT or QED – it can be expected to have an effect on pupils’ performance’

⁹³ Code-switching occurs when a speaker alternates between two or more language in the context of a single conversation. Code-switching relies of both communicators to be fluent to more than one language (Myers-Scotten 1989)

- A lack of training of knowledge for teachers to deal with heterogeneous groups.
- Teachers taking a negative view of students from particular backgrounds.
- ‘Linguistic schizophrenia’ in classrooms, where reading and writing occurs in the second language, while oral instructions takes place in the first language.
- The use of unfamiliar language.

As discussed in the previous section, the IE survey finds that the majority of pupils are not school-ready at least with respect to the language of instruction when they enter school, because they do not speak Kiswahili at home. The survey also found that teachers rarely language-switch in class to assist pupils in understanding lessons.

Despite this apparently obvious need to improve school readiness for the majority of children, the IE qualitative research found a different perception among parents. Many parents across case study schools stated that language readiness was not a concern for them as their children learn Kiswahili at school. This indicates that at least some parents do not perceive pre-primary as relevant. Clearly this is likely to affect their willingness to send their children to pre-school classes if given the opportunity. Indeed the IE qualitative study found little evidence that parents would be willing to do this. Some case study schools either offered pre-school classes or had previously offered pre-school classes. Those that had closed down their pre-school classes did so due to limited classroom availability, reliance on volunteers and low pupil attendance. The IE survey found that current classroom ratios are 78 pupils per classroom, further supporting the limited classroom availability within schools. Parents of pupils in these schools indicated that they would be unable to send small children on the long walks to the school and that the costs of pre-primary are unaffordable.

Overall, although the wider evidence suggests that pre-primary education could increase school readiness in general, the evidence that pre-school classes would sustainably increase language acquisition for non-Kiswahili speaking pupils is weak. There is not sufficient evidence regarding the duration and intensity of intervention required for pre-school classes to be effective on measures of literacy and language. In addition, the extent to which there is an adequate cadre of specialists to foster the development of pre-school teachers in literacy, language and pedagogy skills at scale is also unknown at this time. In addition, IE qualitative research suggests that the assumption that that communities and parents will send their children to pre-primary education classes if offered may not hold.

Mechanism 2: There will be a more conducive learning environment for girls because teachers will receive gender sensitisation and the number of female teachers in rural schools will increase after scholarships are offered to Form 4 female leavers.

This mechanism is supported by three causal links and the assumptions for each of these are discussed below.

Link 2a: The provision of gender sensitisation training to teachers will lead to a more conducive learning environment for girls

For the provision of gender sensitisation training to teachers to lead to a more conducive learning environment for girls, it is assumed that gender sensitisation works at scale and that male and female teachers are willing and able to change their gendered attitudes, values and behaviours towards (boy and girl) children.

Evidence suggests that gender sensitisation is more likely to work when training is participatory, allowing learners to revise their presumptions about cultural expectations. In one study in Zambia,

when learning was information based, where trainers merely deliver information stating that girls and boys are equal, participants were found to be unable to recall the subject matter (Evans 2014). Other success factors include facilitators being committed to gender equality and when a critical mass of 'neighbours' also change their views. Evans finds that gender sensitisation at scale often has limited effects as recruiting and mobilising trainers with a commitment to gender equality at scale is particularly difficult.

Therefore, although the IE survey and qualitative research found evidence of gender bias in the school and classroom environment (cited in the previous section under link 1b), the extent to which the assumption that gender sensitisation training at scale will change teachers' behaviour is mixed/contextual.

Link 2b: Offering scholarship to Form 4 leavers to undertake teacher training, especially girls, will lead to an increase in the number of female teachers in rural schools

This causal link is underpinned by three assumptions. First, that a sufficient number of female pupils from programme districts will complete Form 4. Second, that Form 4 leavers will take up scholarship offers, complete a teacher training course and be recruited for rural posts. Third, that these female teacher training graduates are able and willing to live and work in rural locations after completing their training.

In relation to the first assumption, the number of male and female pupils from government schools within the programme districts registering for Certificate of Secondary Education Examinations (CSEE) in 2011 was extremely low. The CSEE is open to any pupils who have completed four years of secondary education at both government and non-government registered schools and passed Form two secondary education examinations. The examination identifies pupils with the capacity to continue learning to the advanced secondary school level and other learning institutions. Therefore, in order for this assumption to hold, the number of girls in government schools in the programme areas would have to increase, and/or girls from non-government schools would need to be targeted for the scholarship programme. It seems unlikely that the pull-effect of the scholarship programme would be sufficient to boost numbers in government schools, and it is unclear how the programme could affect non-government school numbers, and therefore this assumption is weak.

The second assumption is that Form 4 leavers will take up scholarship offers, complete a teacher training course and be recruited for rural posts. Evidence for this is quite weak. A multi-country study focussing on strategies to increase the number of female teachers in rural schools found that the development of facilities for secondary education for girls in rural areas and long-term teacher education training programmes specifically designed for less qualified rural girls are required (UNESCO PROAP, 2000). The study found several limitations to increasing the number of rural female teachers including rural women's inability to compete with urban males in recruitment processes and the low academic performance of rural women in teacher training institutions. In Nepal, a quota system was introduced, whereby 60 percent of new candidates were required to be female, yet the quota remained unfulfilled as women failed to be recruited. The study found that pre-service teacher training did not equip women teachers with the skills they need to cope with actual classroom problems in rural areas (including large classes, multi-grade classrooms and classroom management), leading to retention issues. The study found that pre-service training lacked contextual relevance for rural women deployed to rural schools.

The IE qualitative study found that teachers were generally ill-prepared for the rural environment in Tanzania. There were large gaps between teacher expectations and experiences regarding rural

teaching and many teachers cited ways in which pre-service training should be adapted to assist teachers to cope with the contextual challenges of the rural environment.

Finally, the EQUIP-T TOC seems to require a substantial increase in the number of female teachers in order to see the desired effect. However, the available international evidence, combined with that from the baseline IE qualitative study, does not provide strong evidence that the increase will be substantial.

The third assumption that female Form 4 leavers are able and willing to work in rural locations is contested. The wider evidence base suggests that in order to increase the number of female teachers in rural schools, teacher housing issues (possibly hostels at the cluster school level) and consideration of the mobility and transport issues women teachers face in remote areas (UNESCO PROAP, 2000) need to be addressed. In the IE qualitative research, the environment of rural schools was described by EQUIP-T and government stakeholders as inhospitable for female teachers, stating that some communities view the appointment of females to head teacher roles as unacceptable, that the rural environment is unappealing to female teachers and one stakeholder recalled a community 'booting out' a female invigilator during exams because 'they have never worked under a women teacher'. These descriptions indicate that female teachers entering the working environment of rural schools will require specific coping and support strategies (as suggested in UNESCO PROAP, 2000). It seems likely that the extent to which this third assumption holds will differ depending on the specific strategies in place including facilities for female teachers and the attitudes of the community and school stakeholders within specific contexts. So this assumption is classified as mixed/contextual.

Given that the first assumption is necessary for the second and third, on balance the assumptions underpinning this causal link were judged to be weak. Another factor which contributes to this assessment, is time-lag. Even if the three assumptions above hold, the time required to increase the number of female teachers in rural schools is likely to be beyond the EQUIP-T programme life cycle.

Link 2c: An increase in the number of female teachers in rural schools will lead to a more conducive learning environment for girls.

As stated earlier, evidence suggests that a conducive learning environment for girls is characterised by access to adequate sanitation facilities, safety and security in and on the way to school and being able to participate effectively in class (FAWE 1998 and UNICEF 2010). For an increase in the number of female teachers to lead to a more conducive learning environment for girls, it is assumed that:

- Female teachers enable girls to participate more effectively in class;
- An increase in female teachers will challenge traditional gender stereotypes in schools; and
- Female teachers will influence decision-making in the schooling environment to advocate for gender sensitive approaches.

The IE baseline survey found that 30 percent of teachers in the treatment districts interacted more with boys than girls and 16 percent of teachers interacted more with girls than with boys. Further analysis found that male teachers were slightly more likely to show bias towards boys but that these differences were not statistically significant. On the basis of this contextual evidence, the assumption that female teachers enable girls to participate more effectively in class seems weak.

The assessment of the assumption that an increase in female teachers will challenge traditional gender stereotypes in schools is not straightforward. There is evidence to support a correlation

between girls' enrolment in secondary school and the number of female teachers in primary school, particularly in sub-Saharan Africa (UNESCO 2012b). However, the extent to which this correlation is a reflection of the society's general level of women's empowerment (affecting both variables) or the extent to which girls' increased enrolment can be attributed to the number of female teachers is unknown. For example, a nation with a high level of women's empowerment will most likely have higher participation rates for women across a number of measures. This does not mean that one has caused the other, but rather that a change in the social relations between men and women affected several measures.

The DFID Rigorous Literature Review on interventions to enhance girls' education and gender equality (Unterhalter et al, 2014) found a number of studies providing evidence that having female teachers in local schools increased girls' enrolment probabilities and reduces dropout. However other studies were mixed regarding the effects of male and female teachers on learning outcomes. One study found that when students had more exposure to male teachers, girls did better than boys, while another found favourable outcomes for girls associated with employing women teachers. The review also notes the complexity in attempting to disentangle any possible correlation between women teachers and girls' learning outcomes.

In relation to the third assumption, there is limited evidence that attempts to raise the number of females teachers has corrected the gender bias in subject selection and performance in traditionally male-dominated subject areas, or that an increase in the number of females in an institution results in positive outcomes for all females (a shift in gender relations). The gender bias in mathematics, science and technology has not been corrected in various education subsectors across a range of geographical locations, despite significant attempts for over a decade (Zuga 1999; Minich 1990). Teacher expectations have been found to be self-fulfilling in all subjects, but appear to be the most important in mathematics. There is also evidence to suggest that the effect does not depend on a pupil being the same gender as the teacher (Woolston 2008). Overall, it is not clear that male teachers are primarily responsible for gendered expectations of female pupils.

Research has found that male and female teachers reinforce socially determined ideas defining what roles and activities are deemed appropriate for women and men, reaffirming gendered ideas regarding the productive and reproductive roles of women (Gowlett, 2011; Outhred et al 2012). This was observed in several IE qualitative research case study locations, where an increase in the number of women teachers was perceived to be positive on the grounds that women are the traditional caretakers.

The evidence to support the assumption that female teachers will influence decision-making in the schooling environment to advocate for gender sensitive approaches is weak. The key limitation of the 'just add women and stir' approach is that it attempts to socialise women and girls into the existing hierarchy, rather than interrupting unequal power relations within the school system.

More fundamentally, it is not immediately clear that there is a need for substantially greater numbers of female teachers in programme treatment schools on average. The IE survey found that 55% of teachers in treatment schools were female, indicating that if there is a lack of females in rural schools, the constraints that women currently within the system face with regards to teaching in rural schools should also be addressed. The qualitative study found that the deployment of female teachers, the domestic duties of female teachers and the rural environment were all constraining factors.

However, the IE survey found that only 16% of head teachers were women. The fact that male and female teachers are similarly qualified and have similar years of experience, implies that there may be some gender-bias in promotion for women.

On balance, given the findings from the IE baseline and the wider literature, the assumptions underpinning this causal link overall seem weak. Even if these assumptions hold, the likely time-lags to see the effects of an increase in the number of female teachers in rural schools on the learning environment for girls is substantial and unlikely to be detectable within the programme lifecycle.

Mechanism 3: Teachers will perform better because they will receive training, a teacher morale toolkit will be introduced and incentives linked to a PMS, they will be trained in literacy and language, and head teachers will lead schools more effectively.

This mechanism is supported by three causal links and the assumptions for each of these are discussed below.

Link 3a: Teacher training will lead to better teacher performance

Two assumptions underpinning this link are that teacher training can be effective in increasing the performance of teachers, and that teachers will be more willing and able to take on new skills, roles or ways of working.

The evidence suggests that teacher training can be effective in increasing the performance of teachers. There are several key components of teacher training delivery that have been shown to improve teaching quality, which reflect the planned EQUIP-T model for in-service training. Networking has been identified as effective in establishing connections for mutual support and learning. This is affirmed in a systematic review, where peer support and networking were “described in eight studies as being significant factors in teachers’ implementation of effective practices across Ghana, Bangladesh, India, Benin, Cambodia, Egypt and Lao PDR...”(Westbrook et al, 2013: 62). Research suggests that shorter and more frequent workshops are more effective (Orr et al, 2011: 50), group lectures are less effective as they fail to prompt critical reflection or allow for questions to be raised (Orr et al, 2011: 50), school clusters are an effective approach to maximise resources (Orr et al, 2011: 54) and teacher manuals/guides support “the translation of newly acquired theoretical knowledge into concrete practice” (Westbrook et al, 2013: 60).

Sustainability is a key challenge for training programmes (both teacher and leadership), as educators often revert back to old habits (Orr et al, 2011: 40). A related challenge is where programmes effectively achieve changes to beliefs or attitudes, but these fail to translate to everyday practices (Jull et al, 2014: 78). The translation of education and training into better teaching and learning practices can also be affected by trainer experience and relevance of course content. Westbrook et al (2013) cites limited experience and thus narrow repertoire of trainers, and misalignment between curricular intentions and realities of classroom as hindrances to preparing trainees for the job (Westbrook et al, 2013: 61). Orr et al (2011) highlights the need for high quality educators/mentors, appropriate frequency of visits, support from school management, and the need for trainers themselves to have on-going training (Orr et al, 2011: 51-52).

Therefore, the wider evidence seems fairly strong that, under key conditions, the assumption that high quality teacher training can be effective in increasing the performance of teachers is likely to hold. In the case of the EQUIP-T in-service training programme (both the training for the in-service leader, and for the teachers), it will be essential that there is sufficient time in the school day for these sessions, and that they are linked to the performance management frameworks for teachers. Having confident and well-equipped trainers is also essential, and this means high-quality recruitment and high-quality train-the-trainer preparatory sessions.

On the second assumption: the willingness of teachers take on new skills, roles or ways of working, an EQUIP-T manager spoke frankly regarding a perceived risk to the teacher training, stating that:

“The institution has to cover some of the costs and the teachers have to take their own time, their own effort to learn something... That is a risk and what can we do to motivate the teachers... to sit together twice in a month, maybe they spare two hours but we are not going to give any allowance for these teachers. No food, no transport even some of the people they are saying if I go to EQUIP-T, I don't get anything.”

This highlights the reliance on mutually reinforcing mechanisms within the EQUIP-T TOC, in that if teachers are not motivated to attend training through other strategies within the programme, the quality of training will be of little consequence. Therefore, the extent to which teachers will be more willing and able to take on new skills, roles or ways of working as a result of EQUIP-T activities is likely to be highly contextual, and dependant on the success of the teacher motivation and morale interventions.

The need for teacher professional development is evident from the IE survey results. These show, that teachers have some weaknesses in subject knowledge in mathematics (especially on higher-level curriculum materials) and in Kiswahili (on particular topics including grammar and punctuation). The qualitative research also found that knowledge of the curriculum is weak. In terms of pedagogy, only a small group of teachers displayed a range of effective skills including basic participatory methods.

So, given the clear need for teacher training in EQUIP-T schools, and assuming that the interventions on teacher motivation and morale are effective, on balance the assumptions underpinning this causal link were judged to be reasonably strong.

Link 3b: The provision of a Teachers' Morale Toolkit and incentives linked to a PMS will lead to better teacher performance

This causal link is underpinned by two assumptions: that teachers value career development and advancement, and that incentive programmes can be effective.

The Prova de Promoção programme in Brazil was found to be effective in improving teacher performance through the use of promotion as an incentive. Teachers could opt-in to a selective, highly remunerated career track by passing a difficult exam. There was a cap of 44, 500 new entrants per year and teachers had to wait three years before taking another, more demanding exam to access the next level of pay (Bruns et al, 2012: 59-61).

Based on the qualitative data collected from national, regional, district and ward stakeholders, the previous system of responsibility allowance has now been abolished in Tanzania. While some stakeholders discussed the need for a responsibility allowance to incentivise staff to improve their teaching in the interests of career development, other stakeholders discussed the status, power and flexibility that is associated with career advancement.

Therefore, there is some evidence that teachers value career development and advancement in principle but only given certain management contexts which, in this case, EQUIP-T may have little scope to influence, so on balance this assumption is assessed as mixed/contextual.

There is some international evidence supporting the second assumption that incentive programmes can be effective in improving teacher performance. Examples include the IMPACT programme in the United States (Dee and Wyckoff, 2013), Bonus Pay in Pernambuco, Brazil

(Bruns et al, 2012) the Seva Mandir camera monitoring programme in India (Duflo et al, 2012) and the World Bank/Azim Premji Foundation performance pay initiative, also in India (Muralidharan and Sundararaman, 2011).

The IMPACT ‘stick and carrot’ strategy of pay and threat of dismissal created dual mechanisms of teacher effort and teacher attrition. Teachers who were already high performing were recognised and retained in the sector; other teachers capable of higher performance improved the performance due to both negative and positive incentives. Teachers who could not or were unwilling to improve their performance left the sector or were dismissed. The evidence indicates that successful bonus pay programmes rely on system and policy features and strong governmental institutional support for reform (Bruns et al, 2012).

Again, there is evidence that incentive programmes can improve teacher performance, but design and contextual assumptions are critical to success, and so this assumption is also judged as mixed/contextual.

One of the most important results from the IE survey and qualitative research is that low teacher morale and motivation is pervasive. Low morale and motivation has various manifestations, but perhaps the most damaging for pupil learning is that the majority of teachers are not attending their timetabled lessons (even though they are on the school premises). The qualitative study finds that teachers face a number of inhibiting factors, of which the most commonly cited are teacher housing shortages, high transport costs, low and late pay and a perceived lack of respect for the profession.

These IE survey results illustrate the need to vastly improve teacher motivation and morale. It is the perception of most stakeholder groups in the IE qualitative research that the vast majority of teachers are willing and will be able to take on new skills, roles or ways of working should the teaching environment become more conducive through skills training and through various incentives. The international literature suggests that incentive programmes (particularly payment incentives supported by other systemic changes) can improve motivation and morale, but only under certain design and contextual conditions. In the EQUIP-T context, it not clear that the actual design and scale of the EQUIP-T intervention (Morale Toolkit) will meet the necessary conditions. For example, while payment incentives if part of the design may improve motivation at some level, they are unlikely to address all the systemic inhibiting factors highlighted by the IE qualitative research, and hence teacher performance may remain limited. This is a real risk for the overall success of the programme, since many other causal links are linked to this one.

Link 3c: Head teachers leading schools more effectively will lead to better teacher performance

This causal link is underpinned by the assumption that school leadership behaviour affects teacher performance.

Adeyemi (2010) in a study of heads and teachers in Nigeria showed that the leadership style of the head teacher impacts on the quality of teaching. Waters et al. (2003), using a meta-analysis of some 70 research projects estimated low to moderate effects of school leadership in the school on pupil achievement. They also found a number of relationships between facets of leadership and pupil achievement ranging from a low effect size related to the visibility of the leadership in the school to a moderate effect size related to the intellectual stimulation provided and situational awareness. Robinson (2007, p5) found that most of the effect of leadership is mediated by teachers. Waters et al. (2003) also drew attention to a wide variation in the impact of school

leadership (ranging from negative effects through to strong effects). Therefore, bad leadership can negatively affect teacher performance.

The IE qualitative study found a wide variation of school leadership behaviour displayed by head teachers and found two extreme cases of head teacher leadership with the nine case studies. One of these cases found an association between the way in which the head teacher conceptualised and enacted the key elements of school leadership and the way in which teachers engaged in reflecting on their practice.

Combined with the wider literature, this indicates that there is substantial evidence to support the assumption that head teachers leading schools more effectively will lead to better teacher performance, so the assumption supporting this causal link was judged as strong.

