Tanzania - Education Quality Improvement Programme Impact Evaluation Endline Survey 2018

Oxford Policy Management Ltd

Report generated on: November 11, 2019

Overview

Identification

ID NUMBER
TZA_2018_EQUIPIE-EL_v01_M

Version

VERSION DESCRIPTION
Version 2.1: Edited, anonymous dataset for public distribution.

PRODUCTION DATE
2019-07-02

NOTES
Version 2.1 consists of six edited and anonymised datasets (at school, teacher, pupil, lesson and training session level) with the responses to a small number of questions removed; these were removed for confidentiality purposes or because they were not needed for the analysis. Some of the datasets also contain selected constructed indicators prefixed by n_. These constructed indicators are included to save data users time as they require complex reshaping and extraction of data from multiple sources (but they could be generated by data users if preferred).

Overview

ABSTRACT
Education Quality Improvement Programme in Tanzania (EQUIP-T) is a six-year (2014-20) Government of Tanzania programme, funded by the United Kingdom Department for International Development (DFID), which seeks to improve the quality of primary education and to improve pupil learning outcomes, especially for girls. The programme 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. Initially, the programme was intended to run for four years, with activities targeted at seven of the most educationally disadvantaged regions in Tanzania. In 2017 the programme was extended for a further two years, and the extension introduced some new sub-components to the seven regions, and introduced a reduced package of interventions to two new regions.

The independent Impact Evaluation (IE) of EQUIP-T is a five-year study funded by DFID. It is designed to: i) generate evidence on the impact of EQUIP-T on primary pupil learning outcomes, including any differential effects for boys and girls; ii) examine perceptions of effectiveness of different EQUIP-T components; iii) provide evidence on the fiscal affordability of scaling up EQUIP-T post-endline; and iv) communicate evidence generated by the impact evaluation to policy-makers and key education stakeholders. The evaluation uses a quasi-experimental approach to quantitative estimation of impact that combines propensity score matching (PSM) with difference-indifferences (DID).

The research priorities for the quantitative endline IE are captured in a comprehensive endline evaluation matrix (see Annex C in the 'EQUIP-Tanzania Impact Evaluation. Endline Quantitative Technical Report, Volume I: Results and Discussion' under Reports and policy notes). The matrix sets out evaluation questions linked to the programme theory of change. It asks questions related to the expected results at each stage along the results chain (from the receipt of inputs to delivery of outputs, and contributions to outcomes and impact) under each of the programme's components. The aim is to establish: (i) whether changes have happened as expected; (ii) why they happened or did not happen (i.e. whether key assumptions in the theory of change hold or not); (iii) whether there are any important unanticipated changes; and (iv) what links there are between the components in driving changes.

The main IE research areas are:

-Impact of EQUIP-T on standard 3 pupil learning in Kiswahili and mathematics.

-Impact of EQUIP-T on teacher absence from school and from classrooms.

-Impact of EQUIP-T on selected aspects of school leadership and management.

The IE uses a mixed methods approach that includes:

-A quantitative survey of 100 government primary schools in 17 programme treatment districts and 100 schools in 8 control districts in 2014, 2016 and 2018 covering:

*Standard three pupils and their parents/caregivers;

*Teachers who teach standards 1-3 Kiswahili;

*Teachers who teach standards 1-3 mathematics;

*Schools;

*Head teachers; and

*Standard two lesson observations in Kiswahili and mathematics.

-Qualitative fieldwork in a few treatment schools that overlap with a sub-set of the quantitative survey schools, in 2014, 2016 and 2019, consisting of key informant interviews (KIIs) and focus group discussions (FGDs) with head teachers, teachers, pupils, parents, school committee (SC) members, PTP members, region, district and ward education officials and EQUIP-T programme staff.

The endline data available in the World Bank Microdata Catalog are from the EQUIP-T IE quantitative endline survey conducted in 2018. The endline qualitative research will take place in mid-2019 with results available in early 2020.

KIND OF DATA
Sample survey data [ssd]

UNITS OF ANALYSIS
- School

- Teacher

- Pupil

- Lesson (not sampled)

- EQUIP-T training session

Scope

NOTES
The scope of the EQUIP-T IE Endline Survey includes:

-HEAD TEACHER/HEAD COUNT/SCHOOL RECORDS: Head teacher background information, frequency/type of school leadership and management in-service training received, availability and contents of school development plan, teacher management, frequency of staff meetings, ward education officer supervision and support to the school, JUU club activities, school committee, school information system, Parent-Teacher-Partnership activities, community engagement, head teacher morale and other conditions of service, head teacher attendance, reasons for head teacher and teacher absenteeism (reported by head teachers), teacher attendance (from school records and by headcount on the day of the survey), teacher punctuality, pupil attendance (from school records and by headcount on the day of the survey), pupil enrolment, school background information (teachers, physical facilities, school timetable, number of days school open), school characteristics, infrastructure and funding, receipt of in-kind resources.

-STANDARD 3 PUPILS: Pupil background information, pupils' school experience, Kiswahili Early Grade Reading Assessment (EGRA) and Early Grade Mathematics Assessment (EGMA) based on standards 1 and 2 national curriculum requirements. Note: The same pupils were assessed in both Kiswahili and mathematics.

-PARENTS OF SAMPLED STANDARD 3 PUPILS: household and parental characteristics (that can be used to convert scores into poverty likelihoods based on a pre-existing instrument), pupil background, home support for schooling, parent-school communication.

-TEACHERS WHO TEACH STANDARDS 1-3 KISWAHILI AND/OR MATHEMATICS: Interview including background information, qualifications, frequency/type of in-service training received, classroom teaching and pupil assessment practices, access to teaching and learning resources, support for teaching (lesson planning, observation, meetings), frequency/nature of performance appraisal and teacher morale and other conditions of service.

-LESSON OBSERVATION: Standard 2 Kiswahili and mathematics lesson observations of inclusive behaviour of teachers with respect to gender and spatial location of pupils, key teaching practices in the classroom, pupils' reading and teacher support, availability of lesson plan, availability of seating, availability and use of teaching and learning materials during the lesson.

