

## **BASELINE REPORT**

# **Impact Assessment of IDP Foundation, Inc. and Sesame Workshop's Techniques for Effective Teaching Program in Ghana**

NOVEMBER 20, 2018

Prepared by NORC at the University of Chicago for the IDP Foundation, Inc.

## Table of Contents

<b>1. Executive Summary</b> .....	<b>2</b>
1.1 Project Background .....	2
1.2 Evaluation Objectives .....	2
1.3 Findings and Conclusions .....	3
<b>2. Project Background</b> .....	<b>6</b>
<b>3. Evaluation Purpose &amp; Evaluation Questions</b> .....	<b>8</b>
3.1 Evaluation Purpose .....	8
3.2 Evaluation Questions .....	8
<b>4. Evaluation Design &amp; Data Collection</b> .....	<b>10</b>
4.1 Evaluation Design Overview .....	10
4.2 School and Learner Sample .....	13
4.3 Baseline Data Collection .....	14
4.4 Midline and Endline Data Collection .....	19
<b>5. Baseline Results</b> .....	<b>21</b>
5.1 Basic demographic characteristics of learners and teachers .....	21
5.2 Attitudes towards Teaching and Learning .....	22
5.3 Discipline Practices .....	28
5.4 Gender in the Classroom .....	30
5.5 Teacher Motivation and Satisfaction .....	31
5.6 Teacher and Child Attendance .....	37
5.7 KG2 Learner Assessments .....	38
5.8 P2 Learner Assessments .....	41
5.9 School and Home Environment .....	48
<b>6. Final Comments</b> .....	<b>52</b>

# 1. Executive Summary

## 1.1 Project Background

---

Despite reform attempts, the dominant pedagogy in Ghana remains teacher-centered. This ‘chalk-and-talk’ instruction places learners in a passive role and limits their activity to memorizing facts and reciting them back to the teacher. In low-fee private schools, the vast majority of teachers are also under-trained and have very limited opportunities for professional development. The IDP Foundation, Inc. (IDPF) has partnered with Sesame Workshop to develop the Techniques for Effective Teaching (TFET) program to address these challenges through strengthening the capacity of low-fee private school kindergarten and primary school teachers to prepare for, conduct, and assess pupil-centered learning.

The initiative targets nurseries, kindergartens and primary schools participating in the IDP Rising Schools Program and began in 2012 with a two-day training for proprietors and teachers from 92 schools. The program seeks to equip teachers with the child-centered skills necessary to make teaching and learning fun. Sesame Workshop provides in-service, in-person teacher training alongside a kit comprised of teacher training video modules, interactive training materials, and durable classroom learning materials. The program follows a train-the-trainer model, with those teachers expected to train their colleagues. Based on implementation experience, and internal and external research, the program has been adapted and expanded for a second phase of piloting. In 2018, around 50 schools participating in the IDP Schools Rising project in the Greater Accra and Central regions were invited to attend two-week long training sessions – the first in February and the second in August.

## 1.2 Evaluation Objectives

---

NORC at the University of Chicago, has been charged by IDPF with conducting an external evaluation of the TFET program. The primary interest is to assess the degree to which teachers are able to effectively adopt and implement the key TFET pedagogical principles and whether this translates to better development and learning outcomes for their pupils in Kindergarten 2 and Primary 2 classes.

Specifically, the study aims to assess the impact of the TFET program on a range of outcomes, including, among others:

- Quality of classroom instruction
- Promotion of child-friendly classrooms

- Positive discipline in the classroom
- Pupil-centered learning
- Gender equality in the classroom
- Teacher motivation and job satisfaction
- Emergent and early literacy; emergent and early numeracy
- Children's socio-emotional development
- Children's motivation to attend school

In addressing these research objectives, the NORC team is using a combination of quantitative, qualitative and secondary research methods. The study is collecting quantitative data at baseline and endline through learner development and learning assessments, learner interviews, teacher and head teacher interviews, and classroom observations. In addition to quantitative data collection, NORC is also reviewing the TFET program training materials, observing master and step-down training sessions, conducting classroom observations, and facilitating focus group discussions. These qualitative assessments will help provide a more holistic picture of the TFET program impact in addition to providing Sesame Workshop with information to feed into ongoing adaptations of the program implementation model.

### 1.3 Findings and Conclusions

---

In this report we present findings from the baseline data collection activity that took place in February and March 2018. These baseline findings will also contribute to the broader literature on teacher attitudes and behavior and student learning outcomes in Ghana.