8.3 EQUIP-T component 2: Improving school leadership and management

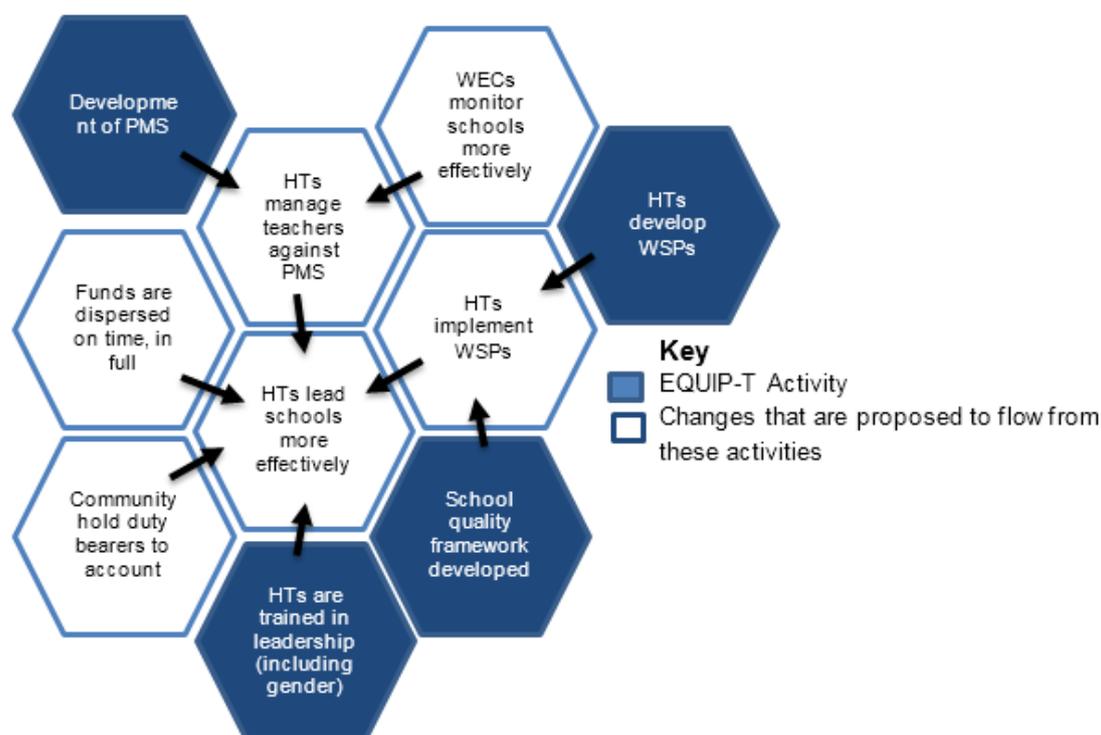
This section examines the causal links and assumptions which support the second component of EQUIP-T, which aims ensure that head teachers lead schools more effectively. As discussed above, this change itself is linked to improving teacher performance in the TOC.

Causal mechanisms

The review of EQUIP-T programme documents and interviews with key stakeholders including EQUIP-T programme staff suggests the mechanisms behind the EQUIP-T TOC for this component are:

- Head teachers will lead schools more effectively because they develop and implement WSDPs, manage teachers against a PMS, funds will be disbursed on time and in full, community members will hold duty bearers to account and head teachers will be trained in school leadership;
- Head teachers will manage teachers against a PMS because they are monitored more effectively by WECs; and
- Head teachers will implement WSDPs because there is a school quality framework and because they are monitored by WECs.

Figure 48 Causal mechanism: Improving school leadership and management



Assessing the assumptions underpinning the causal links

The causal links that contribute to the three mechanisms expected to achieve the desired changes shown in Figure 48 are underpinned by a range of implicit and explicit assumptions as indicated by the arrows. A summary of the causal links and strength of the assumptions underpinning these are shown in Table 11 for EQUIP-T component 2 and are discussed in the following sections.

Table 11 Summary of causal links and strength of underpinning assumptions

Improving school leadership and management		Strength of assumptions underpinning causal links		
Causal links	Strong	Weak	Mixed/Contextual	
1a: WECs monitoring school more effectively and the development of PMS will lead to head teachers managing teachers against a PMS, and therefore, head teachers will lead schools more effectively.	X			
1b: The development of a school quality framework will lead to head teachers developing and implementing WSDPs, and therefore, head teachers will lead schools more effectively.	X			
1c: Funds being disbursed on time and in full will lead to head teachers leading schools more effectively	X			
1d: Head teachers being trained in leadership (including gender sensitisation) will lead to head teachers leading schools more effectively			X	
1e: Communities holding duty bearers to account will lead to head teachers leading schools more effectively.			X	
<p>Strong: The wider literature base and contextual data from the IE baseline survey provide substantial evidence that the main assumptions underpinning the link are likely to hold.</p> <p>Weak: The wider literature and contextual data from the IE baseline survey provide little to no evidence that the main assumptions underpinning the link are likely to hold.</p> <p>Mixed/contextual: There is some evidence from the wider literature and the contextual data from the IE baseline survey that the main assumptions underpinning the link are likely to hold. However, the assumptions may only hold under certain conditions and may therefore lead to heterogeneous results across treatment districts.</p>				

Mechanism 1: Head teachers will lead schools more effectively because they will develop and implement WSDPs, manage teachers against a PMS, funds will be disbursed on time and in full, community members will hold duty bearers to account and head teachers will be trained in school leadership.

Link 1a: WECs monitoring school more effectively and the development of PMS will lead to head teachers managing teachers against a PMS, and therefore, head teachers will lead schools more effectively.

This link is underpinned by the assumptions that (1) WEC monitoring will focus on head teacher management and provide required feedback to head teachers to improve their management and that (2) clear guidelines of roles and responsibilities linked to an appraisal system will improve the performance of teachers. This is, in turn, premised on the assumption that these guidelines are fair and relate well to the role of head teacher. As with teachers, so also for Head teachers, what they are expected to know and to do can be contested, and so the establishment of clear guidelines can be a complex task.

While focusing on school inspections, rather than WEC visits, a national audit in Tanzania found that a focus on pupil performance is not sufficiently prioritised in school inspections and that school inspection materials are not currently focused on school management for better pupil performance (Twaweza, no date cited). In addition, the audit found that feedback mechanisms to schools often do not function effectively. The IE qualitative study found that while WECs conceptualise their roles as including providing support to head teachers, the majority of teachers and head teachers felt WECs focussed on disciplinary actions and engagement with the community, rather than focussing on head teacher management. Therefore, the introduction of monitoring materials to support a focus on school management and the provision of feedback to head teachers is likely to orientate WEC monitoring towards a focus on these key areas. Overall, the assumption that WEC monitoring will focus on head teacher management and provide the required feedback to improve head teacher management, is strong.

There is very little literature available from developing contexts regarding the association between teacher and head teacher guidelines, appraisals and performance.

Militello et al. (2013) examined how head teachers put prescribed leadership standards into practice, in a small sample of 61 head teachers in US schools. They found three different types of response; none of these types of response, however, led to a clear and unambiguous implementation of standards. The nexus between standards and performance was varied and not always clear. There was at times a gap described by the authors as ‘dissonance’ between the standards and the practice of the head teachers. Nevertheless they concluded that meaningful outcomes were achievable. In another US study, Davis et al. (2013) compared the impact of college administrator credential programmes, on-the-job experiences, and the Educational Leadership Policy Standards on 101 randomly selected head teachers. They found that on-the-job experience had the strongest impact, and the standards the weakest. However the authors commented that all the respondents viewed the standards as important. From this it might be concluded that the standards had had some impact on them, even if it was not manifest in their practice.

A now old, but widely cited paper by McEnrue (1984), which focused more generally on the relationship between clarity of role and job performance suggested that the effectiveness of clarity on job performance is mediated by the perceived competence of the person. This implies that clarity in the guidelines is not of itself sufficient for improving the performance of head teachers. The qualitative study found that many head teachers failed to conceive of their roles outside of the day-to-day management of teachers. This indicates that within the treatments districts, there is a

disconnect between the roles and responsibilities of a head teacher required to increase school performance and the perceived roles and responsibilities of the incumbent.

Collectively, this disconnect and the wider literature suggest that the assumptions underpinning this link will hold, should the guidelines be fair and relate well to the role and capacity of head teachers.

Link 1b: The development of a school quality framework will lead to head teachers implementing WSDPs, and therefore, head teachers will lead schools more effectively.

This link is based on the assumption that schools do not currently have or implement WSDPs and that financial constraints alone do not determine whether a school performs well or not.

The EQUIP-T IE survey found that only 20 percent of head teachers could provide a WSDP, only 2 percent of schools had WSDPs that had a budget, teaching and learning objectives, and baseline data and targets. This indicates that this is a very relevant area of intervention.

Schools' finances can be constrained in recurrent or capital expenditures or both. It is likely that constraints on either will impact on the quality of education and how well a school performs. Since teacher salaries are recurrent expenditure, and the quality of the teaching in a school is an important determinant of pupil learning, financial constraints here may have more impact, especially at a systemic level.

Lee and Zuze (2011) recently examined the link between pupil achievement and resourcing of schools in Botswana, Malawi, Namibia, and Uganda. They found a 'strong' association between funding and resourcing levels and pupil achievement at Grade 6 in Mathematics and Reading. A similarly strong effect was identified in OECD nations (Haegeland, Raaum, & Salvanes, 2012). So, clearly, funding is important across a range of contexts. However, a study in the USA (Heyneman, 2013) showed, using PISA data that high levels of funding do not necessarily translate into high levels of pupil achievement. Many US schools were seen on these measures to be inefficient. Also in the USA, Guryan (2001) showed that increased spending improved 4th grade but not 8th grade test scores (most improvement in the performance, however, was by low-scoring pupils).

There is some evidence that cleverly targeted expenditure can have positive effects on pupil learning. For example, Holden (2013) using a regression discontinuity design that took advantage of a sudden change in policy and resourcing, found that a one-off payment of just under \$100 per pupil for textbooks in California had significant positive effects on pupil achievement (an increase of 0.14 standard deviations in reading and mathematics scores).

Within schools and communities, frustrations with the capitation grant were predominantly related to low and late release and also frustrations with the approval processes and decision-making around how funds are utilised.

Therefore, decision-making processes and management skills within schools, and the health and wellbeing of pupils and communities, are examples of other constraints that determine whether a school performs well. This indicates that while resourcing is important, financial constraints are not the only determinant of school performance.

Link 1c: Funds being dispersed on time and in full will lead to head teachers leading schools more effectively

This link is based on the assumption that school planning will improve when the school management and the community are confident about when and how much the school will receive

in capitation grants. There is quite limited evidence about positive experience with funds, but plenty of evidence to suggest that late and partial grants are problematic.

Twaweza suggest that capitation grant payments in Tanzania are so unpredictable in their timing that proper planning is impossible. The EQUIP-T IE survey found that only 17 percent of schools received their expected capitation grants in 2013. Osei et al. (2009) examined the use of capitation grants in Ghana finding that one of the reasons for a lack of impact on education outcomes was delays in funding.

Similar reports have been made about delays in provision of capitation grants in Uganda. One report commented:

“UPE [Universal Primary Education] schools are frequently forced to incur debts with local suppliers while waiting for the disbursement of UPE capitation grants. This not only results in a reduction in quality and quantity of services but also limits schools ability to organise and track their spending. As capitation grants are spent before disbursement, balancing of schools’ budgets becomes increasingly difficult—if not impossible.”

The late and reduced-level release of capitation funding also affected the ability of schools to raise funding through IGA in several schools as head teachers planned to engage the school in IGAs but missed investment opportunities (or seasonal opportunities for agricultural activities) due to the late release of funds.

Overall, therefore, the available evidence supports the idea that late, unpredictable and partial funding is problematic, which suggests that the reverse – funding that is on time and in full – should help head teachers manage more effectively.

Link 1d: Head teachers being trained in leadership (including gender sensitisation) will lead to head teachers leading schools more effectively

This link is underpinned by a number of assumptions. The first is that the challenges facing head teachers can be overcome through training in leadership skills.

The range of challenges a head teacher may face is wide. This assumption embraces, therefore, many facets of the work of a head teacher. Chapman (2005 p1) noted that many programmes designed to train and head teachers are deficient and therefore, there is a necessary assumption that training provided is fit-for-purpose and of high quality. An evaluation of the Managerial Skills Training Programme for Primary School Heads in Zimbabwe by Mandebvu and Chitekuteku (2000) found that despite the very uneven roll out of the programme, there was some evidence that it had had a positive impact. The programme seemed to have had most impact on skills related to budgeting and similar work. Jull et al. (2014) reported on the Leadership for Learning (LfL) programme of school leadership in Ghana. They report results from questionnaire data gathered from 125 head teachers who participated in the LfL programme in Ghana between 2009 and 2011. Their evidence suggests some successes, although, as the research was part of a formative evaluation, it did not strongly focus on outcomes. Ng, examining a programme in Hong Kong suggested that the impact of training may vary according to the domain, for example, with new head teachers feeling most comfortable with issues related to teaching, and less so with strategic directions and policy development. Zachariou et al. (2013) examined school leaders’ education and training for sustainable development in Cyprus. They concluded that head teachers were poorly prepared and the professional development programmes were in need of revision (which implies they believe that these challenges can be overcome by improved training and skills).

The IE survey and the qualitative study found the majority of head teachers in the case study schools had never received training and the qualitative study found that many head teachers lacked knowledge and confidence in enacting their leadership roles.

The evidence base and the low levels of knowledge and confidence of head teachers indicates that the challenges facing head teachers can be potentially be overcome through improved training and leadership skills, depending on the quality of the training given the readiness of head teachers.

The second assumption underpinning this link is that training on the skills required of head teachers will lead to the improved management of schools. This assumption is based on the sub-assumption that those skills, their relative importance and difficulty levels are known. It also assumes that the scope and depth of what is required to manage a school is known.

There is some evidence and a strong general belief that training leads to improved outcomes for head teachers. There is, however, very little literature about the connection between training and the management of schools in general, rather the literature focused upon facets of management. For example, Eller (2010) looked at professional development around technical, socialization and self-awareness (role clarification) skills of head teachers. He saw these as being of most importance, but does not systematically link them back to a more generic notion of 'school management'. There is a wide spread assumption that training and other professional development is important for head teachers. For example, Bandur (2012) examined policy changes in Indonesia which led to increased school-based management. He advocated regular professional development to consolidate the gains made by this policy change.

One important assumption underpinning this causal link is that head teachers are willing and able to consider and respond to gender implications. As with teachers, this relies on the approach to gender sensitisation training, the social setting within which gendered interactions take place and the extent to which a critical mass of 'neighbours' also change their views.

Overall, therefore, the evidence regarding the extent to which this assumption holds is mixed/contextual and relies on how the programme is implemented.

Link 1e: Communities holding duty bearers to account will lead to head teachers leading schools more effectively.

This link in the causal mechanism is underpinned by the assumptions that communities currently rarely hold duty bearers to account and that if they did, head teachers would respond by leading more effectively.

The meaning of the term 'community' is contested, and its implications for other policy imperatives are also contested. For example, the idea of a community implies that it consists of a group with shared characteristics, which in turn implies there are others outside of this community. It is not stretching credibility to see how this could give rise to issues of exclusion, misrepresentation and unfairness. Furman (2012) provides a useful overview of the history of the notion of community and the debates that it has provoked. Tembo (2013) makes a particularly important contribution on the precariousness of the notion of community as part of his review of the Mwananchi Programme in six African countries. What is clear from this work and the debates Furman reports is that the notion of community is less benign than it first seems.

The qualitative study supports the finding that communities do not necessarily share key characteristics and rather, include members with divergent views, levels of engagement and levels of participation in the school environment. Within the nine schools communities were representative of diverse groups of parents with different levels of involvement in the school. There

was strong evidence in many case study schools that some community members do currently hold duty bearers to account, but these members were not always representative of the community as a whole.

The second assumption is that head teachers will respond to being held to account by communities by managing schools more effectively. Tembo (2013) describes an apparently growing consensus that has developed since the publication of 2004 World Development Report (WDR) 'Making Services Work for Poor People' that direct citizen engagement is essential "for the formulation and implementation of better public policies, deepening democracy, and achieving better development outcomes".

Blimpo and Evans (2011), in a methodologically sound evaluation in the Gambia, investigated the extent to which a school-based management and capacity building programme called Whole School Development was effective. One of their findings was that in villages with high literacy, the programme impacted more on pupils' learning outcomes compared with villages with low literacy is low. This might suggest that the validity of EQUIP-T assumptions are contingent upon the context in which they are operating, with literacy levels a key factor.

Therefore, the evidence to support the assumptions underpinning this link are mixed and will likely elicit heterogeneous responses based on context.

8.4 EQUIP-T component 3: Strengthening district and regional education management

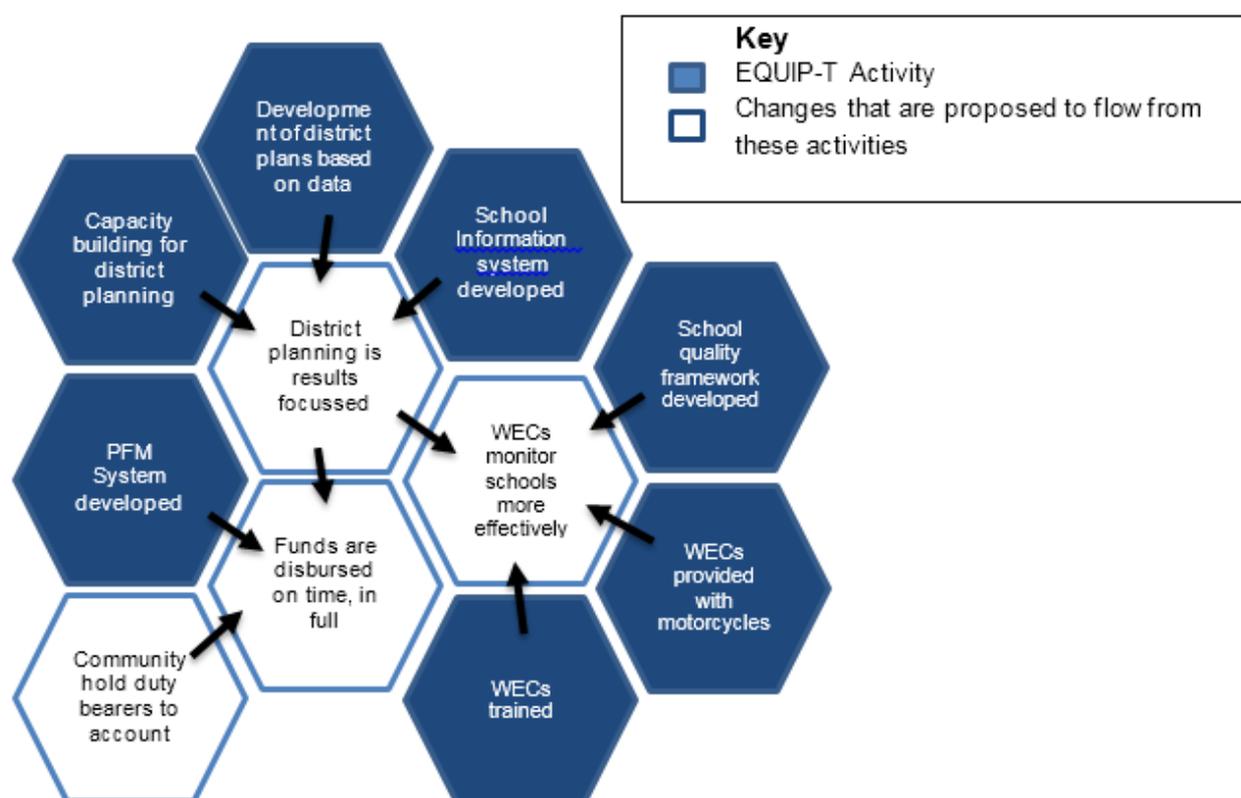
This section examines the causal links and assumptions which support the third component of EQUIP-T, which aims to strengthen district and regional management of education. In terms of the wider TOC, there is a key causal link between strengthening district and regional management and head teachers leading schools more effectively.

Causal mechanisms

The review of programme documents and interviews with key stakeholders including EQUIP-T programme staff suggests the three causal mechanisms for this component are:

- District planning will be results-based if there is a School Information System, district plans are developed (based on data), training in district planning is provided and a Public Financial Management system is in place;
- Funds will be disbursed on time and in full if district planning is results-based, a PFM system is developed and communities hold duty bearers to account; and
- WECs will monitor schools more effectively if a school quality framework is developed, district planning is results-based, and they are provided with motorcycles and training.