-TEACHER GROUP INTERVIEW (including In-service Training Coordinator): Frequency and nature of all EQUIP-T early grade teacher training sessions that have taken place away from school and in school since baseline, INCO background information, participation in ward cluster reflection meetings.

TOPICS

Topic	Vocabulary	URI
Education		
Primary education		
Impact evaluation		
Girls’ Education		

KEYWORDS

Primary education, Education quality, Pupil learning, Student learning, Pupil learning assessment, Early Grade Reading Assessment, EGRA, Early Grade Mathematics Assessment, EGMA, Impact evaluation, Mixed methods evaluation, Teaching practices, Pedagogy, Teacher motivation, Teacher absenteeism, Classroom absenteeism, School leadership and management, Teacher support, District education management, Instructional time, Community participation, In-service training

Coverage

GEOGRAPHIC COVERAGE

The survey is representative of 17 EQUIP-T programme treatment districts. The survey is NOT representative of the 8 control districts. For more details see the section on Representativeness in 'EQUIP-Tanzania Impact Evaluation. Final Baseline Technical Report, Volume I: Results and Discussion' and 'EQUIP-Tanzania Impact Evaluation. Final Baseline Technical Report, Volume II: Methods and Technical Annexes' under Reports and policy notes.

The 17 treatment districts are:

-Dodoma Region: Bahi DC, Chamwino DC, Kongwa DC, Mpwapwa DC

-Kigoma Region: Kakonko DC, Kibondo DC

-Shinyanga Region: Kishapu DC, Shinyanga DC

-Simiyu Region: Bariadi DC, Bariadi TC, Itilima DC, Maswa DC, Meatu DC

-Tabora Region: Igunga DC, Nzega DC, Sikonge DC, Uyui DC

The 8 control districts are:

-Arusha Region: Ngorongoro DC

-Mwanza Region: Misungwi DC

-Pwani Region: Rufiji DC

-Rukwa Region: Nkasi DC

-Ruvuma Region: Tunduru DC

-Singida Region: Ikungi DC, Singida DC

-Tanga Region: Kilindi DC

Producers and Sponsors

PRIMARY INVESTIGATOR(S)

Name	Affiliation
Oxford Policy Management Ltd	

FUNDING

Name	Abbreviation	Role
Department for International Development UK	DFID	

Metadata Production

METADATA PRODUCED BY

Name	Abbreviation	Affiliation	Role
Harb, Jana		Oxford Policy Management Ltd	Data analyst

DATE OF METADATA PRODUCTION
2019-07-02

DDI DOCUMENT VERSION
Version 1 (July 2019)

DDI DOCUMENT ID
DDI_TZA_2018_EQUIPIE-EL_v01_M

Sampling

Sampling Procedure

Because the EQUIP-T regions and districts were purposively selected (see 'EQUIP-Tanzania Impact Evaluation. Final Baseline Technical Report, Volume I: Results and Discussion' under Reports and policy notes), the IE sampling strategy used propensity score matching (PSM) to: (i) match eligible control districts to the pre-selected and eligible EQUIP-T districts (see below), and (ii) match schools from the control districts to a sample of randomly selected treatment schools in the treatment districts. The same schools are surveyed for each round of the IE (panel of schools) and a cross section of standard 3 pupils and Standard 1-3 teachers will be interviewed at each round of the survey (no pupil panel or teacher panel).

Identifying districts eligible for matching

Eligible control and treatment districts were those not participating in any other education programme or project that may confound the measurement of EQUIP-T impact. To generate the list of eligible control and treatment districts, all districts that are contaminated because of other education programmes or projects or may be affected by programme spill-over were excluded as follows:

-All districts located in Lindi and Mara regions as these are part of the EQUIP-T programme but implementation started later in these two regions (the IE does not cover these two regions);
-Districts that will receive partial EQUIP-T programme treatment or will be subject to potential EQUIP-T programme spillovers;
-Districts that are receiving other education programmes/projects that aim to influence the same outcomes as the EQUIP-T programme and would confound measurement of EQUIP-T impact;
-Districts that were part of pre-test 1 (two districts); and
-Districts that were part of pre-test 2 (one district).

Sampling frame

To be able to select an appropriate sample of pupils and teachers within schools and districts, the sampling frame consisted of information at three levels:

-District;
-School; and
-Within school.

The sampling frame data at the district and school levels was compiled from the following sources: the 2002 and 2012 Tanzania Population Censuses, Education Management Information System (EMIS) data from the Ministry of Education and Vocational Training (MoEVT) and the Prime Minister's Office for Regional and Local Government (PMO-RALG), and the UWEZO 2011 student learning assessment survey. For within school level sampling, the frames were constructed upon arrival at the selected schools and was used to sample pupils and teachers on the day of the school visit.

Sampling stages

Stage 1: Selection of control districts

Because the treatment districts were known, the first step was to find sufficiently similar control districts that could serve as the counterfactual. PSM was used to match eligible control districts to the pre-selected, eligible treatment districts using the following matching variables: Population density, proportion of male headed households, household size, number of children per household, proportion of households that speak an ethnic language at home, and district level averages for household assets, infrastructure, education spending, parental education, school remoteness, pupil learning levels and pupil drop out.

Stage 2: Selection of treatment schools

In the second stage, schools in the treatment districts were selected using stratified systematic random sampling. The

schools were selected using a probability proportional to size approach, where the measure of school size was the standard two enrolment of pupils. This means that schools with more pupils had a higher probability of being selected into the sample. To obtain a representative sample of programme treatment schools, the sample was implicitly stratified along four dimensions:

-District;
-PSLE scores for Kiswahili;
-PSLE scores for mathematics; and
-Total number of teachers per school.