A number of facts stem from our analysis of the baseline data:

- Teachers tend to respond positively to statements that describe learner-centered pedagogical approaches (and negatively to statements that are antithetical to learner-centered teaching) although there is considerable variation in the percentage of teachers strongly agreeing across statements. For example, only 26% of teachers strongly agree with the statement that “pupils should be allowed to think of solutions to practical problems by themselves before the teacher shows the solution,” and only 19% strongly disagree that “teachers are not supposed to play with children during the school day.” This is in contrast to the 52% of teachers who strongly

agree that “it is important that teachers help pupils to learn to control their behavior to make them responsible.”

- An aggregated index of learner-centered beliefs shows considerable variation across teachers. This enables us to identify more and less learner-centered teachers, which will be important in exploring the extent to which prior beliefs hinder or support the uptake of TFET methods.
- While P2 learners tend to agree with statements that describe their teachers as learner-centered, over half (53%) of our P2 learners indicate that their “teacher does not think that is okay to make mistakes.” This, together with the teachers’ responses, suggests that classroom practices do not tend to encourage children to explore, experiment and solve problems independently.
- The use of emotional and physical punishment seems common among KG2 and P2 teachers, and learners report fear of going to school because of discipline practices. Caning in particular is reported by a majority of teachers as an effective way to discipline learners.
- Learner absenteeism is an important issue that gets compounded by teacher’s absenteeism. Almost 50% of P2 pupils interviewed reported missing at least one day of school during the prior week, and attendance records indicate an average of 31 and 35 percent absenteeism in P2 and KG2 in the previous school term, respectively. Records indicate that KG2 and P2 teachers are absent on 19% and 10% of teaching days respectively.
- Strong gender inequitable attitudes prevail amongst both teachers and learners. For example, only 22% of teachers strongly disagree with the statement that “boys are naturally better than girls at math and science.” The P2 learners’ views show even stronger evidence of gender bias, particularly amongst boys. When we combine the very true and true categories, we find that 83% of boys and 60% of girls agree with the statement “it is more important for boys to do well in school than it is for girls.” Results are similar for the statement “boys are usually more intelligent than girls.”
- The literacy and numeracy assessments of P2 learners, largely mirror findings in the 2013 and 2015 national EGRA and EGMA assessments. The vast majority of learners are unable to correctly identify at least one word in the dominant local language. The average oral reading fluency in English is 33 words per minute which is higher than the national average, but still substantially below the national benchmark of 45 words per minute. It is the comprehension subtasks (both listening and reading comprehension) where results are most concerning with P2 learners showing very low levels of understanding. The analyses of the math assessments show that, in general, learners perform better in numeracy subtasks that are procedural and

repetitive than in conceptual ones. P2 learners tend to solve problems and operations using inefficient strategies, such as tick marks or their fingers, which preclude them from solving addition or subtraction tasks that are not very simple.

- Similar to the P2 learners, the KG2 learners show greater difficulty with the more conceptual and spatial activities such as matching and sorting and putting a simple puzzle together than the more procedural or repetitive subtasks such as number identification.

This baseline report has provided a rich picture of a challenging context with multiple opportunities for the program to potentially impact practices and outcomes. Learner outcomes are poor, particularly for tasks with higher cognitive demand such as problem-solving and comprehension. Learner and teacher responses suggest that learners are not encouraged to explore, experiment and solve problems independently. Absenteeism is high amongst learners and this is compounded by teachers missing on average one to two of every ten teaching days. Corporal punishment is widely considered effective and the majority of learners report that children are sometimes afraid to attend school for fear of punishment. Both learners and teachers exhibit substantial gender bias.

## 2. Project Background

IDPF has partnered with and funded Sesame Workshop to create the Techniques for Effective Teaching Program (TFET). Operating in Ghana, TFET aims to improve early childhood education and development in low-fee private schools by focusing on child-centered learning and teacher capacity to provide higher quality learning environments. In line with other developing countries, changes to the basic education curricula in Ghana involved a shift from teacher-centered to learner-centered approaches with the 1987 educational reforms promoting the adoption of hands-on activities and learner-centered approaches.<sup>1</sup> Despite these reform attempts, curriculum delivery in Ghana “remains entrenched in a rote learning style, which is neither child-centered nor activity based.”<sup>2</sup> In low-fee private schools, teachers tend to be under-qualified and under-trained and there are very limited professional development opportunities. Teachers in these schools also have limited access to resources. In this context, the TFET project conducts a teacher training program designed to strengthen the capacity of low-fee private school kindergarten and early primary school teachers to prepare for, conduct, and assess pupil-centered learning. The initiative targets kindergartens and primary schools participating in the IDP Rising Schools Program.

In 2004, Ghana adopted the Early Childhood Care and Development (ECCD) policy, which sets out a framework to guide various stakeholders in the effective implementation of ECCD with the aim of promoting the growth and development of all children under eight years of age.<sup>3</sup> The TFET program fits squarely into the mandate of development partners and NGOs by developing appropriate curricula, building capacity and providing training. In addition, the program also supports private proprietors in their responsibility to deliver quality ECCD services.