Figure 49 Causal mechanisms: Strengthening district and regional management



Assessing the assumptions underpinning the causal links

The causal links that contribute to the three mechanisms expected to achieve the desired changes shown in Figure 49 are underpinned by a range of implicit and explicit assumptions as indicated by the arrows. A summary of the causal links and strength of the assumptions underpinning these are shown in Table 12 for EQUIP-T component 3 and are discussed in the following sections. In addition to these links, *all* of the causal links for this component are underpinned by the assumption that political elites and vested interests will not disrupt or block elements of the EQUIP-T programme.

The Tanzanian based REPOA group (Manara & Mwombela 2012) observed the following about the Primary Education Development Programme in Tanzania (PEDP): “Although PEDP has increased school autonomy, the role of local governments at district, ward and village levels in the delivery of primary education remained unclear. Interference by the District Primary Education Office (DPEO), Ward Education Coordinators (WECs) and village government officials have been creating tensions among the school committee members and teachers. The school committee is the lynchpin of the success of the PEDP at the community level (Tanzania Education Network, 2003), but members do not have full mandate on school management.”

An education programme that interrupts a political equilibrium is likely to elicit a disruptive response from political elites and interest groups. At least two of the EQUIP-T programme core areas have the potential to disturb equilibria;

- Strengthening district planning and management: The objective here is to strengthen systems and human resource capacity to support sub-national management of education; and

- Strengthening community participation and education accountability: The objective here is to empower communities to take an active role in improving school outcomes and services by equipping them with tools, resources and accessible and appropriate information.

The qualitative study found that REOs, DEOs and WECs acknowledge political or vested interests disrupting the use of resources and weakening the education planning cycle. Empowering head teachers, teachers and communities will necessarily shift power away from the centre.

At this point, therefore, available evidence suggests that political elites and vested interests may disrupt or block elements of the programme and this risk is present across all the causal links discussed below.

Table 12 Summary of causal links and strength of underpinning assumptions

Strengthening district and regional management	Strength of assumptions underpinning causal links		
	Strong	Weak	Mixed/Contextual
Causal links			
1a: School information system development will lead to results-based district planning	X		
1b: The development of district plans based on data will lead to results-based district planning	X		
1c: Capacity building for district planning will lead to results-based district planning			X
2a: Results-based district planning will lead to funds being disbursed on time and in full		X	
2b: The development of a PFM system will lead to funds being disbursed on time and in full		X	
2c: Communities holding duty bearers to account will lead to fund being disbursed on time and in full		X	
3a: The development of school quality frameworks will lead to WECs monitoring schools more effectively	X		
3b: WECs being provided with motorcycles, and a grant for recurrent costs, will lead to WECs monitoring schools more effectively	X		
3c: Results-based district planning will lead to WECs monitoring schools more effectively			X
3d: The training of WECs will lead to WECs monitoring schools more effectively	X		
<p>Strong: The wider literature base and contextual data from the IE baseline survey provide substantial evidence that the main assumptions underpinning the link are likely to hold.</p> <p>Weak: The wider literature and contextual data from the IE baseline survey provide little to no evidence that the main assumptions underpinning the link are likely to hold.</p> <p>Mixed/contextual: There is some evidence from the wider literature and the contextual data from the IE baseline survey that the main assumptions underpinning the link are likely to hold. However, the assumptions may only hold under certain conditions and may therefore lead to heterogeneous results across treatment districts.</p>			

Mechanism 1: District planning will be results-based if there is a School Information System, district plans have been developed, training in district planning is proved and a PFMS is in place.

Link 1a: School information system development will lead to results focussed district planning

This link relies on the assumption that the timely submission of school-based data to districts will make it possible to use EMIS data for planning purposes. This rests on the assumption that these data are of good quality and relevant to planning purposes. It also assumes that a country has the capacity to collect and use these data. This was recently investigated by UNESCO (2010), which

examined this in Lesotho, Madagascar, Mozambique, South Africa, Swaziland, Tanzania and Zambia. The report found widely varying quality and capacity across these countries. Some of the issues identified included:

- A lack of clarity around who was responsible for the data collection and data quality standards;
- A lack of clarity about acceptable levels of data quality;
- A lack of technical knowledge;
- Limited collaboration between EMIS and other education sector stakeholders resulting in non-harmonized and regularized data formats and collection approaches;
- Low levels of coordination between regions and the centre; and
- A lack of capacity in schools to collect and process data.

The scope and level of detail required for data to be accurate, comparable and useful for planning purposes is extensive, and infrastructure is also needed (Hare 2007). Ombui (2009) found that the use of ICT by school heads in Kenya indicated that much support is required for technology to be used effectively and efficiently. Attfield and Vu (2013) report the successful use of a national data system to track school standard compliance in improving pupil achievement in Vietnam. Wamakote et al. (2010), looking at EMIS in East Africa noted that in Uganda and Kenya district systems were not networked to the national centre.

The qualitative study found that REOs, DEOs and WECs are familiar with the process for using EMIS data for planning purposes, but do not currently trust the available data and find it difficult to verify data due to limited resources. Therefore, while there is little research to either support or refute the assumption, knowing the number of pupils enrolled, staff teachers and schools and the levels of resourcing and pupil performance at each site is fundamental to the education planning process. Therefore, while implementation of a new EMIS may be challenging due to infrastructure limitations and data harmonisation difficulties, the contextual evidence suggests that the timely submission of school-based data to districts will make it possible to use EMIS data for planning purposes, though whether or not data are used for planning will depend on other political economy and capacity factors.

Link 1b: The development of district plans based on data will lead to results-based district planning

While there is a wide consensus on the fundamentals, dimensions and stages of education planning, over time there have been changes in the tasks, approaches and actors that shape its development (UNESCO IIEP no date cited). Education planning is not merely a technical exercise, but rather an organised social process involving a variety of actors. The main criticisms of traditional planning have included too much focus on plan preparation and not enough on implementation, plans being prepared in a top-down way manner and not enough consideration to the changing environment. Therefore this link relies on the assumptions that the district education plan will be participatory, iterative, flexible, change orientated and will emphasise plan implementation.

Link 1c: Capacity building for district planning will lead to results-based district planning

This link assumes that district education staff who acquire new skills will be willing and able to change behaviours and put their learning into practice. This assumption is highly contextual and relies on the individual attitudes, behaviours and knowledge of district officials. Assessing these individual traits of officials at baseline is not possible. Therefore, in the absence of data, logic would indicate that the extent to which this assumption will hold will be mixed or contextual due to

the differing social and political settings in each district and individual preferences of the education officers involved.

Mechanism 2: Funds will be disbursed on time, in full because district planning is results focused, there is a PFM system developed and communities hold duty bearers to account

Link 2a: Results-based district planning and the development of a PFM system will lead to funds being disbursed on time and in full; and Link 2b: The development of a PFM system will lead to funds being disbursed on time and in full

These links are underpinned by the assumption that a PFM system will lead to better financial management and expenditures being focused to a greater extent on agreed and budgeted education priorities.

There is very little evidence from the existing literature regarding the extent to which a PFM system leads to better financial management and a greater focus on agreed and budgeted education priorities. A CfBT Education Trust (2013) evaluation of the Zimbabwe School Grants Pilot found that the financial planning component of the pilot appeared to have worked, but the study focused on school and community based expenditure, rather than district level expenditures.

The IE qualitative study found a lack of systems in place to ensure that agreed education priorities as per the budget were expended as per the budget allocation. Reasons for the diversion of education funds were not perceived to be based on capacity, but rather a lack of accountability.

Therefore, the extent to which this assumption is likely to hold is weak.

Link 2c: Communities holding duty bearers to account will lead to funds being disbursed on time and in full

This link is underpinned by the assumptions that community bearers rarely hold duty bearers to account and that officials will respond to being held to account by communities and parents. As stated above, many community members do hold duty bearers to account, however they do not always represent the whole community. The qualitative study finds that community members and parents hold school level duty bearers to account by raising concerns with the village committee, the WEC or the DEO. It is unclear what mechanisms parents and community members could use to hold DEOs to account for fund disbursement. There would likely need to be a critical mass of pressure from many communities to support the assumption that officials will respond to being held to account by communities.

Therefore, the evidence to support the assumptions underpinning this link are weak.

Mechanism 3: WECs will monitor schools more effectively because there is a school quality framework, WECs are provided with motorcycles, district planning is results-based and WECs are trained.

Link 3a: The development of school quality frameworks will lead to WECs monitoring schools more effectively

School quality frameworks hold schools accountable against specific indicators, based on the evidence regarding what domains effect educational quality. These domains can include pupil performance, teaching and learning, family and community, leadership and collaboration and pupil access. This link is based on the assumption that WECs do not currently assess school quality against indicators that predict pupil performance. As mentioned above, a national audit of school

inspection was undertaken. The audit found that a focus on pupil performance is not sufficiently prioritised in school inspections and school inspection materials are also not focussed on pupil performance.

The IE qualitative study found that while WECs support schools across some of these domains (community engagement and teacher attendance), other domains were not systematically assessed by WECs within their roles (for example, teaching and learning and leadership and collaboration). The evidence suggests that school inspection materials are not currently orientated towards focussing on aspects (Twaweza, no date cited).

In addition, the qualitative study found divergent perceptions from WECs regarding what makes a good school and what makes a good teacher. Therefore, systematically standardising indicators across these domains is likely to increase the effectiveness of WEC monitoring. This indicates that the assumption underpinning this link is likely to hold, provided WECs are appropriately incentivised and capacitated to support schools.

Link 3b: WECs being provided with motorcycles, and a grant for recurrent costs, will lead to WECs monitoring schools more effectively

This link in the causal chain is based on the assumptions that transportation is a major problem for WECs and that WECs effectively monitor school when they visit them.

The IE qualitative study found that transport is a major constraining factor for WECs. Some WECs within the study have already been issued with motorcycles and still cited transport as a major constraint due to the lack of resources for petrol and maintenance of motorcycles. Given that the programme will also provide a grant to cover these recurrent costs, at least during the lifetime of the programme, this intervention is likely to contribute to WECs monitoring schools more frequently. Currently WECs are visiting primary schools on average 6.6 times per year, compared to 12 times set in the guidelines. Head teachers typically seem to appreciate their support, although it is clear there is scope for improving their effectiveness. On balance the assumptions underpinning this link are likely to hold.

Link 3c: Results-based district planning will lead to WECs monitoring schools more effectively.

Results-based planning relies on a focus on achieving goals, effectiveness and impact, flexibility and is often high risk (Chang 2006). Results-based planning increases the accountability of stakeholders as performance is judged by outcomes and impact, rather than inputs. This link is underpinned by the assumption that results-based planning will include the required incentives and sanctions for WECs to be held accountable for outcomes and impacts. This will predominantly be determined based on the extent to which DEOs and REOs buy in to and enforce results-based planning.

Thus, this link is likely to be mixed and contextual depending on the region and district.

Link 3d: The training of WECs will lead to WECs monitoring schools more effectively.

As with teachers and head teachers, this link is based on the assumption that training can improve performance. Similarly, regular professional development has been found to improve performance, should training be of high quality and fit-for-purpose. Therefore, the evidence to support this link in the causal chain is strong, but depends on the quality and relevance of training.

8.5 EQUIP-T component 4: Strengthening community participation and accountability

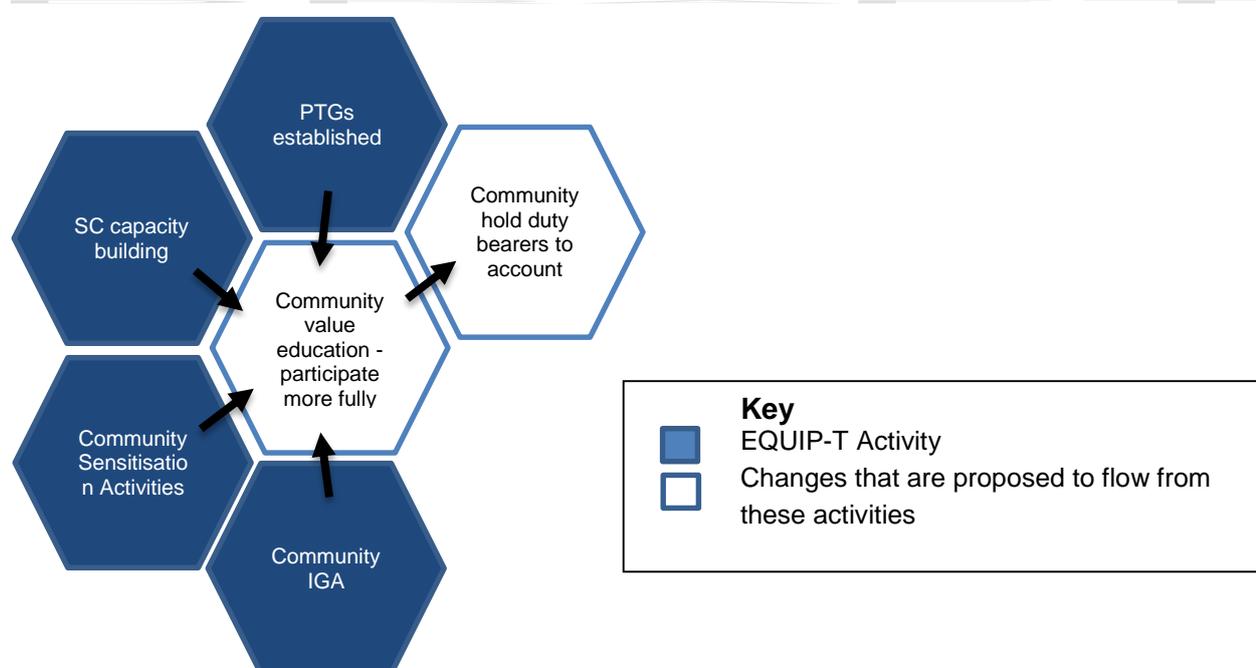
This section examines the causal links and assumptions which support the fourth component of EQUIP-T, which aims to strengthen community participation and accountability. In terms of the wider TOC, there is a key causal link between strengthening community participation and accountability and head teachers leading schools more effectively.

Causal mechanisms

The review of programme documents and interviews with key stakeholders suggests the causal links for this component are:

- Community members valuing education and participating more fully will lead to them holding duty bearers to account;
- The establishment of PTGs will lead to community members valuing education and participating more fully;
- Building the capacity of the SC will lead to community members valuing and participating in education more fully;
- Community sensitisation will lead to community members valuing and participating in education more fully; and
- Access to capital to engage in IGAs will lead to communities valuing education and participating more fully.

Figure 50 Causal mechanism: Strengthening community participation and accountability



Assessing the assumptions underpinning the causal links

The causal links that contribute to the three mechanisms expected to achieve the desired changes shown in Figure 50 are underpinned by a range of implicit and explicit assumptions as indicated by the arrows. A summary of the causal links and strength of the assumptions underpinning these are shown in Table 13 for EQUIP-T component 1 and are discussed in following sections.

Table 13 Summary of causal links and strength of underpinning assumptions

Strengthening community participation and accountability	Strength of assumptions underpinning causal links		
	Strong	Weak	Mixed/ Contextual
Causal links			
1a: Community members valuing education and participating more fully will lead to them holding duty bearers to account			X
1b: The establishment of PTGs will lead to community members valuing education and participating more fully			X
1c: Building the capacity of the SC will lead to community members valuing and participating in education more fully			X
1d: Community sensitisation will lead to community members valuing and participating in education more fully			X
1e: Access to capital to engage in IGAs will lead to communities valuing education and participating more fully	X		
<p>Strong: The wider literature base and contextual data from the IE baseline survey provide substantial evidence that the main assumptions underpinning the link are likely to hold.</p> <p>Weak: The wider literature and contextual data from the IE baseline survey provide little to no evidence that the main assumptions underpinning the link are likely to hold.</p> <p>Mixed/contextual: There is some evidence from the wider literature and the contextual data from the IE baseline survey that the main assumptions underpinning the link are likely to hold. However, the assumptions may only hold under certain conditions and may therefore lead to heterogeneous results across treatment districts.</p>			

Mechanism 1: Communities hold duty bearers to account because they value education and participate more fully, community members value education and will participate more fully if PTGs are established, SC capacity is built through training, the community has received sensitisation and the community is given access to capital to engage in IGAs.

Link 1a: Community members valuing education and participating more fully will lead to them holding duty bearers to account

As with the School Leadership component regarding holding duty bearers to account, this link is underpinned by the assumption that communities rarely hold school management and teachers to account for the delivery of education. As discussed previously, the qualitative study found that in many schools, members of the community often hold school management and teachers to account. However, this was not always coordinated or organised collective action. In other schools, this power relationship was inverted, with parents fearing schools reporting them to the village committee, resulting in unaffordable fines.

Therefore the evidence regarding the assumption is mixed and therefore, the strength of this link in the causal chain is likely to be context specific.

Link 1b: The establishment of PTGs will lead to community members, valuing education and participating more fully

This link is built on the assumption that an active community will ensure that pressure is applied through the PTG on the SC and school to report to the community on their actions.

Banerjee et al. (2010) examined, among other outcomes, the impact of a participatory programme on village education committee awareness. These committees were established by the Indian government linking community members with district educational authorities. This 3ie evaluation study was methodologically strong. It found that the committees did not affect parental awareness of school, parental involvement with school, parental knowledge of education, nor their understanding of the problems the schools faced in the village. Nor were the committees associated with any increase in the provision of extra nonteaching resources. The authors note:

“This study and most previous studies show that community or citizen participation in publicly provided services in the health and education sectors is not a magical remedy to improve health and education final outcomes because their effectiveness depends heavily on local context. To improve final outcomes, creative solutions need to be found that involve direct action of citizens in local service delivery.”

In another study in India, Pandey et al. (2011) investigated the Sarva Shiksha Abhiyan (SSA), a nation-wide public scheme initiated in 2001. The SSA, among other objectives, aimed to increase accountability of schools to communities by cultivating greater involvement of village education committees and parent-teacher groups. The study was methodologically strong, and found that:

- There had been a significant improvements in the functioning of school committees;
- School committees were more active. (The number of school committee meetings and attendance at them, school visits increased significantly); and
- School committees had improved knowledge of school accounts and of committee’s roles and responsibilities.

Therefore, within these two examples alone, the effectiveness of PTGs to apply pressure on the SC and the school to report on their actions is mixed.

The qualitative study reflected contextual conditions regarding the extent to which parents and communities can apply pressure on the school to report on their actions. There were several examples of community members holding schools to account, however this was acknowledged as a risky exercise, as parents feared teachers would fail to teach their children as a result of complaints. In some case study schools teachers and head teachers felt parents yielded power in terms of identifying individuals required to account for their behaviour to the village council or the WEC. However, neither parents nor school-level stakeholders viewed communities as being able to effect change within the school. Rather, in some case study schools, communities were able to apply pressure for resolutions to specific incidences.

Therefore the establishment of PTGs is likely to lead to this outcome in some contexts and not in others.

Link 1c: Building the capacity of the SC will lead to community members valuing and participating in education more fully

This link in the causal chain is underpinned by the assumptions that (i) training of the SC will improve the capacity to plan and manage financial resources; and (ii) SCs do not have necessary skills to fulfil their intended functions (PFM, planning, management and knowledge of their statutory role).

The assumption that training school committees will improve their capacity to plan and manage financial resources is framed by the understanding that this training would be of good quality (that is, with appropriate content and delivered well).

Pradhan et al. (2011), in a 3ie evaluation investigated the educational impacts of four Indonesian interventions aimed at increasing community involvement by strengthening school committees. The interventions were (i) the provision of training to school committees, (ii) organising elections of committee members, (iii) creating linkages between the school committees and local governments with the aim of establishing and helping support the committees' authority and (iv) a grant of money to support the work of the committee members. The study was methodologically rigorous.

It found, among other findings that the intervention increased understanding of the role of the committee and the rights of members.

The evidence stated earlier from Blimpo and Evans (2011) that a programme in the Gambia impacted more on pupils' learning outcomes in villages with high literacy compared with villages with low literacy, is relevant here too.

The IE qualitative study found that SC members face an access to information problem when fulfilling their intended functions (PFM, planning, management and knowledge of their statutory role). The qualitative research also found that in addition to their intended functions SCs serve as a problem solver across the school with almost unlimited functions and responsibilities. However, rarely do SC members see representing parents view as one of their core functions.

This indicates that capacity building of the SC will lead to different outcomes in different contexts.