Stage 3: Selection of control schools

As in stage one, a non-random PSM approach was used to match eligible control schools to the sample of treatment schools. The matching variables were similar to the ones used as stratification criteria: Standard two enrolment, PSLE scores for Kiswahili and mathematics, and the total number of teachers per school.

The endline survey was conducted for the same schools as the baseline and midline surveys (a panel of schools). However, the IE does not have a panel of pupils or teachers as a pupil only attends standard three once (unless repeating) and there is high teacher turnover. Thus, the IE sample is a repeated cross-section of pupils and teachers in a panel of schools.

Stage 4: Selection of pupils and teachers within schools

Pupils were sampled within schools using systematic random sampling based on school registers. The within-school sampling was assisted by selection tables automatically generated within the computer assisted survey instruments. Per school, 15 standard 3 pupils were sampled. The parents of these 15 sampled pupils were then interviewed using the poverty scorecard instrument.

For the teacher interviews, as at midline, all teachers of Standards 1-3 who teach Kiswhaili or maths were interviewed to boost the sample size as many schools are small (as opposed to baseline where up to three teachers were sampled within each school for the interviews).

Lesson observations were not randomly sampled. Instead, one maths and one Kiswahili Standard 2 lessons were selected within each school using convenience sampling to be observed on the day of the survey.

Replacement sample

At baseline, if a selected school could not be surveyed it was replaced. In the process of sampling, the impact evaluation team drew a replacement sample of schools, which was used for this purpose (reserve list) and the use of this list was carefully controlled. Five out of the 200 original baseline sample schools were replaced during the fieldwork. At midline and endline, all of the 200 schools surveyed at baseline were visited again (no replacements).

Sample sizes

The actual sample sizes at endline are:

-200 schools (100 treatment and 100 control).
-2,999 standard 3 pupils assessed in both Kiswahili and mathematics.
-2,992 poverty scorecards were administered to the assessed pupils' parent(s).
-889 teachers who teach standards 1 to 3 Kiswahili and/or mathematics interviewed.
-196 standard 2 Kiswahili and mathematics lessons observed (treatment schools only).
-99 teacher group interviews were conducted (treatment schools only).

Note that the lesson observation and the small group teacher interview were only conducted in treatment schools, because the information generated could not be used in the impact modelling and so collecting information in control schools was not necessary.

Representativeness

The results from the treatment schools are representative of government primary schools in the 17 EQUIP-T programme

treatment districts. However, the results from the schools in the 8 control districts are NOT representative because these districts were not randomly sampled but matched to the 17 treatment districts using propensity score matching.

Response Rate

Unit response

Actual sample sizes at endline are close to target sample sizes. See 'Sampling Procedure' section for actual sample sizes.

- All 200 schools surveyed at baseline and midline were also surveyed at endline (100% response rate).
- For tested Standard 3 pupils, response rate is 99.9% (target: 3,000).
- For parents of tested Standard 3 pupils, response rate is 99.7% (target: 3,000).
- For interviewed Standards 1-3 teachers, response rate is 99.2% (target: 906 which is all teachers at endline who are teaching maths and/or Kiswahili to Standards 1-3), including 11% of teachers who were absent or unavailable on the day of the survey and were later interviewed by phone.
- The target for lesson observations (not sampled) was 200, but under the new Standards 1 and 2 curriculum, maths and Kiswahili (either reading or writing) lessons often run sequentially without a break, and this enabled 95 maths lessons to be observed and 101 Kiswahili lessons, in total more than the target (but maths lessons less than the target of a 100).
- For the teacher group interview, response rate is 99% (target: 100).

Item response

Item response rates were generally high.

Weighting

The survey is only representative of the 17 EQUIP-T programme districts and therefore survey weights were only constructed for schools, pupils and teachers in the treatment group (not for the control group).

To obtain results that are representative of the EQUIP-T programme treatment areas, treatment estimates should be weighted using the provided survey weights that are normalised values of the inverse probabilities of selection into the sample for each unit of analysis. The relevant probabilities of selection differ depending on whether analysis is carried out at school, pupil or teacher level, and survey weights for each of these units of analysis are included in the datasets.

School weights (treatment group only)

The probability of being selected of each school depended on the total number of schools being selected and its size relative to the total number of enrolled pupils across all schools in the 17 programme districts. Formally, the probability of a given school being selected into the sample equals the total number of schools sampled multiplied by the ratio of the number of pupils in the given school and the total number of pupils in all schools in the relevant programme areas. The school weights are appropriately normalised inverses of these probabilities.

Note: Refer to the end of this section for the strata, weights and finite population correction factor variables included in the dataset.

Pupil weights (treatment group only)
---
15 standard 3 pupils were randomly sampled at each school. The probability of selection of a pupil in a given school equals the school weight (defined above) multiplied by the ratio of the number of pupils selected per school (15 in all schools except in the schools that had less than 15 pupils present on that day) and the total number of eligible pupils in the given school. The pupil weights are appropriately normalised inverses of these probabilities.

Note: Refer to the end of this section for the strata, weights and finite population correction factor variables included in the

dataset.

Teacher weights (treatment group only)

The probability of selection of a teacher in a given school equals the school weight (defined above) multiplied by the ratio of the number of teachers that were selected for a given teacher instrument per school and the total number of teachers eligible for the given instrument. The teacher weights are appropriately normalised inverses of these probabilities.

NOTE:
-For data from the teacher interviews, the teacher interview weights should be used: weight_tchint. Since all teachers eligible for the interview in each school were interviewed, this means that the selection probability for each teacher is equal to one in this case.
-For data from the teacher roster, the teacher roster weights should be used: weight_teacherroster. Since all teachers in each school are included in the roster, this means that the selection probability for each teacher is equal to one in this case.

Note: Refer to the end of this section for the strata, weights and finite population correction factor variables included in the dataset.

Stratification, clustering and finite population corrections

The survey weights should be used within a survey set-up that takes into account stratification, clustered sampling and finite population corrections.