Link 1d: Community sensitisation will lead to community members valuing and participating in education more fully

Two assumptions underpin this link in the causal chain. The first is that community sensitisation will increase buy-in by parents and the second is that this will, in turn, ensure that the community is more supportive of schools.

While sensitisation is a necessary part of many development initiatives, sensitisation has long effect lags and the effectiveness is notoriously contextual. While some WECs described the communities within their wards as slowly changing and adapting to value education, others indicated very strong resistance from communities as a whole, some parents within the community, or powerful members of the community. In addition, a number of the communities have already received sensitisation with varying degrees of success.

Therefore the strength of this link in the causal chain is contextual and where community sensitisation does lead to community members valuing and participating in education more fully, the effects may not be seen within the programme life cycle.

Link 1e: Access to capital to engage in IGAs will lead to communities valuing education and participating more fully

This link is underpinned by the assumptions that (i) financial resources for education will be generated by upgrading existing school based IGAs or establishing new IGAs, (ii) income generated by IGAs will be used on education projects and (iii) communities will continue to engage in IGAs after the initial capital injections because they value education.

The qualitative study finds that many of the IGAs within schools are constrained by the late and reduced-level capitation grants, as the investments required are disbursed after debts have been incurred or after the required season for agricultural activities has passed. This indicates that the provision of capital to engage in IGAs will mitigate some of these constraints at the school level. Community led IGAs have been instigated in many of the case study communities. Education officials reported positive results, however the sustainability of the initiatives were questioned as many IGAs were not repeated after initial investments. Within these case study schools, multiple stakeholders reported that the income generated from community led IGAs was used for education purposes. This indicates that the assumptions that financial resources for education will be generated by upgrading existing school based IGAs and that the income generated will be used for education purposes are likely to hold. However, the issue of sustainability beyond initial capital investments remains. The qualitative study also found that transparency of communication and

realistic expectations of what can be financed through IGA is crucial for ongoing community support.

In addition, the evidence suggests other complexities may be present. The Indonesian study cited earlier (Pradhan et al. 2011) found that conflicts arose between the Head teacher and the committees over control of the grant monies, and that the fair election of members was at time jeopardised by the attempts of existing members to preserve the status quo. This indicates that shifts in the power dynamics between stakeholders will likely cause some stakeholders to resist change and disrupt programme outcomes.

Therefore, the qualitative findings indicate that the assumptions are likely to hold. However, the issue of sustainability and unpredictable resistance strategies in the face of shifting power dynamics, bring a level of volatility to this link in the causal chain.

8.6 The complete expanded theory of change

The expansion of the EQUIP-T programme TOC and assessment of the assumptions underpinning each of the causal links provides a picture of the likely effects of the various programme activities along each link within the causal pathway (see Figure 51 for a complete picture). One weak link within a mechanism or pathway may not be detrimental to the achievement of programme outputs and outcomes if there are sufficient other links reinforcing the mechanism or causal pathway. In summary, the TOC analysis finds that:

- The majority of the causal mechanisms expected to deliver changes in the programme outputs and outcomes are based on substantial evidence, should the EQUIP-T inputs be relevant (for example, teacher training content) and of quality. Improving teacher performance, improving the effectiveness of head teacher leadership, and improving the effectiveness of WEC monitoring are all based on strong evidence.
- The strength of the causal link between providing both a Teacher Morale Toolkit and incentives linked to a PMS and better teacher performance will rely heavily on the detailed design of the intervention. It will probably need to tackle systemic inhibiting factors such as poor teacher housing. Without this, teacher performance may remain limited and absenteeism remain critically high, with implications for many other key causal links in the TOC.
- A number of causal mechanisms rely on links that are only likely to be realised under certain contextual conditions. The qualitative study finds that these conditions are present in some locations, but not all. EQUIP-T activities within communities will only likely lead to communities valuing education and participating more fully in some locations, and there will also likely be heterogeneity of results within communities themselves. The study finds predominately strong evidence (with some links relying on certain conditions to be present) for EQUIP-T activities to lead to results-based district planning, but weaker evidence that this would lead to funds being disbursed on time and in full.
- Two causal links were found to be particularly weak. First, the evidence did not support the assumptions underpinning the causal link for scholarships for Form 4 female school leavers to lead to a more conducive learning environment for girls. The number of girls eligible to enrol in teacher training, the capacity of teacher training colleges to prepare teachers for the challenges of rural classrooms, and the inability of rural females to compete with urban males during recruitment processes were all contributing factors. Second, the evidence to support the assumptions underpinning the causal link that providing support for pre-primary classes will lead to children being school-ready, with language and literacy skills was limited. While the evidence suggests pre-primary classes may result in increased school readiness in relation to

-  **Strong:** The wider literature base and contextual data from the baseline survey provide substantial evidence that the main assumptions underpinning the link are likely to hold
-  **Weak:** The wider literature based and contextual data from the baseline survey provide little to no evidence that the main assumptions underpinning the link are likely to hold.
-  **Mixed/contextual:** There is some evidence from the wider literature base and the contextual data from the baseline survey that the main assumptions underpinning the link are likely to hold. However, the assumptions may only hold under certain conditions and may therefore lead to heterogeneous results

It is worth bearing in mind that the budget for the different EQUIP-T interventions shown in Figure 51 varies widely. Obviously it is particularly important to consider the findings of the TOC analysis which relate to the areas of greatest planned investment. A summary of the EQUIP-T budget by component is given in Volume II Annex N (Section N2.1), which shows that teacher development (component 1) easily accounts for the largest share of the budget. In fact within the teacher development budget, the vast majority of planned spending is on INSET with a far smaller proposed investment in the teacher morale and school readiness interventions.⁹⁴

⁹⁴ A more detailed budget analysis of EQUIP-T will be available in the Fiscal Affordability Concept Paper which will be produced in 2015.

PART D: Conclusions and next steps going forward

9 Conclusion

9.1 Pupil learning and quality of education

The IE baseline survey finds that the majority of standard three pupils in treatment schools have fallen considerably behind curriculum expectations in Kiswahili and in mathematics. There is a particularly large group of pupils that have fallen more than two standards behind in Kiswahili. The learning gaps are noticeably acute for pupils from non-Kiswahili-speaking homes in both subjects. The fact that Kiswahili is the medium of instruction appears to be a considerable barrier to learning for this majority of pupils (77%), who do not benefit from continuous exposure to Kiswahili at home.⁹⁵ Boys and girls perform similarly in Kiswahili, but boys have significantly stronger skills in mathematics than girls on foundational (standard one) level material. Teachers', head teachers' and parents' perceptions of differential learning outcomes between boys and girls are often inaccurate and based on perceptions of pupil behaviour and innate capacity rather than on actual performance.

This picture of learning levels at baseline confirms that the core objective of the programme – 'to provide better-quality education, especially for girls' – is highly relevant, and that large-scale change will be needed. It may be important to ensure that pupils who do not speak Kiswahili at home are considered as a target group in the core objective, as well as girls.

Looking at the four core mechanisms in the expanded TOC that are expected to improve school quality as a result of programme interventions, the IE baseline findings suggest that each of the mechanisms is clearly related to constraints identified at baseline, and that the links between the expected changes and improving school quality are generally consistent with available evidence from Tanzania and elsewhere.⁹⁶ The IE baseline results give some insight into the extent of change needed in each intervention area:

- *Children 'school-ready' with literacy and language skills:* Major change needed – for the majority of pupils, current low levels of attainment of foundational skills appear to be at least partly related to the mismatch between their home language and the medium of instruction. Very few teachers are currently using language switching in class, which puts pupils who do not speak Kiswahili at home at a real disadvantage and has negative implications for their learning.
- *More conducive learning environment for girls:* Change needed – there is a gender gap in foundational mathematical skills, there exists a group of teachers who appear to disadvantage girls during classroom interactions, and, most alarmingly, the qualitative research suggests that girls sometimes face serious harassment within the school or on the way to school, and are also sometimes required to take on inappropriate domestic work within the school.⁹⁷

⁹⁵ Of course, there may be other differences between the linguistic groups that are also correlated with academic attainment.

⁹⁶ This means that the wider evidence base suggests that changes of this type can improve school quality under various specified design and contextual conditions.

⁹⁷ There is also a group of teachers who appear to disadvantage boys in classroom interactions, although this group is smaller. The qualitative research also found that boys are sometimes also required to carry out 'chores' during school hours.

- *Children attend school more regularly:* Major change needed – on the day of the IE baseline survey, one-third of pupils (both boys and girls) in standards one to three were absent.
- *Teachers perform better:* Change needed across multiple areas of professional capacity, performance, motivation and morale (the IE baseline study finds both strengths and weaknesses in the current situation, discussed in the next section) – one area where major change is clearly required is teacher classroom absenteeism, as the majority of teachers are not attending their scheduled lessons, despite being on school premises.

While most of the links between the planned programme interventions and the change expected in the first three areas above (the last, teacher performance, is discussed below) appear to be supported by existing evidence, the IE baseline study finds two important potential weaknesses. First, on the school-readiness intervention, while there is strong evidence that pre-school classes can increase school-readiness on a wide range of important measures, these do not include language acquisition, which is clearly critical in this context. Second, on improving the learning environment for girls, the assumptions required for the Form 4 scholarship programme to lead to more female teachers into rural schools are numerous, and at best only likely to hold in certain contexts. It also seems likely that any change stemming from this intervention is likely to occur outside the programme life cycle.

9.2 Teacher performance

Improving teacher performance is the core aim of the first component of the EQUIP-T programme. This is a highly relevant objective given the potentially strong link between teacher performance and improving school quality (discussed above), and because baseline IE findings show low levels of teacher performance in critical areas. Teachers in programme schools vary considerably in their current levels of professional capacity and performance, but teacher motivation and morale are low for the majority. These findings suggest that the areas of programme intervention in this component (see bullets below) are highly relevant to improving teacher performance. The IE baseline findings related to these intervention areas give some insight into the extent of change needed:

- *Ongoing professional development, including gender sensitisation:* Variable levels of change needed across the different elements of teacher capacity and performance:
 - Teacher subject knowledge in mathematics is fairly good on lower curriculum materials, but varies enormously for upper curriculum material. There is less variation in Kiswahili subject knowledge; teachers scored reasonably on the lower curriculum material but struggled with particular topics including grammar and punctuation.
 - Knowledge of the current curriculum, and new syllabus, is very weak according to the qualitative research.
 - Both strengths and weaknesses were observed in teacher pedagogy, but overall only a small group of teachers displayed skills conducive to higher-order thinking among pupils, including basic participatory methods.
 - A small majority of teachers currently display gender-balanced interactions with pupils during lessons.
- *Morale toolkit and incentives:* Major change needed – generally low levels of teacher motivation and morale appears to be leading to behaviour that is detrimental to pupil learning. Teachers are frequently absent from school (more than one in 10 were absent on the day of the survey), and the majority of those who attend school arrive late. Even when teachers are present in school, the majority are not attending their scheduled lessons. This combined effect reduces available instructional time markedly. The qualitative study found various explanations

for low teacher motivation and morale, the most commonly cited being a lack of appropriate teacher housing close to schools, and low salaries necessitating teachers engaging in other IGAs, late salary payments, and a perceived lack of respect for the teaching profession.

While the link between professional development training and improving teacher performance is supported by the literature,⁹⁸ support for the morale and motivation intervention depends very heavily on intervention design and context. It seems questionable at this point whether the Teacher Morale Toolkit and implementation of a PMS will be able to address many of the systemic inhibiting factors identified by the qualitative research.

Given the clear need for behaviour change among the majority of teachers, this intervention is critical to programme impact on pupil learning and seems at risk. Another factor that could undermine the effectiveness of this component, especially in relation to the professional development intervention, is the high level of teacher turnover found at baseline (only half of teachers of standards one to three have been in their current school for one year or more).

9.3 School leadership and management

The core aim of the second programme component is to ensure that head teachers lead schools effectively, so that schools meet quality standards, including those set for teacher performance. The TOC analysis finds evidence that the link between head teachers leading schools effectively and improving teacher performance (the most direct link in the TOC to improving the quality of education for this component) is strong in principle.⁹⁹ Given the generally weak SLM capacity observed at baseline, this appears to be a highly relevant objective to improving educational quality.

Overall, the IE baseline study finds that head teachers' capacity to lead and manage schools is generally weak across a range of leadership competency areas. Indeed, the qualitative study found that head teachers were not typically aware of many basic school leadership responsibilities. Given this baseline context, the areas of programme intervention seem highly relevant. The IE baseline findings shed light on the extent of change that is likely to be needed in each area.

- *Leadership PMS supported by WECs:* Change needed – there is limited awareness or enactment of core school leadership competencies in the case study schools, and current leadership is undermined by high levels of head teacher absence (close to one in five head teachers were absent on the day of the survey). WECs are, however, visiting schools on a reasonably frequent basis (roughly once every two months on average) but the qualitative research finds that they often do not carry out key roles and responsibilities.
- *Professional development training:* Change needed – very few head teachers have received any recent professional development training on SLM, and the qualitative study found that most head teachers had never received any leadership or management training.
- *Whole school planning and financial management systems:* Major change needed – very few schools currently have WSDPs that contain critical basic elements such as budget, teaching and learning objectives, and baseline data and targets. The lack of the budget element is indicative of weak financial management systems.

⁹⁸ Assuming various design and contextual conditions.

⁹⁹ In the expanded TOC, of all the changes expected to flow from head teachers leading schools effectively, the pathway to improving teacher performance is the most direct link to improving education quality. Of course there are numerous other indirect pathways.

- *School information systems and management processes (underpinned by school quality framework):* Major change needed – for example, teacher performance management, a key element of a school quality framework, is currently focused on checking lesson plans and to a lesser extent lesson observation, but critically teachers appear to receive very little concrete feedback to improve their skills. The vast majority of teachers said that they had not had a one-on-one meeting to discuss their performance or professional development needs in the previous school year. There is also limited evidence of head teachers rewarding good performance, either verbally or materially. It is telling that only a small percentage of head teachers cited punctuality and attendance as the most important measure of teacher performance in the context of such high levels of teacher classroom absence. The qualitative research found that some head teachers did not feel empowered to manage teachers in this respect, while others were simply unwilling to enforce the rules. Many teachers also viewed WECs as their first line of accountability for attendance, rather than head teachers.

The links between the planned programme interventions¹⁰⁰ under this component and improving school leadership were found to be supported in principle by existing evidence, with the exception of leadership professional development (including gender sensitisation), where the extent to which this is likely to hold is mixed/contextual, partly depending on the scope and quality of training, as well as the readiness of head teachers to receive training and then apply new skills. Other factors that could potentially undermine the effectiveness of head teacher training are the high level of turnover (nearly four in 10 head teachers have been in their current school for one year or less) and that 15% of head teachers are due to retire in the next four years.

9.4 District and regional education management

The core objective of the third programme component is to ensure that regions and districts have the capacity to effectively manage the education system. Three key changes are sought: for WECs to monitor schools more effectively; for funds to be disbursed to schools on time and in full; and for district education planning to be results-based. The links between the first two changes and improving head teacher leadership are the main pathways to improving the quality of education for this component (results-based planning will support both of these changes).¹⁰¹ The IE TOC analysis finds these are strong links in principle. This suggests that, given the weaknesses in district and regional management observed at baseline, these changes are relevant programme objectives.

Overall, the IE baseline study finds capacity gaps across a range of areas related to achieving the changes outlined above. These gaps include:

- *Physical resources:* Education officials overwhelmingly cited these constraints as the biggest challenge to fulfilling their responsibilities; WECs suggested that monitoring of schools is currently compromised by a lack of funds for transport.
- *Lack of transparency in key financial and budget processes,* including the capitation grant for schools – Less than one school in five reported receiving capitation grants that met or exceeded expectations.

¹⁰⁰ Development of PMS, HTs develop WSDPs, School quality framework developed, HT trained in leadership (including gender)

¹⁰¹ In the expanded TOC, the pathways between these three changes (themselves interlinked) and improving head teacher leadership is the most direct link to improving education quality. Of course there are numerous other indirect pathways.

- *Mixed performance in current support systems for school improvement:* WECs are fairly active in visiting schools, managing an average of one visit per school every two months, but the qualitative study found that WECs are mainly enacting disciplinary and trouble-shooting roles, rather than supporting school improvement more broadly. Systems for ensuring delivery of in-kind resources to schools seem to be functioning at some level, given that the vast majority of schools received some in-kind resources in the previous two school years (mainly textbooks).
- *Weaknesses at all stages of the education planning cycle:* The lack of reliable data from schools for planning purposes was emphasised as a key constraint by district education officials, while implementation of education plans is perceived to be mainly undermined by incomplete and late disbursement of funds.

While the majority of links between the planned interventions¹⁰² and the desired changes were found to be supported by existing evidence (given contextual factors), three interventions appear to have weak links to the changes sought. First, that results-based district planning or that communities holding duty bearers to account will contribute to funds being disbursed on time and in full. Second, that providing WECs with motorcycles will lead to more frequent monitoring of school, partly because reliable recurrent support for fuel and maintenance is also necessary. Third, and more generally, the IE study identifies the risk of political elites and vested interests disrupting or blocking certain elements of the programme.

9.5 Community participation and accountability

The objective of the fourth programme component is to inform, equip and empower communities to engage more effectively in school operations and outcomes. This component targets two key changes: communities valuing education and participating more fully in schools; and communities holding duty bearers (including the head teacher) to account for school performance. In the TOC, these are interlinked, with the first supporting the second. The link between holding duty bearers to account and improving head teacher leadership is the main pathway to improving the quality of education for this component.¹⁰³ The IE TOC analysis finds mixed/contextual evidence to support this pathway in principle, and the IE baseline study finds limited evidence that communities are well informed, equipped and empowered currently. Taken together, this affirms the relevance of this component's objectives.

The more detailed IE findings on community–school relationships at baseline give a sense of the scale of the changes likely to be needed from this component's interventions:

- *Informed communities:* Change needed – communication mechanisms between schools and communities are often not working well. Only half of schools in the survey have a public noticeboard, most with limited relevant information. Monitoring pupil academic progress is largely perceived as a parental responsibility, but one quarter of teachers surveyed admitted that they do not report on academic progress to parents. Letters sent home with pupils are the main form of communication between schools and parents in the case study schools, but this is channel is sometimes unreliable.
- *Equipped communities:* Change needed – SCs are present in all schools, and generally understand their broad roles and responsibilities, but capacity to implement these is currently

¹⁰² PFM system developed, capacity building for district planning, development of district plans based on data, WECs provided with motorcycles, WECs trained.

¹⁰³ In the expanded TOC, the pathways between these two changes (themselves interlinked) and improving head teacher leadership is the most direct link to improving education quality. Of course there are numerous other indirect pathways.

perceived to be weak, especially with respect to school finances. SCs in the case study schools reported they have limited convening power with parents, and most did not mention representing parents' views as a core responsibility. Parents reported providing some level of material support to the case study schools (either financial or in-kind), but did not always know how their contributions had been used. About a third of schools were unable to show evidence of a previous parents-teachers meeting.

- *Empowered communities*: Change needed – the qualitative study found that while parents generally felt that their ability to hold teachers to account was limited, teachers and head teachers reported fearing parental attempts to hold them to account, based on previous experiences, especially if the village council and other spheres of government were involved.

All of the links between the planned interventions¹⁰⁴ and communities valuing education and participating more fully, were supported by mixed/contextual evidence, except the IGA intervention, where there is evidence of a stronger link. Community-led IGAs have been instigated in many of the case study communities already, and it was a common view that the capital invested, when available, ultimately generated funds which were used for education purposes. The broader research highlighted the risk of a lack of sustainability of IGAs after the initial investment, and also the potential for disruption and conflict due to shifting power dynamics linked to resources.

10 Follow-up quantitative survey and qualitative research

The quantitative survey conducted at baseline will be repeated in 2016 and 2018 to enable the IE to measure impact of the EQUIP-T programme as a whole on key programme impact and outcome indicators over time (see Volume II, sections 3.4 and 3.5). The quantitative survey data collected at midline will also be used to enable the qualitative research post-2016 to explore selected unexpected changes in key outputs and outcomes.

The qualitative and quantitative baseline fieldwork investigated a set of contextual factors which when combined with information on the programme design helped to inform the expanded programme TOC. The expanded TOC sets out the key causal pathways and makes explicit the contextual and main design assumptions underpinning each link in the chain. This provides a very useful and detailed frame for the next round of research. The baseline findings together with the analysis of the expanded TOC at this stage, suggests the following as some of the key areas of focus for the qualitative research during the midline round. There will also be some refocusing of the quantitative instruments in the light of these priorities.

- *Home language of pupils and how it affects their classroom and school experience*, including how teachers use Kiswahili and mother tongue in classrooms, as well as other pedagogical strategies to support non-Kiswahili speaking pupils; and what constraints and strategies pupils who don't speak the medium of instruction have and use to cope in the classroom and in school.
- *Provision and nature of pre-school provision*, including the basic features of pre-school provision in schools, particularly the literacy, language and pedagogy skills and knowledge of pre-school teachers; and, on the demand-side, the extent to which parents are willing to send their children to pre-primary education classes and reasons for their choice.
- *Teacher motivation and morale and related interventions*, including factors driving the extremely high level of classroom absence by teachers; links between the direct interventions

¹⁰⁴ Parent-teacher groups established; SC capacity building; Community sensitisation activities; Community IGAs.

to improve teacher morale and motivation, as well as indirect links via head teacher management practices, WEC supervision, and the reporting of absenteeism at school level and up the chain.

- *Relationship between the presence of female teachers in schools and a more conducive learning environment for girls*, with a particular focus on the risk of sexual harassment and violence, as well as seeking to understand more about the learning experience for girls in mathematics lessons compared with other subjects.
- *Role and performance of WECs and related causal pathways in the TOC*, including more detailed research into the nature and quality of services that WECs provide to schools, their capacity in terms of knowledge, skills, time and other constraints, and expectations of their role from different stakeholders.
- *Community and school accountability*, especially the characteristics of schools and communities, and mechanisms, where accountability relationships work well and pressure brings about positive change or resolution of conflict.

The results from the detailed analysis of the EQUIP-T budget, which will be part of the Fiscal Affordability Study Concept Paper (to follow in 2015), will also help to guide the mid-line research. It is difficult to be specific at this point, but areas of the programme which are costly, and are part of an area of the TOC where there are weak links will be an important area for follow-up research. Conversely, areas of the TOC where the baseline research suggests that large changes are needed, yet there is very little investment planned and few interventions will also merit particular attention.

More generally, at mid-line the qualitative research will start to compare the perceived effectiveness of the main EQUIP-T components, as well as investigating signs of heterogeneous impact across case-study sites. This analysis will be indicative at best, given that the programme will only be two years into implementation at this point, and some components started before others. These aspects of the research will be consolidated in the end-round in 2018.

Bibliography

- Abadzi, H. (2006) *Efficient Learning for the Poor: insights from the frontier of cognitive neuroscience*. World Bank, Directions in Development Series.
- Adeyemi, T.O. (2010) Principals' leadership styles and teachers' job performance in senior secondary schools in Ondo State, Nigeria *Journal of Education Administration and Policy Studies*, 2(6), 83–91.
- Alderson, P. and Morrow, V. (2004) *The Ethics of Social Research with Children and Young People*. Barking: Barnardos.
- Attfield, I. and Vu, B.T. (2013) A rising tide of primary school standards—The role of data systems in improving equitable access for all to quality education in Vietnam. *International Journal of Educational Development*, 33(1), 74–87.
- Bandur, A. (2012) School-Based Management Developments: Challenges and Impacts. *Journal of Educational Administration*, 50(6), 845–873.
- Banerjee, A.V., Banerji, R., Duflo, E., Glennerster, R. and Khemani, S. (2010) Pitfalls of Participatory Programmes: Evidence from a Randomized Evaluation in Education in India. *American Economic Journal: Economic Policy*, 2(1), 1–30.
- Berlinski, Galiani and Manacorda, (2008). Giving Children a Better Start: Preschool Attendance and School-Age Profiles. *Journal of Public Economics*, 2008, vol.92, iss.5-6, pp.1416-40.
- Birdsall, N., Ibrahim, A. and Gupta, G. (2004) Interim Report on Achieving the Millennium Development Goal of Universal Primary Education United Nations Development Programme.
- Blimpo, M. and Evans, D. (2011) *School-Based Management and Educational Outcomes: Lessons from a Randomized Field Experiment*: The World Bank.
- Bruns, B et al (2012). *Achieving World-Class Education in Brazil : The Next Agenda*. World Bank Report. <https://openknowledge.worldbank.org/handle/10986/2383> License: CC BY 3.0 IGO.
- Bush, T and Glover, D. (2003) *School Leadership: Concepts and Evidence*. Summary Report. National College for School Leadership.
- Cambridge Education (2014a) Draft EQUIP-T Programme Description, 6 January.
- Cambridge Education (2014b) Final EQUIP-T Inception Report, 5 February.
- Carvalho, S. and White, H. (1997) 'Combining The Quantitative and Qualitative Approaches to Poverty Measurement and Analysis', Technical Paper 366. Washington DC: World Bank.
- Cartwright, N, and Hardie, J., (2012) *Evidence Based Policy: A Practical Guide To Doing It Better*. Oxford: Oxford University Press.
- Casey, K., Glennerster, R. and Miguel, E. (2013) *The GoBifo project evaluation report: Assessing the impacts of community-driven development in Sierra Leone (Vol. 3ie Impact Evaluation Report 3)*.
- Cataldo, F., Kielmann, K., and Seeley, J., (2011). *Introduction to Qualitative Research Methodology*. Research for Development.
- CfBT Education Trust (2013) *Zimbabwe: School Grants Pilot Baseline Evaluation*: www.unicef.org/evaldatabase/files/Zimbabwe_2013-010_Documentation_of_SIG_pilot_Report.pdf
- Chang, 2006. *National Education Sector Development Plan: A results-based planning handbook*. UNESCO IIEP.
- Chapman, D., Snyder, C. and Burchfield, S. (1991) *Teacher incentives in the Third World (Vol. ERIC Document: ED354621)*. Washington: Agency for International Development.

- Chapman, J. (2005) Recruitment, retention, and development of school principals: UNESCO, International Academy of Education, International Institute for Educational Planning.
- Chaudhury, Nazmul, Jeffrey S. Hammer, Michael, Kremer, Karthik Muralidhara, and F. Halsey Rogers. Missing in Action; Teacher and Health Worker Absence in developing countries.
- Cogill, J (2008). Primary teachers' interactive whiteboard practice across one year: Change in pedagogy and influencing factors. EdD thesis. King's College University of London.
- Cueto, S et al. (2009) Psychometric characteristics of cognitive development and achievements instruments in Round 2 of Young Lives. Young Lives Technical Note #15, January 2009.
- DataVision (2013) Survey on Flow and Release of School Capitation Grants in Tanzania: Survey Report and Analysis. Dar es Salaam: DataVision International.
- Davis, S. H., Leon, R.J. and Fultz, M. (2013) How Principals Learn to Lead: The Comparative Influence of On-the-Job Experiences, Administrator Credential Programmes, and the ISLLC Standards in the Development of Leadership Expertise among Urban Public School Principals. *International Journal of Educational Leadership Preparation*, 8(1).
- Day, Sammons, Hopkins, Harris, Leithwood, Gu, Penlington, Mehta and Kington (2007). The Impact of School Leadership on Pupil Outcomes. Interim Report. University of Nottingham.
- Dee, T and Wyckoff, J Incentives, Selection, and Teacher Performance: Evidence from IMPACT. NBER Working Paper Series, No. 19529 <http://www.nber.org/papers/w19529>
- Dehejia, R.H. and Wahba, S. (2002). Propensity-Score Matching Methods for Nonexperimental causal studies. *The Review of Economics and Statistics*, 84(1): 151–161
- Devarajan, S. (2013) Disrupting Low-level Political Equilibria. from <http://blogs.worldbank.org/futuredevelopment/disrupting-low-level-political-equilibria>
- DFID (2013) Intervention Summary. Education Quality Improvement Programme in Tanzania (EQUIP-T). London, UK: Department for International Development.
- DFID (1998). Investigating bilingual literacy: Evidence from Malawi and Zambia – Education Research Paper No. 24, 1998, 99p Department for International Development Education Papers.
- Duflo, E. et al (2012) Incentives Work: Getting Teachers to Come to School. *American Economic Review* 102, 1241–1278.
- Duflo, E. et al (2014) School Governance, Teacher Incentives, and Pupil-Teacher Ratios: Experimental Evidence from Kenyan Primary Schools. Working Paper.
- Dunne, M. and Leach, F. (2005) Gendered School Experiences: the Impact on Retention and Achievement in Botswana and Ghana. London: DFID.
- Education Partnerships Incorporated (No date) Coaching Teachers: An Important Principal Role: ERIC document ED538322.
- Eller, J.F. (2010) An Evaluation of a Development Programme for New Principals. *The Qualitative Report (ERIC doc: EJ896180)*, 15(4), 956–965.
- Evans, A. (2014). Does gender sensitisation work? Accessed at <https://www.opendemocracy.net/alice-evans/does-gender-sensitisation-work>
- Evans, L. (1992) 'Teacher Morale: an individual perspective'. *Educational Studies*, 18:2: 161–171.
- EQUIP-Tanzania (2014) Draft National School Leadership Competencies Framework
- Forum for African Women Educationalist (FAWE) (1998). Experiences in Creating a Conducive Environment for Girls in School.
- Furman, G. (ed.) (2012) *School as community: From promise to practice*: SUNY Press.
- Galab, S. et al (2013) Community-based accountability for school improvement.

- Garbarino, S., and Holland, J., (2009), 'Quantitative and Qualitative Methods in Impact Evaluation and Measuring Results', Issues paper commissioned by DFID. GSDRC Emerging Issues Service.
- Gertler, P.J., Martinez, S., Premand, P., Rawlings, L. and Vermeersch, C. (2011) *Impact Evaluation in Practice*. Washington, DC: World Bank.
- Gertler, P et al (2012) Parental Empowerment in Mexico: Randomized Experiment of the "Apoyos a La Gestion Escolar (Age)" Program in Rural Primary Schools in Mexico. Society for Research on Educational Effectiveness.
- Getzel, J. W., & Guba, E. G. (1957). Social behaviour and the administrative process. *The School Review*, 65, 423-441.
- Glewwe, P., Maïga, E.W.H. (2011) The impacts of school management reforms in Madagascar: do the impacts vary by teacher type? *Journal of Development Effectiveness* 3, 435–469. doi:10.1080/19439342.2011.604729
- Global Partnership for Education (2013) *All Children Learning Report*.
- Government of Tanzania (2007) *Mtaala wa elima ya msingi*. Dar es Salaam.
- Gowlett (2011) Injurious assumptions: Butler, subjectification and gen(d)erational poverty. In Outhred, R. *Reconceptualising Equity in Higher Education. International Journal of Inclusive Education. Volume 16. No 9*.
- Guest, G., MacQueen, K., Namey, E., (2010). *Applied Thematic Analysis*. Sage Publications.
- Guryan, J. (2001) Does Money Matter? Regression-Discontinuity Estimates from Education Finance Reform in Massachusetts *NBER Working Paper No. 8269*. National Bureau of Economic Research.
- Haegeland, T., Raaum, O. and Salvanes, K.G. (2012) Pennies from Heaven? Using Exogenous Tax Variation to Identify Effects of School Resources on Pupil Achievement. *Economics of Education Review*, 31(5), 601–614.
- Hallinger, P., and Heck, R. (1996). The Principal's Role in School Effectiveness: An Assessment of Methodological Progress, 1980-1995. In K. Leithwood, *The International Handbook of Research on Educational Leadership and Administration*. New York: Kluwer Press.
- Hardman, F, and Dachi, H. 2012. *Evaluation of School-Based INSET Pilot Programme*. York, UK: Institute for Effective Education, University of York.
- Hare, H. (2007) *ICT Education in Tanzania. Survey of ICT and Education in Africa*. from www.infodev.org/infodev-files/resource/InfodevDocuments_432.pdf
- Hattie, J. (2003) *Teachers Make a Difference. What is the research? Distinguishing expert teachers from novice and experienced teachers*. University of Auckland. Australian Council for Educational Research.
- Heinrich, C, Maffioli, A. and Vázquez, G. (2010). *A Primer for Applying Propensity-Score Matching. Impact-Evaluation Guidelines Technical Notes No. IDB-TN-161*. Washington, DC: Inter-American Development Bank.
- Heyneman, S.P. (2013) The International Efficiency of American Education: The Bad and the Not-So-Bad News. *Current Issues in Comparative Education*, 16(1), 74–89.
- Holden, K.L. (2013) *Buy the Book? Evidence on the Effect of Textbook Funding on Student Achievement. Working Papers*. from http://pages.uoregon.edu/kholden/uploads/1/3/8/7/13879025/bythebook_holden.pdf
- Hopkins, J., Harris, A., and Jackson, D. (1997). Understanding the School's Capacity for Development: Growth states and strategies. *School Leadership and Management*, 17(3), 401-412.

- Jull, S., Swaffield, S. and MacBeath, V. (2014) Changing perceptions is one thing...: barriers to transforming leadership and learning in Ghanaian basic schools. *School Leadership & Management*, 34(1).
- Kane, E. (1996). Gender, Culture and Learning. Advancing Basic Education and Literacy, Phase 2. Education Development Centre.
- Khandker, S.R., Koolwal, G.B. and Samad, H.A. (2010) *Handbook on Impact Evaluation. Quantitative Methods in Practices*. Washington, DC: World Bank.
- Kiggundu, E., Moorosi, P. (2012) Networking for school leadership in South Africa: perceptions and realities. *School Leadership & Management* 32, 215–232. doi:10.1080/13632434.2012.688738
- Koerner T. F. (1990). Developing staff morale. *The Practitioner*, 16, 4.
- Korach, S. and Agans, L.J. (2011) From Ground to Distance: The Impact of Advanced Technologies on an Innovative School Leadership Programme. *Journal of Research on Leadership Education*, 6(5), 216–233.
- Kremer, Michael, Edward Miguel, and Rebecca Thornton. (2009). "Incentives to Learn." *The Review of Economics and Statistics* 91(3): 537-56.
- Krippendorff, K., (2004). Content Analysis: An Introduction to Its Methodology. Sage Publications.
- Kunje, D., Selemani-Meke, E. and Ogawa, K. (2009) An Investigation of the Relationship between School and Pupil Characteristics and Achievement at the Basic Education Level in Malawi. *Journal of International Cooperation in Education*, 12(1), 33–49.
- Leach, J (2002). The curriculum knowledge of teachers: a review of the potential of large-scale electronic conference environments for professional development. *Curriculum Journal*, 13(1), pp. 87–120.
- Lee, V.E. and Zuze, T.L. (2011) School Resources and Academic Performance in Sub-Saharan Africa. *Comparative Education Review*, 55(3), 369–397.
- Leithwood. (2010) Reasons for school failure. Chapter 1 in *Leading school turnaround*.
- Lewin, K. (2010) Looking back to see the future: Four decades of development in educational planning. In M. Bray and N. V. Varghese (eds.), *Directions in educational planning; international experiences and perspectives*. Paris: UNESCO IIEP.
- Lipham, J. M., Rankin, R. E., & Hoeh, J. A. (1985). The principalship: Concepts, competencies, and cases. New York: Longman.
- Lumadi, M., L. (2010). The Need For Training Gender-Sensitive Teachers: Addressing Education Challenges for Gender Sensitive National Development. Contemporary Issues in Education Research. March 2010 Volume 3, Number 3.
- Manara, K. and Mwombela, S. (no date) The Governance of the Capitation Grant in Primary Education in Tanzania: Why Civic Engagement and School Autonomy Matter. In REPOA (ed.). Dar es Salaam: REPOA.
- Mandebvu, O.S. and Chitekuteku, S.R. (2000) An Evaluation of the Managerial Skills Training Programme for Primary School Heads in Zimbabwe. Harare Regional Office: UNICEF.
- Mannathoko, C. E. (1994) Democracy in the management of teacher education in Botswana. *British Journal of Sociology of Education*, Volume 15(4), pp.481-495.
- Martinex, S, Naudeau, S and Pereira V. (2012). The Promise of Preschool in Africa: A Randomized Impact Evaluation of Early Childhood Development in Rural Mozambique. 3ie.
- Mathaba, R.S.R. (Undated) *External whole school evaluation of underperforming secondary schools in Mpumalanga Province*. (Doctor of Technology), Durban University of Technology, South Africa. Retrieved from http://ir.dut.ac.za/bitstream/handle/10321/1073/MATHABA_2014.pdf?sequence=1

- Mays, N., and Pope, C. (1995). Rigour and Qualitative Research, *British Medical Journal* 311(6997): 109-112.
- McEnrue, M.P. (1984) Perceived competence as a moderator of the relationship between role clarity and job performance: A test of two hypotheses. *Organizational Behavior and Human Performance*, 34(3), 379–386.
- McKenzie, D. and Woodruff, C. (2012) What are we learning from business training and entrepreneurship evaluations around the developing world? *Policy Research Working Paper 6202*: The World Bank Development Research Group.
- Mendel, P. C. (1987). An investigation of factors that influence teacher morale and satisfaction with work conditions. Unpublished Dissertation, University of Oregon, Eugene.
- Mestry, R. and Schmidt, M. (2010) Portfolio Assessment as a Tool for Promoting Professional Development of School Principals: A South African Perspective. *Education and Urban Society*, 42(3), 352–373.
- Militello, M., Fusarelli, B., Alsbury, T. and Warren, T.P. (2013) How Professional Standards Guide Practice for School Principals. *International Journal of Educational Management*, 27(1), 74–90.
- Ministry of Labour. (2003). United Republic of Tanzania: National Ageing Policy. Ministry of Labour, Youth Development and Sports.
- Minnich, E. K. (1990). Transforming knowledge. Philadelphia, PA: Temple University Press.
- MoEVT (2009a) Basic standards for pre- and primary education in Tanzania. Dar es Salaam: Ministry of Education and Vocational Training. www.ed-dpg.or.tz/pdf/PE/Basic%20Standards%20for%20Pre%20and%20Primary%20Education%20in%20TZ_2009.pdf
- MoEVT (2009b) *Handbook for School Inspectors Third Edition 2009*. Dar es Salaam: Ministry of Education and Vocational Training. www.ed-dpg.or.tz/pdf/PE/Inspector's_Handbook_Final%20March%202010.pdf
- MoEVT (2009c) *Guidelines for School Supervision*. September, 2009. Dar es Salaam: Ministry of Education and Vocational Training. www.ed-dpg.or.tz/pdf/PE/School%20Supervision_guidelines_Final%20March%202010.pdf
- Molander, C. and Winterton, J. (1994) *Managing Human Resources*. London: Routledge.
- Muralidharan, K. and Sundararaman, V. (2011) Teacher Performance Pay: Experimental Evidence from India. *The Journal of Political Economy*, 119(1), 39–77.
- Muzale and Rugemalira (2008). Researching and Documenting the Languages of Tanzania. *Language Documentation and Conservation*. Vol. 2 pp 68-108.
- Myers-Scotton, Carol (1989). Codeswitching with English: types of switching, types of communities. *World English's* 8 (3): 333–346.
- Nag S, Chiat S, Torgenson C, Snowling MJ (2014) Literacy, Foundation Learning and Assessment in Developing Countries: Final Report. Education Rigorous Literature Review. Department for International Development.
- Ng, S.W. (2013) Equipping Aspiring Principals for the Principalship in Hong Kong. *Educational Management Administration & Leadership*, 41(3), 272–288.
- OECD (2008) Improving School Leadership Volume 1: Policy and Practice. www.oecd.org/education/school/41200009.pdf
- Ombui, S.C. (2009) Challenges facing principals in Integrating Information Communication Technology (ICT) for efficient management in secondary schools In Nyamira County, Kenya. from <http://ir-library.ku.ac.ke/handle/123456789/7267>

- OPM (2014a) 'Impact Evaluation of the Education Quality Improvement Programme in Tanzania (EQUIP-T): Inception Report and Baseline Design'. Oxford, UK: Oxford Policy Management.
- OPM (2014b) 'EQUIP-T Impact Evaluation. Preliminary Findings from the Baseline Quantitative Survey for Programme Treatment Districts'. Draft Report. Oxford, UK: Oxford Policy Management.
- Osei, R., Owusu, G., Asem, F. and Afutu-Kotey, R. (2009) Effects of capitation grant on education outcomes in Ghana: Institute of Statistical Social and Economic Research.
- Oulai, D., Lugaz, C., Minas, A. and Teklehaimanot, H. (2011) Analysis of capacity development in educational planning and management in Ethiopia. Paris: International Institute for Educational Planning.
- Palomer, C.G. and Paredes, R.D. (2010) Reducing the Educational Gap: Good Results in Vulnerable Groups. *Journal of Development Studies*, 46(March), 535–555.
- Pandey, P. et al (2009) Community participation in public schools: impact of information campaigns in three Indian states. *Education Economics* 17, 355–375. doi:10.1080/09645290903157484
- Pandey, P., Goyal, S. and Sundararaman, V. (2011) Does Information Improve School Accountability? Results of a Large Randomized Trial (Vol. Report No. 49): International Bank for Reconstruction and Development/ and The World Bank – South Asia: Human Development Unit.
- Pole de Dakar (2012) Tanzania Education Sector Analysis. Beyond Primary Education, the Quest for Balanced and Efficient Policy Choices for Human Development and Economic Growth. Geneva: United Nations Educational Scientific and Cultural Organization.
- Ponera, G.E., Mhonyiwa, J.E. and Mrutu, A.S. (2011) Quality of Primary School Inputs in Tanzania Mainland. Policy Brief Number 2 (October 2011): www.sacmeq.org/sites/default/files/sacmeq/reports/sacmeq-iii/policy-brief/tan_school_inputs_15oct2011_final.pdf [09 September 2014]
- Pradhan, M. et al. (2011) Improving Educational Quality through Enhancing Community Participation: Results from a Randomized Field Experiment in Indonesia *Policy Research working paper; no. WPS 5795*. Washington, DC: The World Bank.
- Prendergast, C. Intrinsic motivation and incentives. *Am Econ Rev* 2008; 98: 201-5 doi: [10.1257/aer.98.2.201](https://doi.org/10.1257/aer.98.2.201).
- Pritchett, L. and Beatty, A. (2012) 'The negative consequences of overambitious curricula in developing countries' *Faculty Research Working Paper Series RWP 12-035*. Harvard Kennedy School.
- Ramanujam, V. and Venkatraman, N. (1987) Planning System Characteristics and Planning Effectiveness. From <http://dspace.mit.edu/bitstream/handle/1721.1/48833/planningsystemsc00rama.pdf?sequence=1&origin=publicationDetail>
- Richards, J.C. and Theodore, S (2001) Approaches and Methods in Language Teaching (2nd Ed.). Cambridge, New York: Cambridge University Press.
- Robinson, V.M. (2007) School Leadership and Student Outcomes: Identifying what works and why *ACEL Monograph series* (Vol. 41).
- Robinson, V. M., Lloyd, C. A., & Rowe, K. J. (2008). The Impact of Leadership on Student Outcomes: An Analysis of the Differential Effects of Leadership Types. *Education Administration Quarterly*, 44(5), 635-674.