Stratification during sampling was used at the primary sampling level, that is, at school level, and not at the lower levels (pupil and teacher). For the estimation set-up, strata for schools are defined by districts and teacher-body size terciles. Although, during sampling, schools were implicitly stratified by primary school leaving examination (PSLE) scores as well, this is a continuous variable that cannot be used to define strata in the estimation set-up.

Clustering is only relevant for pupil and teacher level data, as schools were the primary sampling units within the eligible programme treatment districts. School pupil data is also hierarchical in nature with pupils clustered within schools. Hence, for pupil and teacher estimates, clustering is set at the school level.

Because large proportions of the total eligible population were sampled in many schools at the teacher and pupil levels, the estimation set-up should also account for the finite population correction (FPC) factor. This FPC factor is the square root of the ratio of the population from which the sample is drawn minus the size of the sample and the population from which the sample is drawn minus one. In the case of school level data, the FPC factor is constant across all schools, as the sample of schools was drawn from a constant population of all eligible schools in the programme treatment areas. However, for teacher and pupil level data, the FPC factor changes depending on the school, as population sizes and, in the case of teacher level data, sample sizes vary as well.

Stratification, weight, finite population correction and treatment status variables
---
In the EQUIP-T IE endline datasets the stratification, weight, FPC and treatment status variables are as follows:

-The strata variable is: strata

-The school weights variable is: weight_school
-The school finite population correction factor is: fpc_school

-The pupil weight variable is: weight_pupil
-The pupil finite population correction factor is: fpc_pupil

-The teacher interview weight variable is: weight_tchint
-The teacher interview finite population correction factor is: fpc_tchint

-The teacher roster weight variable is: weight_teacherroster
-The teacher roster finite population correction factor is: fpc_teacherroster

-The treatment status variable is: treatment where 0=control school and 1=treatment school.

NOTE that for lesson-level estimates, the school weights are applied given that lessons were not sampled within each school.

Questionnaires

Overview

Questionnaires

The enumerators administered all of the instruments using Computer Assisted Personal Interviewing (CAPI). All instruments were translated into Kiswahili and administered to all respondents in Kiswahili.

The endline survey round uses a set of survey instruments that retain most of the baseline and midline questions but with some additions to take into account changes in programme context and design and focus of programme implementation. There were two overarching changes to the suite of instruments, compared to the midline set, as follows:

- TDNA instruments dropped: These were designed to measure teachers' Kiswahili and maths subject knowledge, and were introduced at baseline because one of the original objectives of the early-grade teachers' in-service training intervention was to strengthen subject knowledge. However, the final design of EQUIP-T's in-service training chose not to focus on subject knowledge. At endline it made more sense to direct data collection efforts on instruments that are more directly relevant to the programme's interventions, and so the decision was taken to drop the TDNA.
- New small-group interview with teachers (focused on in-service training): The early-grade teacher in-service training is central to the programme's theory of change, and has absorbed a large share of the programme's spending. For this reason, it merits particular focus in the impact evaluation. Attendance at in-service training is already captured in early-grade teacher interviews, but given the high level of teacher turnover that was found at midline, getting a picture of the delivery of in-service training at a school level over the duration of the programme is useful complementary data. By gathering a small group of teachers that have attended the different types of in-service training (3Rs curriculum, Kiswahili, maths and gender-responsive pedagogy), as well as the in-service training coordinator (INCO), this instrument captures the delivery of the various residential in-service training courses, as well as the school-based training sessions.

Apart from these two changes to the group of instruments, these are the main changes that have been made to the other midline instruments:
- Parents of tested standard 3 pupils interview (score card): addition of questions on their child's pre-school attendance (including school readiness programme (SRP)); communication with
the school; awareness of the parent-teachers partnership (PTP); and corporal punishment.
- Standards 1-3 teacher interview: addition of specific questions on EQUIP-T in-service training modules completed since baseline; attendance at ward cluster reflection meetings and school performance management meetings (SPMMs); outstanding non-salary claims; removal of questions on receipt of salary.
- Standard 2 lesson observation: addition of observations related to gender-responsive pedagogy; use of maths learning materials (not textbooks); display of positive and safe learning campaign related materials.
- Head teacher interview and school records: addition of questions related to initiatives to support pupil welfare (e.g. health, hygiene, safety and child protection); initiatives to support
marginalised groups of pupils (girls, children with disabilities, pupils with learning difficulties, pupils that are vulnerable for other socio-economic reasons); new EQUIP-T interventions since midline (tablet-based SIS, business plans and income-generating activities (IGA), SPMMs, JUU clubs, pupil suggestion boxes); PTP grant spending patterns; more detail on head teacher's attendance at in-service training; outstanding non-salary claims; removal of questions on receipt of salary, missing ages of baseline pupils, and information for sampling teachers for TDNA.

The revisions to the midline instruments were trialled during a pre-test held in February 2018.

Refer to the scope section for a description of information collected under each instrument.

Data Collection

Data Collection Dates

Start	End	Cycle
2018-04-16	2018-05-21	Endline

Data Collection Mode

Computer Assisted Personal Interview [capi]

Data Collection Notes

Personnel

Oxford Policy Management's (OPM) Tanzania office conducted the endline IE survey.

The fieldwork management team comprised eight members (including six OPM staff) led by a quantitative survey project manager who had overall responsibility for the design, implementation, management and quality of the fieldwork. Since all the survey instruments were administered using computer assisted personal interviewing (CAPI), the team also included several members with strong computer programming skills in the relevant software (Surveybe). The overall project manager for the IE, who is responsible for the content of the instruments worked closely with the fieldwork team during pre-testing, training, piloting and early fieldwork. 60 enumerators were invited to the training. These were selected based on the following criteria (in order): (i) high performance during the EQUIP-T baseline and midline surveys (24 enumerators from BL and/or ML attended the EL training); (ii) interviewers with strong track record from other OPM-led surveys; and (iii) new recruits that were selected based on their prior survey experience and knowledge of education.