- Romano, Babchishin, Pagani and Kohen (2010). School readiness and later achievement: replication and extension using a nationwide Canadian survey. *Development Psychology*. September 46 (5) 995-1007.
- RTI (2014) 'National Baseline Assessment for the 3Rs (Reading, Writing, and Arithmetic). Using EGRA, EGMA, and SSME in Tanzania.' Study Report. Draft. Washington, DC: USAID.
- Sampa, P.C. et al (2009) Improving quality in Zambia's community schools: Sustainable strategies or quick fixes? Working Paper.
- Scheerens, J. (2004) The quality of education at the beginning of the 21st Century: UNESCO.
- Schenk, K. and Williamson, J. (2005) Ethical Approaches to Gathering Information from Children and Adolescents in International Settings: Guidelines and Resources. Washington, DC: Population Council.
- Schreiner, M. (2013) 'A Simple Poverty Scorecard for Tanzania.' Kansas City: Microfinance Risk Management, L.L.C. www.microfinance.com/#Tanzania
- Sumra, S, and Katabaro, J.K. (2014) Declining Quality of Education: Suggestions for Arresting and Reversing the Trend. Special THDR Issues. ESRF Discussion Paper 63. The Economic and Social Research Foundation, Dar es Salaam.
- Spaull, N and Viljoen, J. (2013) Starting Behind and Staying behind: the case of insurmountable learning deficits in South Africa. Draft.
- Stephens, C. (2013) EQUIP-T Impact Evaluation: Key Literature Analysis. Draft, December.
- Swallow, J. et al (2009) USAID/Zambia Education Program Evaluation. USAID.
- Tembo, F. (2013) Rethinking social accountability in Africa: lessons from the Mwananchi Programme: UKAid.
- UNESCO (2010) Assessing Education Data Quality in the Southern African Development Community (SADC): A Synthesis of Seven Country Assessments: UNESCO.
- UNESCO (2011). Financing education in sub-Saharan Africa. Meeting the challenges of expansion equity and quality.
- UNESCO (2012a) EFA Global Monitoring Report 2011. The hidden crisis: Armed conflict and education. Geneva: United Nations Educational Scientific and Cultural Organization.
- UNESCO 2012 (2012b). World Atlas of Gender Equality in Education.
- UNESCO (2014) EFA Global Monitoring Report teaching and learning: Achieving quality for all: UNESCO.
- UNESCO-IIEP (2011). Gender Equality in Education. Looking beyond Parity. An IIEP Evidence-Based Policy Forum. Paris, France. 3-4 October 2011.
- UNESCO IIEP (no date). Educational Planning: Approaches, challenges and international frameworks. Distance Education Programme on Education Sector Planning. Module I. Accessed at [http://www.moe.gov.jo/Files/\(22-12-2013\)\(1-01-43%20PM\).pdf](http://www.moe.gov.jo/Files/(22-12-2013)(1-01-43%20PM).pdf)
- UNESCO PROAP, (2000). Increasing the Number of Women Teachers in Rural School. A Synthesis of Country Case Studies. South Asia. UNESCO Principal Regional Office for Asia and the Pacific and Asia-Pacific Programme of Education for All.
- UNICEF (2002) Children Participating in Research, Monitoring and Evaluation (M&E): Ethics and Your Responsibilities As A Manager. Evaluation Technical Notes No 1. Geneva: UNICEF Evaluation Office.
- Unterhalter E, North A, Arnot M, Lloyd C, Moleletsane L, Murphy-Graham E, Parkes J, Saito M (2014) Interventions to enhance girls' education and gender equality. Education Rigorous Literature Review. Department for International Development.

- URT (2010) 'Public Expenditure Tracking Survey of Primary and Secondary Education in Mainland Tanzania.' Final report, February 2010. Dar es Salaam: Ministry of Education and Vocational Training.
- UWEZO (2011) Are Our Children Learning? Annual Learning Assessment 2011. Dar es Salaam: UWEZO Tanzania.
- UWEZO (2013) Are Our Children Learning? Annual Learning Assessment Report 2012. Dar es Salaam: UWEZO Tanzania.
- Vogel, I., (2012), 'Review of the use of theory of change in international development', Review paper for the Department of International Development. Department for International Development.
- Wamakote, L., Ang'ondi, E. and Onguko, B. (2010) East African Government Investments in Information and Communication Technologies for Education: Matching Policy to Practice. *Itupale Online Journal of African Studies*, 2, 25–38.
- Washington, R., & Watson, H. F. (1976). Positive teacher morale: The principal's responsibility. *NASSP Bulletin* 60, 4-6.
- Waters, T., Marzno, R. and McNulty, B. (2003) *Balanced Leadership: What 30 years of research tells us about the effect of leadership on student achievement*: McRel.
- Westbrook J, Durrani N, Brown R, Orr D, Pryor J, Boddy J, Salvi F (2013) *Pedagogy, curriculum, Teaching Practices and Teacher Education in Developing Countries*. Final Report. Education Rigorous Literature Review. Department for International Development.
- Westhorp, G., Walker, B., Rogers, P., Overbeeke, N., Ball, D., and Brice, G. (2014). *Enhancing community accountability, empowerment and education outcomes in low and middle-income countries: A realist review*. EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
- What Works Clearinghouse (2012) *WWC review of the report: Teacher incentive pay and educational outcomes: Evidence from the New York City Bonus Program*. U.S. Department of Education, Institute of Education Sciences, Retrieved from <http://whatworks.ed.gov>.
- White, H. (2009) *Theory-Based Impact Evaluation: Principles and Practice*, 3ie Working Paper 3. *International Initiative on Impact Evaluation*.
- Woolston, W. 2008. Do "Great Expectations" Matter? The relationship between teacher expectations and student academic success. The Institute for Research on Education Policy & Practice. Stanford University.
- World Bank (2010) *Does linking teacher pay to student performance improve results?* (Report No. 57782).
- World Bank (2012) 'Service Delivery Indicators. Tanzania.' Washington, DC: World Bank.
- Wright, B and Stone, M. (1979) *Best Test Design*. Chicago: MESA Press.
- Zachariou, A., Kadji-Beltran, C. and Manoli, C.C. (2013) School principals' professional development in the framework of sustainable schools in Cyprus: a matter of refocusing. *Professional Development in Education*, 39(5).
- Zuga, K. (1999). Addressing Women's Ways of Knowing to Improve the Technology Education Environment for All Students. *Journal of Technology Education*. Volume 10, Number 2.

Annex A Original and agreed terms of reference

A.1 Original terms of reference

ITT Volume 2

Terms of Reference

Impact evaluation of the DFID programme EQUIP-T

Objective

To conduct an impact evaluation of a major DFID education programme being delivered in selected district councils in Tanzania, that aims to identify a replicable and affordable approach to improving the quality of primary education that can be rolled out nationally in Tanzania.

Recipient

The recipient of the services is DFID Tanzania.

Background

Context

Findings from (i) the 2010 independent Tanzania Country Programme Evaluation, (ii) the 2010 Committee of Public Accounts (PAC) recommendations on improving DFID's bilateral investment in basic education and (iii) the 2012 Independent Commission for Aid Impact (ICAI) report on DFID's Education Programme in Three East African Countries (including Tanzania); have all reinforced the need to shift support to education from general budget support alone to a mix of modalities that are more likely to lever positive reform changes, better results and increased efficiency savings. This led to the inclusion within DFID Tanzania's Operational Plan 2011- 2015 of a major new programme intervention that seeks to 'improve the quality of primary education so that more children have better literacy and numeracy skills, with a focus on equitable access and completion for girls'.

Education Quality Improvement Programme in Tanzania (EQUIP-T) is a £60 million, four year intervention (2013-2016) to support the improvement of the quality of education in primary schools in Tanzania and to increase the number of children, particularly girls, able to transfer to secondary education. Focusing on approximately 20 - 25 of the 165 local councils and supporting national institutions the programme will develop cost-effective and affordable school improvement models for national scale up by the Government. The programme will deliver teacher training, management and leadership training, capacity building of council education officers and officials, improved data and targeted measures to support girls' transition. The intention is to demonstrate that functional basic education systems can be created and sustained within typical rural councils of Tanzania. More details of the EQUIP-T programme are provided in the Business Case document, approved in late 2012. In tandem to EQUIP-T, DFID will continue to provide both general and education sector budget support to Tanzania in the three year period up to 2015.

The expected outcome of EQUIP-T is better quality education, especially for girls, and a replicable and affordable approach ready for national scale up by the Tanzanian government. This will be demonstrated by:

- Improved teaching of early-grade reading and numeracy resulting in more children able to read with comprehension.
- More time on task for primary school children, resulting in more children passing their end of primary school examinations.
- More girls able to make the transition to secondary school.

The Business Case document incorporates a TOC section pages 19 - 25, a diagrammatic representation is inserted at the end of this TOR, Annex A. The central tenet of this TOC is that in order to achieve better learning outcomes in a sustainable manner a coordinated and multi-faceted programme is required. The education sector has most of the building blocks but the challenge is that the system does not meet the needs of the children. Key to this evaluation will be to test that the interventions chosen contribute to more and better quality 'time on task' teacher ⇔ student interaction and that this leads to better learning outcomes for students. There is scope for permutations of interventions; for example teacher training, school leadership, teacher motivation by financial, recognition or asset incentives, water, sanitation, girl friendly activities, community mobilisation initiatives to be tested, in conjunction with programme requirements to deliver results at scale.

EQUIP-T implementation is scheduled to start in mid-2013, tendering for the programme's Management Agent started in December 2012. It is intended that the detailed design of the Impact Evaluation (immediately prior to baseline field work) will take place in mid-2013, in tandem with the programme's inception phase.

EQUIP-T Programme Targeting

The 20 -25 selected target councils (districts) have yet to be formally agreed and selected, but are likely to compromise the majority of rural district councils in 5 or 6 regions chosen from the following short list of seven regions, Dodoma, Kigoma, Lindi, Mara, Rukwa, Shinyanga and Tabora. These are remoter, rural areas with high incidence of poverty and most are inland from the more economically developed coastal area. Criteria used to select councils include ranking low in terms of educational attainment and financial resourcing. They will not necessarily be the very worst councils since it is possible that there may be insuperable structural constraints to substantive improvements. Councils have been requested to provide written statements of willingness and intent to actively participate in the programme. Criteria and data used to short list councils is available, it should be stressed that the final council selection will not be random, for logistical reasons selected councils will tend to cluster within regions.

The selected councils will generally not have any major development partner financed programmes active, care has been taken to avoid overlap with UNICEF (School based In-service Teacher Training), USAID (TZ-21) and CSO TWaweza (Cash on Delivery pilot) programmes. Given the recent preference for budget support modalities, there are thought to be fewer perturbations caused by geographically targeted programmes.

Another large source of external funding is due to provide all primary schools with a large volume of core text books (at 1:2 ratio per student) in mid – late 2013. How these affect classroom practice and learning outcomes is of potential interest to this evaluation, however the plans for universal book provision should mean that it has a relatively homogenous effect.

Effectiveness, Fiscal Sustainability and Replicability

3.3.1 The design of EQUIP-T has considered three factors. First the needs of the worst performing districts, second a scale large enough to be able to assess wider fiscal affordability and third the actual portfolio of interventions and their costs and effectiveness.

3.3.2 The overall cost of the programme has been justified in economic terms on additional income of primary school graduates within the target population alone - as little as £1.22 extra income per graduate per month is estimated to pay back programme costs - but the fiscal sustainability of this also had to be considered carefully. The programme as currently designed might present fiscal sustainability challenges (annually costing about £9.50 per child covered in addition to the current estimated government expenditure of £55 per child). However the purpose is to test and learn and devise a suite of recommendations that could be absorbed by reasonable assumptions of tax revenue growth and allocation to education, and by improvements in allocative efficiency.

3.3.3 Attention will be given to the overall affordability of support interventions financed in the focal councils, to ensure that the approach is both replicable and sustainable from a fiscal perspective. Programme design will focus on identifying the most cost effective approaches to system strengthening, to ensure that costs to government in future will be substantially lower than the unit costs of EQUIP-T. If scaled nationally by government, programme investments could potentially be multiplied many times over.

3.3.4 EQUIP-T will also explore which approaches are most cost effective to key questions such as optimal approaches to supervision and approaches for teacher attraction and retention, particularly in rural and remote areas. The Evaluation proposed needs to carefully consider these intentions in its design, to ensure that the context in which education outcomes are influenced has the potential to positively influence national education policy and reform.

Governance

3.4.1 The EQUIP-T programme will be managed by a service provider – the *Management Agent* (MA), currently under procurement by DFID Tanzania, which will establish regional support teams and work closely with central and local government authorities to reform and improve primary education services. Reforms and activities implemented by the programme will be additional to the regular government financing and educational programmes. Generic areas of intervention have been described, but it will be to a large extent for the treatment group of Councils to determine their own 'holistic' action plans, with support, mentoring and services from the MA. Some funds may also be transferred to the councils for specific activities, but the scope and mechanisms for these transfers are yet to be determined.

3.4.2 The *Evaluation Service Provider* will be independent, but required to work closely with the government, MA and DFID Tanzania throughout the life of the programme, including agreeing the evaluation design, hypotheses for testing, outcomes of concern, appropriate indicators and a dissemination strategy. Whilst the evaluation must

generate internationally credible evidence, the research and associated findings must have government buy-in and acceptance to be useful for instigating reform and change.

3.4.3 An EQUIP-T *Steering Committee* will be formed at the outset of the programme, to include DFID, the Prime Minister's Office Regional and Local Government (PMORALG), the Ministry of Education and Vocational Training (MoEVT), the MA, and key implementation partners.

3.4.4 The steering committee will appoint a broader evaluation *Reference Group* to provide guidance on the evaluation design and oversight of its implementation and dissemination, this will require expertise in education research and evaluation (including evaluation methods) and financial analysis).

Evaluation purpose, objectives and scope

The purpose of this requirement is to provide technical assistance for a prospective, rigorous impact evaluation of the EQUIP-T programme in Tanzania, including the following services:

Research services: design and oversee all aspects of the implementation of the impact evaluation.

Data collection and quality assurance services: identify and provide direct technical support to the survey firm(s) on data quality assurance related to the design and implementation of the impact evaluation.

Learning, knowledge management and dissemination services: help translate lessons from the Impact Evaluation to benefit Tanzania Education Policy and Practice.

The purpose of the impact evaluation (IE) of the EQUIP-T will be twofold: (i) *to assess if the EQUIP-T interventions in supported councils contribute to better basic learning outcomes amongst primary school age students.* (ii) *Which specific support interventions and measures of quality service provision were most significant in improving learning outcomes and to what extent are these replicable and affordable in the Tanzanian / E. African context.* The IE must ensure that the evidence is used to promote lesson learning, accountability, and understanding of the cost effectiveness and potential of the programme's intervention and approach.

The evaluation will address DFID Tanzania's needs for accountability around this phase of the EQUIP-T programme, and to generate evidence to inform the viability and structure of a potential second phase to the programme beyond 2016. It will increase the international evidence base on influencing learning outcomes in a resource constrained environment for all development partners and national governments. It will contribute to DFID's commitments under its Evaluation Strategy.

The evaluation will also enable the Government of Tanzania to take evidence-based decisions in the design of future education programmes beyond 2015, in particular around teacher strategies, decentralisation by devolution of schools and primary level pedagogy and learning assessment. The evaluation provider will need to demonstrate that an adequate dissemination strategy is in place to facilitate this, and work with the programme's MA to ensure evaluation findings are incorporated into the communication strategy under output six of the programme.

The evaluation will test hypotheses around the following criteria, which are in line with the OECD-DAC evaluation criteria, and with DFID's policy on evaluation. The hypotheses draw on the programme TOC attached below. Bidders will be expected to comment on the hypotheses set out here and propose revisions and prioritisation as appropriate.

Relevance

Do the interventions undertaken through this programme provide an evidence based replicable and affordable approach that can be nationally implemented by the Tanzanian Government to focus primary educational resources on quality improvements?

Effectiveness

Have the programme interventions targeted the most necessary, most economical and appropriate combination of interventions for improvements in the quality of education?

Has student ⇔ teacher 'time on task' been significantly increased in target schools?

Are better pedagogic practices that promote effective learning, demonstrably in place?

Impact

Have the target councils been able to increase learning outcomes for girls / boys, including disadvantaged children, beyond those more generally obtained in comparable areas?

Sustainability

Do councils have costed plans in place that are realistic both fiscally and institutionally for the long term maintenance of quality within schools including provision and quality of teachers, operations, inputs and maintenance of school infrastructure?

These hypotheses will be tested utilising the following outcome related information areas that are derived from the logical framework of the EQUIP-T programme. Attention should be given to cross cutting equity and exclusion issues. Specific indicators will need to be agreed in advance with the MA and the steering committee. Again, bidders will be expected to comment on these at this stage and propose any necessary revisions:

Affordability / costs: unit costs of targeted services, in particular expenditure on teacher training and incentive packages and teaching inputs/materials. Outcomes at household/child level will need to consider out-of-pocket parental expenditure on education, in-kind contributions and opportunity costs.

Utilisation: enrolment, dropout and attendance rates of children in school (and out of school children) and the attendance and actual 'time on task' of teachers.

Quality: outcomes at the impact level are specifically proposed around age specific literacy and numeracy learning outcomes and formal end of primary (grade 7) examination pass rates. The *Opportunity to Learn* approach (described in the attached *Provisional Design Outline*) is potentially a viable basis to construct a proxy quality measure of primary education.

Child Protection: schools are inclusive 'safe spaces' with adequate environmentally friendly water and sanitation provision, with measures to prevent and report abuse, especially of the girl child and disabled.

Data where possible should be disaggregated, in particular by gender, grade, age and disability. Poverty measures should also be included through proxy indicators such as education, household assets, household food security, and household monthly expenditures.

The evaluation should refer specifically to the effectiveness of the intervention in addressing financial and non-financial barriers to learning for the poorest, to addressing issues of exclusion for vulnerable groups, including girls and disabled, and to the particular impact of the intervention upon these groups.

The evaluation should identify any unintended outcomes of the intervention for children, parents, schools, teachers, local authorities or development partners. These may include (but will not be limited to):

- Migration of teachers and students to target schools/districts;
- Impact upon out of school children
- Impact upon pre-primary and secondary school provision.
- Increased parental or community contributions to schools;
- Effects on the local economy and extent of irregular or corrupt behaviour

The final list of evaluation criteria, outcomes and indicators will be agreed by the steering committee of the programme as part of the inception report prior to the baseline survey moving ahead.

Existing information sources

Data where possible will be derived from existing sources including from the Government of Tanzania's demographic, educational and financial statistical data to inform traditional officially reported measures of education access, completion, examination results and cost.

In addition two other major data sources (*UWEZO Learning Outcomes* and *World Bank Service Delivery Indicators (SDI)*) are proposed to be combined with government statistical data and the field survey data collected specifically for this evaluation. These are defined in detail in the *Provisional Design Outline*. Bidders should identify existing relevant evaluations related to learning outcomes in the primary education sub-sector and comment on their implications for this study.

Bidders will also be expected to assess the degree to which use can be made of evidence derived from the programme's M&E systems; and in particular whether such evidence can be considered accurate and precise for use within the impact evaluation. This will be particularly important in terms of assessing the overall affordability and longer-term fiscal implications of any recommendations made as part of the impact evaluation.

Methodology

Bidders are expected to provide a clear description of the methodology they will use to address the issues set out above. They should review and take into account analysis in the attached *Business Case* and *Provisional Design Outline*, and address the following issues.

The evaluation is expected to take a mixed methods approach to respond to the key evaluation questions. Bidders are expected to show how they propose to use a mixed methods approach and how the combination of qualitative and quantitative approaches and data will lead to a robust and credible set of conclusions and recommendations for Government of Tanzania and partners¹⁰⁵. This should include appropriate methods to assess the affordability to Government of Tanzania for national rollout of the recommendations

The core of the evaluation will need be able to show that observed changes in education service provision, learning outcomes, quality and affordability would not have happened in the absence of the intervention (internal validity). Initial investigation indicates that a quasi-experimental design with a 'difference in differences' method is most likely approach to most appropriate and would require identification of a comparable control from the other ~ 140 councils to compare to the 20-25 treatment councils. As important is an overall judgment of the external validity of the model identified for replication nationally by Government of Tanzania. Bidders should clearly set out what approach to assessing external validity they propose and its strengths and weaknesses relative to other approaches.

Bidders are expected to explore whether a quasi-experimental '*difference in differences*' approach is the most feasible, and are encouraged to identify any other viable approaches to construct appropriate counterfactuals/comparisons that fit with the data sources identified. Other possible approaches include randomised phase in of interventions, regression discontinuity (for example comparing districts with similar indicators) and combining matching with a 'difference in differences' approach to allow more detailed gender analysis. Statistical techniques such as propensity score matching may also be used to match the groups on relevant variables. The proposed evaluation design must be able to pass scrutiny of the DFID / World Bank Strategic Impact Evaluation Fund (SIEF) Evaluation QA processes¹⁰⁶.

It is important to ensure that treatment and comparison groups are statistically matched and valid. The Provisional Design Outline contains a brief section on Sampling Considerations with respect to the UWEZO and SDI surveys and how these results and instruments can potentially be exploited. Comments from bidders and proposals on how to address these complex sampling issues in a robust manner are sought; a final decision on whether DFID will fund (separately) these surveys has yet to be made. Bidders will be expected to specify the protocols that will be used to agree and verify the level of accuracy and credibility that is to be expected from this work and how it will be quality assured throughout the evaluation (including by the Steering and Reference Groups). These protocols checklists, which will be annexed to the bids, should be clearly referenced to, or derived from, internationally reputable research sources.

A *mixed methods* approach will allow more qualitative research findings to be integrated with the quantitative field research data collected directly by the evaluation team. Documenting classroom pedagogic practice and researching local community and other stakeholder group preferences and attitudes - in particular those of children, parents, teachers and community leaders - will be essential. Bidders are invited to set out clear proposals on how best to do this in a cost effective manner - such as through the use of case studies and semi-structured interviews, focal group discussions, multi-media recording and school/community score cards.

Operational Research embedded within the programme implementation may also test the processes and variants of education reform under investigation, for example trialling approaches to encourage girls' retention and progression, the impact of textbooks, or how effective teacher incentives can work in remote schools. How best to engage with this should be reviewed during the inception phase.

The IE design will need to be co-ordinated closely between the programme's MA and the evaluation Principal Investigator (PI). The conflicting demands of robustly designed research must be dovetailed with implementation requirements to have a quick and sustained impact. Mutually acceptable pathways must be identified during the inception phase of EQUIP-T. Utilisation of the programme's mainstream M&E programme as suggested in paragraph 5.3 further reinforces this requirement.

It is expected that three surveys will be conducted over a five year period, which will include a baseline (2013), a mid-term review (2015) and a post evaluation (2017). The bid proposal should outline how these surveys will be

¹⁰⁵ In doing so, attention is drawn to DFID's recent work in this area: www.dfid.gov.uk/Documents/publications1/design-method-impact-eval.pdf

¹⁰⁶ An application for additional SIEF resources to extend this evaluation may be made.

co-ordinated with the roll-out of the programme, and how frequently each of the proposed indicators should be measured. Details will be developed as part of the inception report.

A Principal Investigator (PI) should be proposed to lead the impact evaluation of EQUIP-T. The PI will be responsible overall for the design, implementation, analysis and dissemination of results for the impact evaluation.

The Principal Investigator (PI) will need to lead the IE through these three phases and manage the synthesis the data sources for joint analysis and lead the drafting, discussion and dissemination of findings from each phase - each will require a comprehensive, professional output report.

There is a possibility of further downstream evaluation work linked to a possible extension of EQUIP-T beyond 2016, see *Timing* section below for details.

In addition, the service provider should consider how best to address the following issues in their design proposal:

How best to ensure a participatory approach is taken to the design and implementation of the evaluation. Bidders should comment on how they will ensure that this approach incorporates girls / women and marginalised groups.

How best to separate out real changes in service provision and usage from improved measurement of changes (possibly arising from new integrated financial management information system and human resource databases, more accurate demographic information or adjusted learning assessment techniques)

How to assess the merits of alternative explanations for the outcomes which are observed, other than the expected influence of the DFID funded EQUIP-T programme. It should take account of the outcomes and impacts of parallel interventions, such as budget support, local government reforms, Civil Society accountability initiatives and teacher deployment policy changes.

The extent to which the results can be extrapolated nationally, given the non-random selection of district councils proposed.

How best to distinguish between “theory failure” and “implementation failure”. The programme may have failed because of faulty beliefs about causal linkages in the design document, but still have been implemented competently. Or the basic design idea may have been sound, but implementation difficult in practice. These differences have implications for relevant policy advocacy efforts.

Identification of major anticipated risk and strategic approaches to mitigation.

Required Skills Mix, Logistics and Procedures

Within the proposed team a mix of competencies will be necessary that include the following, a mixed international/Tanzanian team is highly desirable: a Principal Investigator (PI) with skills and abilities to successfully manage the evaluation using mixed methods, at least one core member with skills and abilities to deliver quasi experimental designs and their application, strong skills in qualitative research in the education sector, capacity to carry out the financial affordability work. Strong skills in design and use of surveys, education reform and measuring educational performance and gender/poverty analysis.

The evaluation Service Provider will be expected to supply their own logistic requirements including office space and transport.

The evaluation Service Provider is expected to undertake the evaluation independently, deploying personnel for survey design, data collection and analysis, and report production. It will be expected that the same firm will be retained throughout the project period, depending upon satisfactory completion of deliverables, to ensure consistency of survey execution and to build on historical knowledge. Bidders should comment on how independence can be maintained from the programme implementing entity (MA and education offices at the local and central level), given the need for a very close working relationship through the life of this evaluation.

The evaluation will be published in full by the evaluation provider, regardless of findings. Data sets will be made available to other researchers for analysis, with due consideration given for the privacy of respondents.

It is expected that the evaluation should conform to OECD-DAC principles of accuracy and credibility¹⁰⁷, and to the evaluation principles set out in the DFID policy on evaluation for international development. Bidders should set out how they will ensure the study is ethically sound and with which relevant ethical protocols it will comply. Perspectives of those involved in the programme’s geographic scope must be included and diverse and dissenting views be heard, recorded and considered. DFID strongly subscribes to a rights based approach in both its programmematic and evaluation work.

¹⁰⁷ <http://www.oecd.org/dac/evaluation/50584880.pdf>

Outputs

Outputs are expected to include:

- An inception report and work plan including study design with the M&E framework, sampling frame, power calculations, draft tools, and proposed analytical methods by month 2 of the contract;
- A finalised design of the evaluation by month 4 of the contract, including the development of survey instruments that have been field tested;
- Evidence of adhering to ethical protocols, including protocols to interact with children and how to handle suspected cases of abuse;
- A draft baseline survey report within 2 months of the finalisation of the survey and before 31st December 2013;
- A publication and dissemination strategy (for GoT/ DFID consideration) that meets diverse stakeholder needs by month 5 of the contract;
- Regular three-monthly up-date reports to DFID and the steering committee and reference group;
- Production of a series of evaluation reports using quantitative and qualitative methods after the baseline, including a mid-term review and a post assessment;
- A specific high level report using the mid-term review findings to disseminate to senior DFID and government stakeholders the potential for replication and funding extensions to EQUIP-T;
- Analytical reports on specific issues; these will be agreed during the inception phase and set out in the work plan.

Key Performance Indicators (KPIs) will be set to ensure rapid mobilisation and adherence to the design and timely delivery of outcomes of this impact evaluation (see section 8 of this TOR). The KPIs will be identified by DFID in consultation with the EQUIP-T Steering Committee. The exact wording of the sub-criterion for the first phase (baseline) will be agreed between DFID and the Service Provider prior to contract signature, and for the whole contract prior to the completion of the baseline survey report. KPIs will ensure that the management of the contract is undertaken as transparently as possible and to ensure that there is clarity of roles and responsibilities between DFID, the Service Provider and the EQUIP-T Management Agent.

All findings, datasets and methods for the evaluation component project must be published and made available to allow researchers to replicate findings. Publication in peer reviewed journals must be attempted with results made available in a timely fashion.

Timing

The duration of the contract is expected to be from July 2013 to December 2017 with three distinct phases related to the baseline, mid-term and final evaluation work. Break points will be specified in the contract and linked to satisfactory performance, delivery of the specified outputs and continued programme implementation.

A follow on contract of up to 5 years may be awarded (subject to a new competitive tender), if it is decided to lengthen the evaluation and finance a further programmematic extension to EQUIP-T beyond 2016.

An inception report will be provided by September 2013. All timings for outputs will need to be coordinated with the programme steering committee.

Duty of Care and Logistical Arrangements

10.1. The Service Provider is responsible for the safety and well-being of their Personnel (as defined in Section 2 of the Framework Agreement) and Third Parties affected by their activities under this contract, including appropriate security arrangements. They will also be responsible for the provision of suitable security arrangements for their domestic and business property.

10.2. DFID will share available information with the Service Provider on security status and developments in-country where appropriate. DFID will provide the following:

All Service Provider Personnel will be offered a security briefing by the British Embassy/DFID on arrival. All such Personnel must register with their respective Embassies to ensure that they are included in emergency procedures.

A copy of the DFID visitor notes (and a further copy each time these are updated), which the Supplier may use to brief their Personnel on arrival.

The Service Provider is responsible for ensuring appropriate safety and security briefings for all of their Personnel working under this contract and ensuring that their Personnel register and receive briefing as outlined above. Travel advice is also available on the FCO website and the Supplier must ensure they (and their Personnel) are up to date with the latest position.

This Procurement will require the Service Provider to operate in a seismically active zone and is considered at high risk of earthquakes. Minor tremors are not uncommon. Earthquakes are impossible to predict and can result in major devastation and loss of life. There are several websites focusing on earthquakes, including <http://geology.about.com/library/bl/maps/blworldindex.htm>. The Service Provider should be comfortable working in such an environment and should be capable of deploying to any areas required within the region in order to deliver the Contract (subject to travel clearance being granted).

The Service Provider is responsible for ensuring that appropriate arrangements, processes and procedures are in place for their Personnel, taking into account the environment they will be working in and the level of risk involved in delivery of the Contract (such as working in dangerous, fragile and hostile environments etc.). The Service Provider must ensure their Personnel receive the required level of training and safety in the field training prior to deployment.

A.2 Agreed terms of reference

A.2.1 Our understanding of the agreed terms of reference

During the contracting and inception phases it was agreed by DFID and OPM that the scope of the impact evaluation needed to be reduced from that outlined in the original terms of reference (TOR) (see section A.1) and that not all of the original TOR objectives could be met by the impact evaluation¹⁰⁸.

This section begins by setting out the original impact evaluation purpose and then discusses the implications of DFID's design choices during the contracting and inception phases. It goes on to outline what the impact evaluation will measure and the evaluation questions to be answered. Finally, it sets out the revised impact evaluation purpose.

A.2.1.1 Impact evaluation purpose in original TOR

According to the original TOR the purpose of the impact evaluation of the EQUIP-T programme is twofold:

- “Assess if the EQUIP-T interventions in supported councils [districts] contribute to better basic learning outcomes amongst primary school age students.”; and
- Assess “which specific support interventions and measures of quality service provision were most significant in improving learning outcomes and to what extent are these replicable and affordable in the Tanzanian/E. African context.”

The original TOR also specified that:

- “The IE must ensure that the evidence is used to promote lesson learning, accountability, and understanding of the cost effectiveness and potential of the programme’s intervention and approach.”

A.2.1.2 Impact evaluation design options

OPM's technical proposal (May 2013) provided DFID three impact evaluation design approaches to choose between¹⁰⁹.

¹⁰⁸ DFID-OPM correspondence March 10, 2014.

¹⁰⁹ OPM's technical proposal is available on request.

The first option, the gold standard randomised control trial (RCT) approach would have been able to identify which specific EQUIP-T interventions were most effective in improving learning outcomes and programme scalability, but would have required the specification of multiple treatment groups across which to randomise assignment of programme exposure.

The second option, the hybrid approach, offered an intermediate option. Under this approach, a base package of EQUIP-T interventions would have been implemented in all treatment schools. Then the treatment group of schools would have been split into treatment sub-groups with an additional EQUIP-T programme intervention randomly assigned to each sub-group. This would have allowed for the assessment of whether adding specific EQUIP-T interventions to the base package led to further relative improvements in the key indicators and scalability.

For the final option, the basic approach, no attempt would be made to control the roll-out of specific EQUIP-T interventions within the EQUIP-T programme districts, allowing assessment of the impact of the EQUIP-T programme as a whole, but not of relative importance of different EQUIP-T components in improving learning outcomes or scalability.

The technical proposal also provided two options for the sample size for the quantitative baseline survey. The first option was for 100 EQUIP-T (treatment) schools and 100 non-EQUIP-T (control) schools, a total of 200 primary schools. The changes in the baseline proportion of pupils meeting Kiswahili and mathematics proficiency requirements detectable with this sample size are given in the EQUIP-T Impact Evaluation Inception Report (OPM 2014a). This was the minimum sample size to detect the expected effect size changes. A second option that would yield higher precision and improve the ability to detect EQUIP-T programme effects was offered, with a sample of 150 treatment schools and 150 control schools, a total of 300 primary schools.

For all three quantitative design approaches outlined above and provided in the technical proposal, the qualitative research would provide additional detail on issues around gender, reasons for observed changes in pupil learning levels, data on district education management and community participation in and demand for accountability in education. These qualitative data would be complementary to the quantitative survey data, but would not in themselves provide a theory-based evaluation or a rigorous attribution of impact to different EQUIP-T components.

A.2.1.3 DFID design choices

The three quantitative design options: gold standard, hybrid and basic approaches (see section A.2.1.2), were discussed during the contracting and inception phases and DFID selected the **basic approach** because of a preference for implementation of all EQUIP-T programme components in all EQUIP-T districts at approximately the same time and for cost considerations, as the gold standard and hybrid approaches would have required larger sample sizes and additional research activities and therefore would have been more costly than the basic approach.

The two **sample size** options provided: a total of 200 vs 300 schools were also discussed during contracting and inception and DFID selected the 200 school sample size.

Based on discussions with DFID and comments from the Specialist Evaluation and Quality Assurance Services (SEQAS), the qualitative research design was revised to include **development of an enhanced programme TOC** including contextual factors, priority parts of which will be tested during the follow-up rounds of the qualitative fieldwork.

A.2.1.4 What the impact evaluation will measure under the agreed terms of reference

The main focus of the impact evaluation will be to measure any EQUIP-T impact in the EQUIP-T programme districts covered by the IE and to provide accountability for the UK taxpayer in terms of the

impact of resources used. The impact evaluation will also provide evidence on programme cost-effectiveness and fiscal affordability (separate fiscal study), promote lesson learning across districts and provide indications to DFID and the Government on which EQUIP-T programme components may likely be more effective in improving pupil learning outcomes.

Original TOR purpose 1

The evaluation will provide quantitative evidence on **the impact of the EQUIP-T programme on learning outcomes for primary school pupils** supported by qualitative research findings that will probe gender aspects and reasons for changes in pupil learning levels.

Original TOR purpose 2

Given DFID's choice of the basic approach (section A.2.1.3A.2.1.1), the impact evaluation will explore other possibilities for understanding which EQUIP-T components may be more effective in improving pupil learning outcomes (assuming there is impact) and scalability.

The quantitative component will explore whether it will be possible to take advantage of any naturally occurring variation in roll-out of specific EQUIP-T interventions within the evaluation treatment sample, in order to identify **impact of specific interventions**. However, without random assignment of specific interventions or without stratifying the sampling of treatment schools by package of interventions (see above) it is unlikely that there would be enough variation in the sample to robustly identify differential impact.

It should be noted that the original TOR did not specify a theory-based impact evaluation nor was the development of a TOC required beyond that developed by the MA part of the original TOR for the IE. However, in light of SEQAS comments and discussions with DFID, the IE design has been revised to set out a process whereby the qualitative research will develop an enhanced TOC including contextual factors (see chapter 4 in Volume II).

The EQUIP-T programme TOC informs the IE as a whole, but is particularly important for the qualitative component because it permits stronger **generalisation and some attribution of impact**. Specifically, the EQUIP-T TOC change was used to map out EQUIP-T's causal chain and the contextual assumptions that must hold for EQUIP-T activities to lead to the desired impact (White 2009). The IE will use (primarily) qualitative data to conduct 'rigorous factual analysis' on whether the expected links in the causal chains hold and whether the assumptions are valid over time, for some of the links in the causal chain selected on the basis of their perceived importance by key stakeholders.

While this is not a theory-based evaluation in the pure sense because it is not comprehensive on all causal pathways, the IE will use theory to produce results on which components of EQUIP-T are likely to contribute to changes in key outcomes and outputs in different contexts. This will yield **indicative results on which interventions were perceived to be more effective**, and coupled with secondary data analysis of the context in other areas of Tanzania to check whether these contextual assumptions hold there as well, this will enable **consideration of the likely impact of EQUIP-T if implemented at scale**.

Following the discussions and agreements with DFID during the contracting and inception phases, the primary aim of the impact evaluation will be to measure the impact of EQUIP-T over time. To do this the design of the quantitative component seeks to maximise internal validity. The EQUIP-T regions and districts were purposively selected by the MA on the basis of region rankings and district rankings in terms of education performance and financial resources and include primarily rural districts (see OPM 2014a). A large majority of rural districts in Tanzania share similar characteristics and therefore

although the IE impact results will not be statistically generalizable outside the IE sample, it is **reasonable to expect that the findings will have some applicability in other districts as well if sufficiently similar to the treatment districts**, other things being equal. The impact evaluation will use, among other things, the rich dataset compiled for the quantitative baseline sampling frame to compare EQUIP-T districts along several key characteristics including education performance, infrastructure, poverty measures and population density, to similar districts not participating in EQUIP-T to assess the potential for generalisation.