Fieldwork preparation

The early fieldwork preparation consisted of pre-testing the instruments and protocols and refining the instruments and protocols, obtaining permits from the government for visiting schools during the pre-tests, training and pilot and fieldwork, and revising the ML fieldwork manual.

Pre-tests of instruments

A full pre-test of all instruments and protocols took place from 19 to 23 February 2018 in Dodoma. A team of six (four members of the core evaluation team and two experienced survey supervisors who were supervisors during the midline fieldwork) visited eight schools, following one day of classroom based training. The main objectives of the pre-test were to test the functionality of the updated electronic questionnaires in the latest version of the CAPI software (Surveybe); test the changes that were made to the midline instruments, focusing mostly on the head teacher interview; and test the new endline instrument - that is the teacher group interview. The pre-test resulted in the following outcomes:
-Refinement of the instruments and data collection protocols;
-Refinement of the translation of instruments from English to Kiswahili; and
-Significant changes made to the development of the instruments in CAPI (Surveybe).

Permits and reporting

As part of preliminary preparations for any survey in Tanzania, there are two types of governmental permits that have to be obtained prior the beginning of Research work:
-COSTECH Permit - Mandatory for any research activity in Tanzania.
-Ministry Permit - Different partners in the field require Ministry letters, as few recognise COSTECH. These permits give the order to local administration to cooperate with the research and support the field teams.

Upon receipt of the permits, the anticipated fieldwork needs to be reported at the regional and district level. Letters introducing the study to local leaders are obtained in the process. For the endline survey, the COSTECH research clearance and an introduction letter were received more than three months prior to the start of actual fieldwork. For the Ministry

permits, OPM reported to The Prime Minister's Office Regional Administration and Local government (PMORALG) and to the Ministry of Education and Vocational Training (MOEVT). Reporting to MOEVT was relatively fast and simple. The initial submitted letters were followed up in person, and an introduction letter to all 12 Regional Administrative Secretaries (RAS) was received after seven days. Getting government approvals from PMORALG and the RASs was more challenging and time-consuming as it required physical reporting to PMORALG's office in Dodoma as well as physical reporting to all regions and districts that are covered by the endline fieldwork, pre-testing and piloting. However, having learned from midline how challenging this process is, the fieldwork management team devised a plan for endline that started the reporting process early on and involved two members of the fieldwork management team and two supervisors physically reporting in person to all 12 regional and 25 district offices during the month of February. This resulted in all permits and approval letters being obtained at least one month prior to piloting.

Fieldwork manual

Using the midline fieldwork manual as a basis, an extensive endline fieldworker manual was developed that covered basic guidelines on behaviour and attitude, the use of CAPI and data validation procedures, instructions on fieldwork plans and procedures (sample, targets, replacements, communication, and reporting) as well as a dedicated part on the description of all instruments and protocols. Insights from the pre-test were reflected in the manual. Draft versions of the instrument and protocol sections of the manuals were printed, handed out to interviewers as a reference during the training, and used as guidelines by the trainers. The manual was updated on an ongoing basis during the training and pilot phase where updated conventions or additional clarifications were needed. The final version of the manual was printed at the end of the pilot phase and copies provided to the field teams.

Training and pilot

Enumerator training and a field pilot took place in Dar es Salaam and Dodoma from 26 March to 14 April 2018. A total of 60 enumerator trainees participated in the training. The training was delivered by four members of the fieldwork management team, the overall IE project manager and another member of the core evaluation team. The main objective of the training was to ensure that team members would be able to master the instruments, understand and correctly implement the fieldwork protocols, comfortably use CAPI, and be able to perform data validation. Supervisors were furthermore trained on their extra responsibilities of data management, fieldwork and financial management, logistical tasks, and the transmission of data files to the data manager.

The training had two components: a classroom-based training component and a field-based component that included a full scale pilot. The performance of enumerators was assessed on an on-going basis, using written assessments and observation of performance in the field and these scores were recorded. At the end of the training and pilot phase, the final fieldwork team was selected using this information.

A higher number of data collectors than needed for data collection were invited to and attended the training. This allowed for a selection of the best suited candidates at the end of the training and provided a pool of reserve additional trained staff that could be called upon in case of enumerator attrition during data collection.

Fieldwork organisation

The fieldwork plan was designed to cover all 200 schools within all 12 regions and 25 districts for the duration of not more than six weeks. The plan had to cater for the short fieldwork time window dictated by the end of the school mid-term break and the start of exams at the end of the term; rainy season; allowing the fieldwork management team to supervise teams during the first week of implementation; minimising travel days between districts and during the weekdays; suitable allocation of teams to districts to address cultural and language barriers; and flexibility to deal with unforeseen circumstances.

The team composition and fieldwork model at endline were the same as those at midline with the exception of adding one more field team to deal with the shorter timeframe at endline and to ensure that the fieldwork is completed within five to six weeks. At endline there were four treatment teams composed of five enumerators and one supervisor, four control teams of four enumerators and one supervisor each, and one team of five enumerators and one supervisor that visited control and treatment areas. Each team visited and completed one school on one day.

The fieldwork started on 16 April and ended on 21 May 2018 with no major breaks in-between.

Questionnaires

Questionnaires

The enumerators administered all of the instruments using Computer Assisted Personal Interviewing (CAPI). All instruments were translated into Kiswahili and administered to all respondents in Kiswahili.

The endline survey round uses a set of survey instruments that retain most of the baseline and midline questions but with some additions to take into account changes in programme context and design and focus of programme implementation. There were two overarching changes to the suite of instruments, compared to the midline set, as follows:

- TDNA instruments dropped: These were designed to measure teachers' Kiswahili and maths subject knowledge, and were introduced at baseline because one of the original objectives of the early-grade teachers' in-service training intervention was to strengthen subject knowledge. However, the final design of EQUIP-T's in-service training chose not to focus on subject knowledge. At endline it made more sense to direct data collection efforts on instruments that are more directly relevant to the programme's interventions, and so the decision was taken to drop the TDNA.
- New small-group interview with teachers (focused on in-service training): The early-grade teacher in-service training is central to the programme's theory of change, and has absorbed a large share of the programme's spending. For this reason, it merits particular focus in the impact evaluation. Attendance at in-service training is already captured in early-grade teacher interviews, but given the high level of teacher turnover that was found at midline, getting a picture of the delivery of in-service training at a school level over the duration of the programme is useful complementary data. By gathering a small group of teachers that have attended the different types of in-service training (3Rs curriculum, Kiswahili, maths and gender-responsive pedagogy), as well as the in-service training coordinator (INCO), this instrument captures the delivery of the various residential in-service training courses, as well as the school-based training sessions.

Apart from these two changes to the group of instruments, these are the main changes that have been made to the other midline instruments:
- Parents of tested standard 3 pupils interview (score card): addition of questions on their child's pre-school attendance (including school readiness programme (SRP)); communication with
the school; awareness of the parent-teachers partnership (PTP); and corporal punishment.
- Standards 1-3 teacher interview: addition of specific questions on EQUIP-T in-service training modules completed since baseline; attendance at ward cluster reflection meetings and school performance management meetings (SPMMs); outstanding non-salary claims; removal of questions on receipt of salary.
- Standard 2 lesson observation: addition of observations related to gender-responsive pedagogy; use of maths learning materials (not textbooks); display of positive and safe learning campaign related materials.
- Head teacher interview and school records: addition of questions related to initiatives to support pupil welfare (e.g. health, hygiene, safety and child protection); initiatives to support
marginalised groups of pupils (girls, children with disabilities, pupils with learning difficulties, pupils that are vulnerable for other socio-economic reasons); new EQUIP-T interventions since midline (tablet-based SIS, business plans and income-generating activities (IGA), SPMMs, JUU clubs, pupil suggestion boxes); PTP grant spending patterns; more detail on head teacher's attendance at in-service training; outstanding non-salary claims; removal of questions on receipt of salary, missing ages of baseline pupils, and information for sampling teachers for TDNA.

The revisions to the midline instruments were trialled during a pre-test held in February 2018.

Refer to the scope section for a description of information collected under each instrument.

Data Collectors

Name	Abbreviation	Affiliation
Oxford Policy Management Ltd	OPM	

Supervision

Quality control and data checking protocols

At the end of each working day, supervisors collected all interview files from their team members and uploaded them into a shared and organised Dropbox folder that was set up by the data manager. The data manager would receive all files from all nine teams and export them into Stata data files (a statistical programme) and then run daily checks on all files to make sure they are complete and identify potential errors. Several mechanisms were put in place in order to ensure high quality of

the data collected during the survey. These are briefly summarised in turn below.

-Selection and supervision of enumerators
As discussed above, each enumerator was supervised at least once by the training team during the training, piloting and first week of data collection. This allowed a well-informed selection of enumerators and their allocation into roles matching individual strengths and weaknesses.

-CAPI built-in routing and validations
One important quality control means in CAPI surveys are the use of automatic routing and checking rules built into the CAPI questionnaires that flag simple errors during the interview, i.e. early enough for them be corrected during the interview. In each CAPI instrument, validations and checks were incorporated in the design in order to significantly reduce errors and inaccuracies during data collection. In addition to having automatic skip patterns built into the design to eliminate errors resulting from wrong skips, the CAPI validations also checked for missing fields, out of range values and simple inconsistencies within instruments.

-Secondary consistency checks and cleaning in Stata
The EL survey exploited another key advantage of CAPI surveys, the immediate availability of data, by running a range of secondary consistencies checks across all data on a daily basis in Stata. Data received from the field was exported to Stata the following day, and a range of do-files were run to assess consistency and completeness, and make corrections if necessary. The checks comprised the following: ID uniqueness and matching across instruments; completeness of observations: target sample size versus actual; and intra- and inter-instrument consistency and out of range checks. The data manager ran the checking do-file on a daily basis on the latest cleaned data. This would return a list of potential issues which the data manager would then investigate and undertake the necessary cleaning actions. Whenever any issue was flagged, effort to obtain an explanation was undertaken either by reviewing enumerator comments or phoning teams. On a daily basis, the data manager collated, shared and discussed all flagged errors with the supervisors in the field, who in turn discussed them with their team members. Throughout the fieldwork, occurrences of errors were monitored in order to keep an eye on the performance of data collectors and constantly provide them with feedback to improve.

-Monitoring fieldwork progress and performance indicators
In addition to the above checks that were specific to each instrument, the survey team built a dashboard that allowed for daily monitoring of the general progress of the fieldwork and specific indicators revealing the performance of teams and enumerators over time. For example, indicators included number of control/treatment schools completed, number of instruments completed within each school, average interviewing time of each instrument, time of the day when the pupil tests were conducted, number of pupils interviewed for the scorecard instead of their parents, number of teacher interviews conducted over the phone, number of pupils being replaced, etc. These indicators were constructed in a Stata do-file that ran on the latest cleaned dataset and was then uploaded onto the dashboard (that was created using the visual software, Power BI) that would break down each of the indicators by team, enumerator (where applicable) and week of data collection. This was reviewed on a daily basis by the fieldwork management team and used to feedback to weaker teams and to improve performance.

-Back-checking data
The quality assurance protocol involved visits by the fieldwork management team to the field as well as data back-checks. Two members of the fieldwork management team visited a number of schools and households across 8 of the 12 regions over a two-week period. The purpose was to verify that the school and household interviews were conducted properly, to collect any missing information from these schools or clarify certain issues that were flagged as errors during the daily checking process, and to hold debriefs and retraining sessions with the teams in the field.

-Integration of Analysis and Survey Team
Another central element of QA was the strong integration of the fieldwork management team and the members of the quantitative analysis team, including the overall IE project manager. Members of both teams were involved in the fieldwork preparation and implementation, and in the analysis process which followed.

Data Processing

Data Editing

Given the data was electronically collected, it was continually checked, edited and processed throughout the survey cycle.