The impact evaluation will also assess **cost-effectiveness of the EQUIP-T programme** and the fiscal affordability of rolling out EQUIP-T to regions and districts beyond the initial programme areas in a separate fiscal study (see OPM 2014a).

A.2.1.5 Evaluation questions

The original TOR specified key questions related to the OECD-DAC evaluation themes of relevance, effectiveness, impact and sustainability, to be answered by the impact evaluation. These questions are shown in Annex table 1, together with what the impact evaluation will measure, given the EQUIP-T programme design and changes to the TOR agreed with DFID.

Annex table 1 What the impact evaluation will measure

Original terms of reference evaluation questions	Measured by the impact evaluation under agreed TOR
"Have the programme interventions targeted the most necessary, most economical and appropriate combination of interventions for improvements in the quality of education?"	Partly, as the EQUIP-T components had already been determined by the MA during their inception phase and differential roll out of different EQUIP-T components was not deemed possible by DFID. The qualitative component will examine which EQUIP-T components were perceived to raise education quality.
"Has pupil-teacher 'time on task' been significantly increased in target schools?"	Yes.
"Are better pedagogic practices that promote effective learning, demonstrably in place?"	Yes.
"Have the target councils been able to increase learning outcomes for girls / boys, including disadvantaged children, beyond those more generally obtained in comparable areas?"	Yes.
"Do councils have costed plans in place that are realistic both fiscally and institutionally for the long term maintenance of quality within schools including provision and quality of teachers, operations, inputs and maintenance of school infrastructure?"	Partly if possible. The qualitative component will through the district level interviews attempt to collect information on the availability of costed plans, but not their quality, for the EQUIP-T programme councils (districts) selected as qualitative research sites.
"Improved education quality."	Yes.
"Improved teaching of early-grade reading and numeracy resulting in more children able to read with comprehension" and with curriculum appropriate numeracy skills."	Yes.
"Improved teaching of early-grade reading and numeracy resulting in more children able to read with comprehension" and with curriculum appropriate numeracy skills."	Yes.
"More time on task for primary school children, resulting in more children passing their end of primary school examinations"	Yes.
"More girls able to make the transition to secondary school".	No, as the EQUIP-T programme will focus on the early grades and impact of the programme, if any, on this outcome would be highly unlikely to be detectable within the life of the impact evaluation.
Pupil learning results should be disaggregated by gender.	Yes.

The impact evaluation should examine impact for disabled children.	No, because the EQUIP-T programme does not contain any component or activities aimed at this particular group ¹¹⁰ .
The impact evaluation should include poverty measures for pupils.	Yes.

A.2.1.6 Revised purpose of the impact evaluation

Thus the impact evaluation will:

- Generate evidence on impact of EQUIP-T on learning outcomes for pupils in primary education, including any differential impacts for girls and boys;
- Provide evidence on the fiscal affordability of scaling up EQUIP-T beyond the initial EQUIP-T regions and districts (separate fiscal study);
- Assess perceptions of effectiveness of different EQUIP-T components through the qualitative research and explore possibilities to do so through the quantitative component; and
- Communicate evidence generated by the impact evaluation to policymakers and key education stakeholders, including DFID and MOEVT to promote accountability and lesson learning.

A.2.2 Changes to the impact evaluation design since the technical proposal

In addition to the reduction of scope of the TOR outlined in section A.2.1, the following changes to the IE design compared to the technical proposal were made based on discussions with DFID during the inception phase and feedback from the first Reference Group meeting for the impact evaluation¹¹¹.

- Scope of impact evaluation expanded for the qualitative component to examine perceived EQUIP-T contributions to changes in relevant outcomes and outputs, to the extent possible within the scope of the IE;
- Quantitative fieldwork to start in March 2014 (first start date was October 2013, second start date was January 2014);
- Use of EGRA/EGMA style pupil learning assessments instead of UWEZO assessment testing 3,000 standard 3 pupils in Kiswahili (EGRA) and mathematics (EGMA);
- Test standard 3 pupils (Kiswahili and mathematics) instead of standard 2 and standard 5 pupils;
- Administer teacher development needs assessment (TDNA) to standards 1-3 (Kiswahili and mathematics) and 4-7 (mathematics) teachers instead of to standard 2 and standard 5 teachers;
- One standard 2 Kiswahili and one standard 2 mathematics lesson will be observed for each sample school instead of one standard 3 lesson and one standard 5 lesson;
- Replace pupil tracer survey to collect data for poverty measure by data collected at school level (from tested pupils' parents);
- To obtain school sample: in the second stage, match control schools to treatment schools using PSM instead of random selection;
- 17 EQUIP-T programme (treatment) districts in the five EQUIP-T regions covered by the impact evaluation will be surveyed instead of 20 districts due to contamination by other education programmes or projects.

¹¹⁰ MA-OPM and DFID-OPM correspondence December 2, 2013.

¹¹¹ The draft inception report was subsequently further revised based on comments received from the SEQAS review on March 7, 2014.

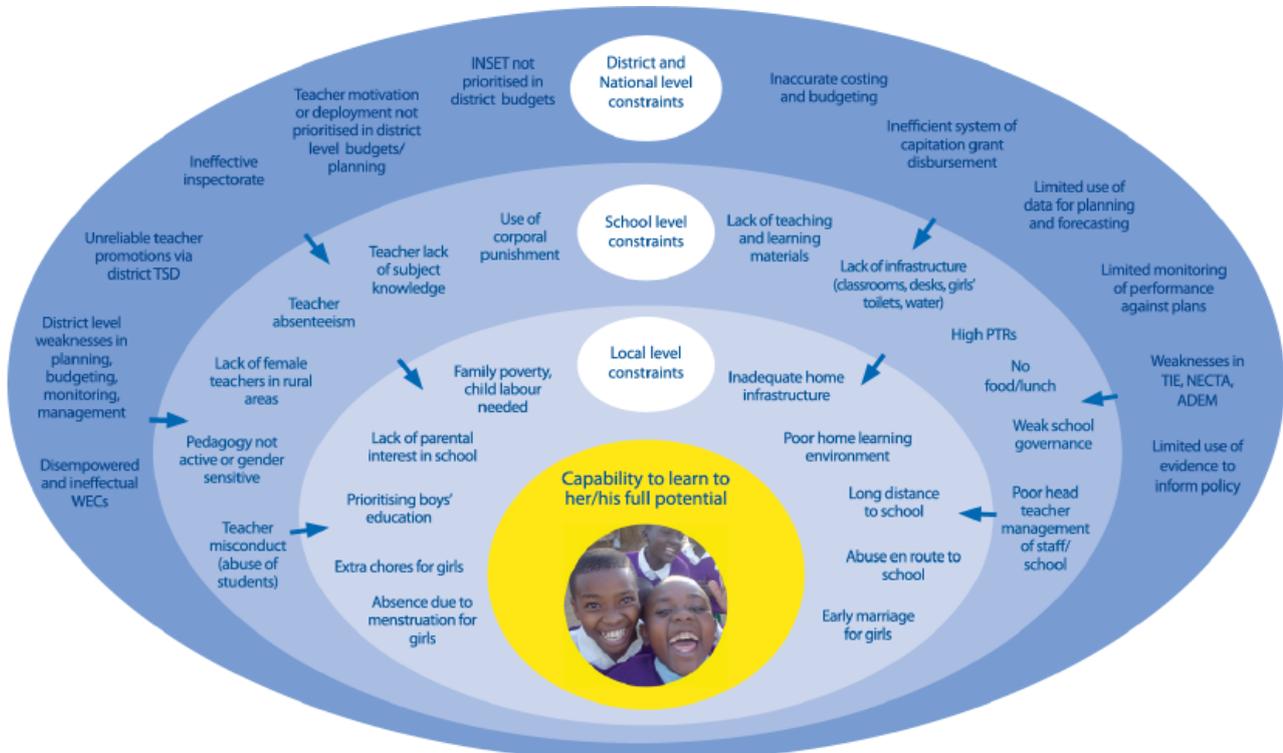
Annex table 2 Impact evaluation regions and districts

Control/treatment	Region	District
Control regions and districts in IE study	Arusha	Ngorongoro DC
	Mwanza	Misungwi DC
	Pwani	Rufiji DC
	Rukwa	Nkasi DC
	Ruvuma	Tunduru DC
	Singida	Ikungi DC Singida DC
	Tanga	Kilindi DC
Treatment districts in IE study (Note: all 17 districts are part of the quantitative survey, * indicates they are also part of the qualitative research)	Dodoma	Bahi DC
		Chamwino DC
		Kongwa DC
		Mpwapwa DC *
	Kigoma	Kakonko DC
		Kibondo DC
	Shinyanga	Kishapu DC *
		Shinyanga DC
	Simiyu	Bariadi DC
		Bariadi TC
		Itilima DC
		Maswa DC
		Meatu DC
	Tabora	Igunga DC
		Nzega DC
		Sikonge DC
		Uyui DC *
Treatment districts that are not part of the IE study	Dodoma	Chemba DC
		Kondoa DC
	Kigoma	Buhigwe DC
		Kasulu DC
		Kigoma DC
		Uvinza DC
		Kahama DC
	Shinyanga	Msalala DC
		Ushetu DC
		Busega DC
	Tabora	Kaliua DC
		Uramba DC
	Lindi	Kilwa DC
		Lindi DC
		Liwale DC
		Ruangwa DC
	Mara	Bunda DC
		Butiama DC
		Musuma DC
		Musuma MC
Rorya DC		

Source: OPM IE team

Annex B Constraints underpinning the EQUIP-T programme theory of change

Annex figure 1 Constraints on children’s capability to learn to their full potential



Source: Cambridge Education 2014, p6.

Annex C EQUIP-T programme summary description

C.1 Component 1: improving teacher performance

Key constraints on teacher performance are poor subject knowledge; poor pedagogy; low teacher morale manifesting in, for example, absenteeism, corporal punishment and avoidance of rural posts.

Component 1 objectives are to:

- To enhance the professional capacity and performance of teachers by creating a comprehensive set of sustainable systems and structures to address the constraints above.
- To ensure awareness of equity issues and application of strategies to support girls and other marginalised groups in education delivery.

Component 1 activities (see Annex table 3). The teacher competency framework will be developed first, so that there is a guiding framework for teacher professional development and teacher performance management.

Annex table 3 EQUIP-T component 1 core activities

Outcome	Core Activity	Input	Output
The enhanced professional capacity and performance of teachers	Teacher Competency Framework	A comprehensive framework that clarifies the type and level of competencies that teachers need	Clear, transparent criteria that links and can be used for accreditation, in-service training, performance management
	Teacher Professional Development	School-based in-service training for 49,000 teachers that focuses on self-learning, application, demonstration	Strengthened subject knowledge and pedagogical skills
	Teacher Performance Management	Performance criteria, targets and appraisals for 49,000 teachers	Feedback to improve practice and reward good performance
	Teacher Morale Toolkit	Strategies specifically targeting the morale of 49,000 teachers	Improved teacher attitudes and behaviours at school
	16-week School Readiness Programme	School Readiness training and resources for Teaching Assistants and St. 1 teachers	Children with the social and language skills needed for Early Grade Literacy
	Teacher Training Scholarships for rural candidates	Scholarships for promising Form 4 leavers (particularly girls) so that they can return to their rural villages to teach.	An increase in the number of female teachers in rural schools

Source: Cambridge Education 2014, p32.

C.2 Component 2: school leadership and management

Key constraints on school leadership and management are weak school systems for planning, budgeting, execution and oversight, and poor human resource management by head teachers.

Component 2 objectives are to ensure schools are:

- Transparent and accountable¹¹²; and

¹¹² The text does not explicitly state which accountability relationships are being targeted by EQUIP-T. From subsequent material it appears the programme is targeting the accountability relationship between schools and communities (including

- Achieving optimum efficiency and effectiveness with available resources in delivery of relevant high quality education.

Component 2 activities (see Annex table 4). The school quality framework will be one of the first activities, so that it can be used as a basis for the whole school planning activity. The development of the school information system SIS is intended to strengthen the use of data and analysis in school planning, and to contribute to the transparency and accountability objective above. SIS will be integrated with the EMIS. A set of head teacher competencies (not listed as a separate development activity) will form a framework for the head teacher performance management system. The draft competency framework is below.

Annex table 4 EQUIP-T component 2 core activities

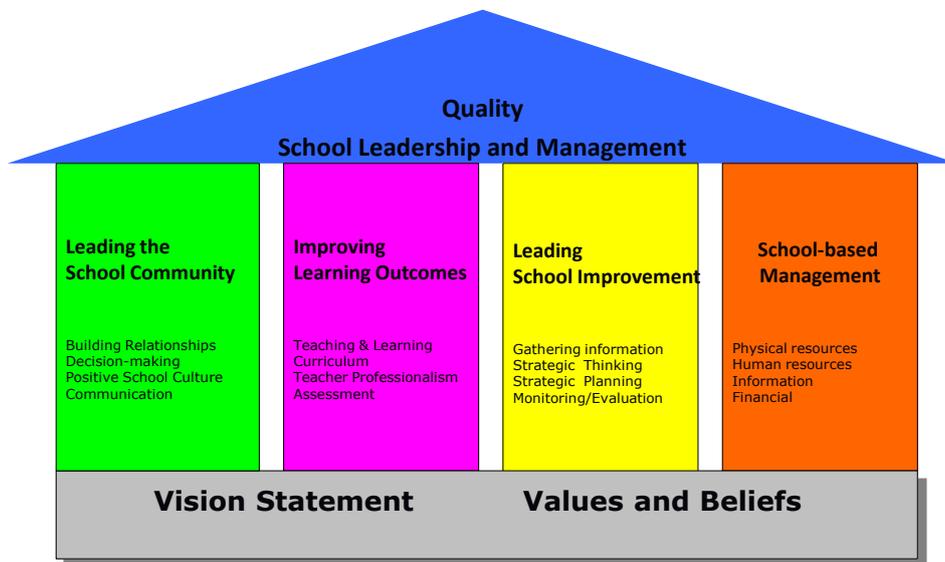
Outcome	Core Activity	Input	Output
Enhanced school leadership and management	Whole School Planning	System development/strengthening for school leadership and management; Professional development for 7500+ Head Teacher/Deputy Head Teachers	More effective Inclusive Whole School Plans, implemented with sound financial management.
	School Quality Framework	Development by MOEVT/PMORALG and stakeholders of a set of minimum standards for primary education; application of minimum standards in school decision-making and performance benchmarking	Generic, minimum achievements incorporated into all school plans to achieve minimum sector goals targeting education outcomes
	School Information System	Data collection, storage, organisation and communication systems for school monitoring, evaluation, reporting and planning. Training of Head Teachers in capturing, analysing and using data.	Transparent, accountable schools able to better determine need and seek resource. Key outputs for EMIS.
	Performance Management System	Development of PMS system (guidelines and tools) for School Leadership providing means of appraisal, professional development pathways and performance recognition. Training of WEC and Head teachers.	Strengthened systems and support for improved school performance.

Source: Cambridge Education 2014, p35.

parents/pupils) and to a lesser extent the relationship between schools and the higher tiers of the education administrative system.

Annex figure 2 National School Leadership Competencies Framework

National School Leadership Competencies Framework



Source: EQUIP-Tanzania (May 2014) 'Draft National School Leadership Competencies Framework'.

C.3 Component 3: district planning and management

Key constraints on district planning and management of education are:

- Weak education information system based on incomplete and poor quality data. This contributes to poor planning, monitoring and evaluation, and a lack of information for education policy making and decisions.
- Inconsistent and unpredictable flow of funds from central to district level, making planning and budgeting very difficult. Some central funds are ear-marked for schools (e.g. capitation grant), resulting in an unpredictable and inconsistent funding situation for schools too. There is considerable friction between districts and schools over funding, partly due to a lack of transparency on the actual flow of funds¹¹³.

Component 3 objective is to strengthen systems and human resource capacity to support sub-national management of education.

Component 3 activities (see Annex table 5). District and regional officials will be provided with formal training and on-the-job mentoring focusing on the core district functions of planning, budgeting, EMIS analysis, public financial management and monitoring and evaluation.

¹¹³ The text says that 'districts are erroneously accused of preventing funding to schools'.

Annex table 5 EQUIP-T component 3 core activities

Outcome	Core Activity	Input	Output
Strengthened systems that support the district and regional management of education	1. Review and assess current District management systems	Baseline survey, Institutional Analysis	Needs Assessment Report and recommendations
	2. Strengthening District planning and management capacity	Capacity building programme to support District education plans and planning cycle	Improved district plans that are results focused, based on school generated data
	3. Support the strengthening of GOT EMIS system	Design, development and implementation of systems to strengthen quality of EMIS data	Districts produce accurate aggregate statistical data, dash boards, and district education indicators
	4. Support the strengthening of District Financial planning, budgeting and management	A comprehensive review and development of PFM tools and district capitation grant mechanism	Districts achieve bench mark targets on management and financial performance
	5. Support the strengthening of district monitoring and Evaluation systems	Internally instituted programme monitoring mechanisms.	Regions/Districts provide 'stock-take' progress reports; hold 6 monthly Education Sector Reviews

Source: Cambridge Education (2014) (pp38-39).

C.4 Component 4: community participation and accountability

Key constraints on community participation and accountability are:

- A lack of transparency and accountability mechanisms between schools and both parents and the wider community. There is a breakdown of both communication and consultation over school planning, operations, challenges and performance;
- School committees are unable to fulfil their mandate, partly because awareness of their mandate is very limited as well as capacity constraints; and
- Public financial resources for schools are seen to be inadequate. The flow of capitation grant funds to schools is unpredictable, and amounts are lower (often much lower) than the policy norm. Parental contributions have been rising, and this, coupled with a lack of transparency in school funding from all sources, contributes to tension between parents and school authorities.

Component 4 objective is to empower communities to take an active role in improving school outcomes and services, by equipping them with tools, resources, and the provision of appropriate accessible information.

Component 4 activities (see Annex table 6): there are five core activities linked to three sub-outcomes which in essence seek to inform, equip and empower communities.

Annex table 6 EQUIP-T component 4 core activities

Sub outcome	Core activity	Description
Communities adequately informed about school operations, needs, plans and performance	1.Improving communication mechanisms	IEC tools and mechanism development. Including use of notice boards, SMS and 3G technology via provision of tablets to all programme schools. Developing clear lines for sharing relevant education information between local government, school committees, parents and the wider community.
Communities and SC have improved capacity and resources to function effectively	2.Community engagement in education planning	Support for a process of participatory needs assessment, planning and resource identification. Community education plans generated by this process will lead to actions to enhance school plans.
	5.School income generating activity	Upgrading existing school-based IGAs or establishing new IGA. Grants provided to develop IGAs based on active engagement in wider EQUIP-T programme, and requirement for clear business plan.
	3.Building the capacity of the school committee	Capacity building workshops for SC, village leaders, led by WECs. WECs to receive capacity building on SC mandate and operations (including PTA formation and role), and teaching methodology. School development plans of the SC will include community priorities and gender analysis.
	4.PTA formation	Formation of PTAs under the guidance of the SC to foster greater class-based teacher-parent collaboration, raise awareness, increase ownership, and develop accountability mechanisms between the school and parents/community. Two seed grants provided for functional PTAs.
Communities engage in school planning, support operations and hold duty bearers to account		

Source: Cambridge Education 2014, pp41-43.

About the project

The independent Impact Evaluation of the Education Quality Improvement Programme in Tanzania (EQUIP-T) is a four-year study funded by the United Kingdom Department for International Development (DFID). It is designed to: i) generate evidence on the impact of EQUIP-T on primary pupil learning outcomes, including any differential impacts for girls and boys; ii) examine perceptions of effectiveness of different EQUIP-T components; iii) provide evidence on the fiscal affordability of scaling up EQUIP-T post-2018; and iv) communicate evidence generated by the impact evaluation to policy-makers and key education stakeholders.

EQUIP-T is a Government of Tanzania programme, funded by UK DIFD, which seeks to improve the quality of primary education, especially for girls, in seven regions of Tanzania. It focuses on strengthening professional capacity and performance of teachers, school leadership and management, systems which support district management of education, and community participation in education.



Oxford Policy Management