A first stage of data checking was done by the survey team which involved (i) checking of all IDs; (ii) checking for missing observations; (iii) checking for missing item responses where none should be missing; and (iv) first round of checks for inadmissible/out of range and inconsistent values. See section 'Supervision' for more details. Additional data processing activities were performed at the end of data collection in order to transform the collected cleaned data into a format that is ready for analysis. The aim of these activities was to produce reliable, consistent and fully-documented datasets that can be analysed throughout the survey and archived at the end in such a way that they can be used by other data users well into the future. Data processing activities involved:

- Computing and merging in the sampling weights,
- Reshaping datasets in order to produce data files for each unit of observation,
- Anonymising data by removing all variables that identify respondents such as names, address, GPS coordinates, etc.,
- Classifying non-response and coding them using a pre-determined classification scheme,
- Reviewing 'Other (specify)' responses by checking if any of the responses fall into existing response categories and can be recoded into the existing category or if there are multiple similar other responses that warrant the creation of a new response category (a decision to be made by the data analysts), and
- Properly naming and labelling the variables in each dataset.

The datasets were then sent to the analysis team where they were subjected to a second set of checking and cleaning activities. This included checking for out of range responses and inadmissible values not captured by the filters built into the CAPI software or the initial data checking process by the survey team.

A comprehensive data checking and analysis system was created including a logical folder structure, the development of a detailed data analysis guide and template syntax files (in Stata), to ensure data checking and cleaning activities were recorded, that all analysts used the same file and variable naming conventions, variable definitions, disaggregation variables and weighted estimates appropriately.

Data Appraisal

No content available

Documentation

Questionnaires

EQUIP-T IE Head Teacher (HT) Endline Questionnaire

Title	EQUIP-T IE Head Teacher (HT) Endline Questionnaire
Author(s)	Oxford Policy Management Ltd.
Date	2018-04-30
Country	Tanzania
Language	English
Filename	EQUIP-T IE head teacher interview (HT) endline questionnaire.pdf

EQUIP-T IE Headcount (HC) Endline Questionnaire

Title	EQUIP-T IE Headcount (HC) Endline Questionnaire
Author(s)	Oxford Policy Management Ltd.
Date	2018-04-30
Country	Tanzania
Language	English
Filename	EQUIP-T IE headcount (HC) endline questionnaire.pdf

EQUIP-T IE School Records (SR) Endline Questionnaire

Title	EQUIP-T IE School Records (SR) Endline Questionnaire
Author(s)	Oxford Policy Management Ltd.
Date	2018-04-30
Country	Tanzania
Language	English
Filename	EQUIP-T IE school records (SR) endline questionnaire.pdf

EQUIP-T IE Lesson Observation (LO) Endline Questionnaire

Title	EQUIP-T IE Lesson Observation (LO) Endline Questionnaire
Author(s)	Oxford Policy Management Ltd.
Date	2018-04-30
Country	Tanzania
Language	English
Filename	EQUIP-T IE lesson observation (LO) endline questionnaire.pdf

EQUIP-T IE Teacher Interview (TI) Endline Questionnaire

Title	EQUIP-T IE Teacher Interview (TI) Endline Questionnaire
Author(s)	Oxford Policy Management Ltd.
Date	2018-04-30
Country	Tanzania
Language	English

Filename	EQUIP-T IE teacher interview (TI) endline questionnaire.pdf

EQUIP-T IE Pupil Background and Learning Assessment (PB) Endline Questionnaire

Title	EQUIP-T IE Pupil Background and Learning Assessment (PB) Endline Questionnaire
Author(s)	Oxford Policy Management Ltd.
Date	2018-04-30
Country	Tanzania
Language	English
Filename	EQUIP-T IE pupil background and learning assessment (PB) endline questionnaire.pdf

EQUIP-T IE Poverty Scorecard (SC) Endline Questionnaire

Title	EQUIP-T IE Poverty Scorecard (SC) Endline Questionnaire
Author(s)	Adapted by Oxford Policy Management Ltd. for the EQUIP-T IE baseline survey from Schreiner, M. (2013) A Simple Poverty Scorecard for Tanzania. Kansas City: Microfinance Risk Management, LLC.
Date	2018-04-30
Country	Tanzania
Language	English
Filename	EQUIP-T IE poverty scorecard (SC) endline questionnaire.pdf

EQUIP-T IE In-service Training Coordinator Interview (INCO) Endline Questionnaire

Title	EQUIP-T IE In-service Training Coordinator Interview (INCO) Endline Questionnaire
Author(s)	Oxford Policy Management Ltd.
Date	2018-04-30
Country	Tanzania
Language	English
Filename	EQUIP-T IE in-service training coordinator interview (INCO) endline questionnaire.pdf

Reports

EQUIP-Tanzania Impact Evaluation. Final Baseline Technical Report, Volume I: Results and Discussion

Title	EQUIP-Tanzania Impact Evaluation. Final Baseline Technical Report, Volume I: Results and Discussion
Author(s)	Pettersson, Gunilla Rawle, Georgina Outhred, Rachel Brockerhoff, Stephanie Wills, Gabrielle Nugroho, Dita Jasper, Paul Kveder, Andrej Beavis, Adrian
Date	2015-01-15
Country	Tanzania
Language	English
Publisher(s)	Oxford Policy Management Ltd.
Filename	EQUIP-T IE Final Baseline Technical Report, Volume I Results and Discussion.pdf

EQUIP-Tanzania Impact Evaluation. Final Baseline Technical Report,

Volume II: Methods and Technical Annexes

Title	EQUIP-Tanzania Impact Evaluation. Final Baseline Technical Report, Volume II: Methods and Technical Annexes
Author(s)	Pettersson, Gunilla Rawle, Georgina Outhred, Rachel Brockerhoff, Stephanie Wills, Gabrielle Nugroho, Dita Jasper, Paul Kveder, Andrej Beavis, Adrian
Date	2015-01-15
Country	Tanzania
Language	English
Publisher(s)	Oxford Policy Management Ltd.
Filename	EQUIP-T IE Final Baseline Technical Report, Volume II Methods and Technical Annexes.pdf

EQUIP-Tanzania Impact Evaluation. Midline Technical Report, Volume I: Results and Discussion

Title	EQUIP-Tanzania Impact Evaluation. Midline Technical Report, Volume I: Results and Discussion
Author(s)	Rawle, Georgina Ruddle, Nicola Pettersson Gelander, Gunilla Wallin, Johanna Binci, Michele Jasper, Paul Harb, Jana Hebbar, Madhumitha Davis, Jean Aldinucci, Alice
Date	2017-03-21
Country	Tanzania
Language	English
Publisher(s)	Oxford Policy Management Ltd.
Filename	EQUIP-T IE Midline Technical Report, Volume I Results and Discussion.pdf

EQUIP-Tanzania Impact Evaluation. Midline Technical Report, Volume II: Methods and Supplementary Evidence

Title	EQUIP-Tanzania Impact Evaluation. Midline Technical Report, Volume II: Methods and Supplementary Evidence
Author(s)	Rawle, Georgina Ruddle, Nicola Pettersson Gelander, Gunilla Wallin, Johanna Binci, Michele Jasper, Paul Harb, Jana Hebbar, Madhumitha Davis, Jean Aldinucci, Alice
Date	2017-03-21
Country	Tanzania
Language	English
Publisher(s)	Oxford Policy Management Ltd.
Filename	EQUIP-T IE Midline Technical Report, Volume II Methods and Supplementary Evidence.pdf

EQUIP-Tanzania Impact Evaluation. Midline Issues Note 1: The Changing Context for Teacher In-Service Training—Reflections on EQUIP-Tanzania's Experience

Title	EQUIP-Tanzania Impact Evaluation. Midline Issues Note 1: The Changing Context for Teacher In-Service Training—Reflections on EQUIP-Tanzania's Experience
Author(s)	Pettersson Gelander, Gunilla Rawle, Georgina Karki, Shrochis Ruddle, Nicola
Date	2017-03-09
Country	Tanzania
Language	English
Publisher(s)	Oxford Policy Management Ltd.
Filename	EQUIP-T IE Midline Issues Note 1.zip

EQUIP-Tanzania Impact Evaluation. Endline Quantitative Technical

Report, Volume I: Results and Discussion

Title	EQUIP-Tanzania Impact Evaluation. Endline Quantitative Technical Report, Volume I: Results and Discussion
Author(s)	Rawle, Georgina Binci, Michele Pettersson Gelander, Gunilla Harb, Jana Jasper, Paul Khan, Safa Medardi, Deo Romarri, Alessio Rorich, Michelle Ruddle, Nicola
Date	2019-06-04
Country	Tanzania
Language	English
Publisher(s)	Oxford Policy Management Ltd.
Filename	EQUIP-T IE Endline Quantitative Technical Report, Volume I Results and Discussion.pdf

EQUIP-Tanzania Impact Evaluation. Endline Quantitative Technical Report, Volume II: Methods and Supplementary Evidence

Title	EQUIP-Tanzania Impact Evaluation. Endline Quantitative Technical Report, Volume II: Methods and Supplementary Evidence
Author(s)	Rawle, Georgina Binci, Michele Pettersson Gelander, Gunilla Harb, Jana Jasper, Paul Khan, Safa Medardi, Deo Romarri, Alessio Rorich, Michelle Ruddle, Nicola
Date	2019-01-28
Country	Tanzania
Language	English
Publisher(s)	Oxford Policy Management Ltd.
Filename	EQUIP-T IE Endline Quantitative Technical Report, Volume II Methods and Supplementary Evidence.pdf

EQUIP-Tanzania Impact Evaluation. Endline Quantitative Technical Report, Volume II: Annex G

Title	EQUIP-Tanzania Impact Evaluation. Endline Quantitative Technical Report, Volume II: Annex G
Author(s)	Rawle, Georgina Binci, Michele Pettersson Gelander, Gunilla Harb, Jana Jasper, Paul Khan, Safa Medardi, Deo Romarri, Alessio Rorich, Michelle Ruddle, Nicola
Date	2019-01-28
Country	Tanzania
Language	English
Publisher(s)	Oxford Policy Management Ltd.
Filename	EQUIP-T IE Endline Quantitative Technical Report, Volume II_Annex G.xlsx

Technical documents

EQUIP-T IE INCO Endline Questionnaire MANUAL

Title	EQUIP-T IE INCO Endline Questionnaire MANUAL
Author(s)	Oxford Policy Management Ltd.
Date	2018-04-16
Language	English
Description	This document is an excerpt of the survey manual and includes a detailed description on how the in-service training coordinator (INCO) questionnaire should be administered at endline by the enumerators. It should be read in conjuction with the INCO questionnaire, uploaded under 'Questionnaires'.
Filename	EQUIP-T IE INCO endline questionnaire MANUAL.pdf

Additional Information on Constructed Indicators

Title	Additional Information on Constructed Indicators
Author(s)	Oxford Policy Management Ltd.
Date	2019-07-01
Language	English
Description	This note includes additional information on a constructed indicator in the school-level public dataset that identifies the respondent to the head teacher (HT) instrument.
Filename	EQUIP-T IE additional information on constructed indicators.pdf

Tanzania Poverty Scorecard Instructions (Schreiner 2016)

Title	Tanzania Poverty Scorecard Instructions (Schreiner 2016)
Author(s)	Schreiner, Mark
Date	2016-06-27
Country	Tanzania
Language	English
Filename	Tanzania poverty scorecard instructions (Schreiner 2016).pdf

Other materials

Map of EQUIP-T IE Districts

Title	Map of EQUIP-T IE Districts
Author(s)	Oxford Policy Management Ltd.
Date	2019-07-01
Country	Tanzania
Language	English
Filename	Map of EQUIP-T IE districts.pdf