The TFET project is in its second phase as of 2018<sup>4</sup>, inviting a cohort of teachers from 40-50 schools participating in the IDP Rising Schools Program to attend two week-long master trainings—the first in February 2018 and the second scheduled for July/August 2018. At the core of the training materials for the two weeks are 14 video modules developed by IDPF and Sesame Workshop. Each of the 14 videos covers a teaching topic and demonstrates what it looks like in action inside a classroom. Each

<sup>1</sup> Nketsia, W., Saloviita, T., Gyimah, E.K., 2016. Teacher Educators' Views on Inclusive Education and Teacher Preparation in Ghana. *International Journal of Whole Schooling*. <https://files.eric.ed.gov/fulltext/EJ1118090.pdf>

<sup>2</sup> Bidwell, K., Parry, K., Watine, L., 2014. Exploring Early Education Programs in Peri-Urban Settings in Africa: Accra Report. [https://www.poverty-action.org/sites/default/files/publications/accra\\_ecd\\_report\\_full.pdf](https://www.poverty-action.org/sites/default/files/publications/accra_ecd_report_full.pdf)

<sup>3</sup> International Labour Organization, 2004. Early Childhood Care and Development Policy. [http://www.ilo.org/dyn/natlex/natlex4.detail?p\\_lang=en&p\\_isn=88527&p\\_country=GHA&p\\_count=116](http://www.ilo.org/dyn/natlex/natlex4.detail?p_lang=en&p_isn=88527&p_country=GHA&p_count=116)

<sup>4</sup> The first phase of the TFET project was implemented in 2012 with 92 schools that were part of the IDP Rising Schools program invited to participate in the trainings. Proprietors and one teacher from each school were invited to attend a two-day training session. They were then expected to pass on the training to their colleagues. Recommendations from both internal and external evaluations of this first phase have informed the adaptation and expansion of the TFET program for the second phase pilot.

video starts with key messages that frame the topic, pauses with questions for reflection, and ends with the key messages. The videos are hosted by Sesame friends Kami and Zobi, and a teacher, Ms. Efia, providing context and entertainment.

The training program, through the video modules and the paired teacher manual, sets out to orient early childhood teachers in the following areas of teaching and learning:

- Classroom and instructional management
- Developing an inclusive, pupil-centered pedagogy that provides lots of fun and opportunities for creativity
- Using and creating teaching and learning materials
- Literacy
- Mathematics

The 14 training modules are named: (1) The fundamentals of teaching, (2) Developing a child-friendly classroom, (3) Positive discipline and classroom management, (4) Developing and teaching with creative, low-cost resources, (5) Pupil-centered learning in English literacy, (6) Pupil-centered learning in mathematics, (7) Time management, (8) Practical ways to make the classroom creative and fun, (9) Differentiated instruction, (10) Evaluation and assessment, (11) Girl's education and inclusion in the classroom, (12) Ensuring an inclusive learning environment, (13) Early education and child development, (14) Learning through play with early learners.

Once the teachers complete the TFET master training week, they are expected to return to their schools and organize a step-down training where they lead other early grade teachers from their school in reviewing the same video modules and other training material.

## 3. Evaluation Purpose & Evaluation Questions

### 3.1 Evaluation Purpose

NORC at the University of Chicago is conducting an evaluation of the TFET program to assess its impact on a range of outcomes, including, among others:

- Emergent and early literacy; emergent and early numeracy
- Children's socio-emotional development
- Children's motivation to attend school
- Teacher motivation and job satisfaction
- Quality of classroom instruction
- Promotion of child-friendly classrooms
- Positive discipline in the classroom
- Pupil-centered learning
- Gender equality in the classroom

In addition to the quantitative data collection for the impact evaluation, a host of qualitative data collection methods will be employed by education experts to provide more in-depth understanding about strengths and challenges of the training materials, training delivery and integrating the training into classroom practice. These qualitative data will provide valuable feedback to the TFET team as the program implementation evolves.

### 3.2 Evaluation Questions

The evaluation research questions for this study focus on broad impacts to teachers and learners as well as on improvements to and recommendations for the training program design and implementation. Specifically, the research questions are:

1. Phase 1 Training versus Phase 2 Training
  - ▶ Which recommendations from the Phase One assessment were effectively incorporated into the second phase, and what were their effects?
  - ▶ How does the teacher experience of the training in phase two compare to phase one (length of training time, video content, teacher trainers instruction & delivery, feedback provided to teachers based on classroom monitoring)?
2. Impact on Teachers

- ▶ What impact does the training program have on teaching techniques in the classrooms?
- ▶ What impact does the training program have on developing greater professional pride?

### 3. Training Delivery Effectiveness

- ▶ How effective were the trainers and the step-down training model at delivering the training material?
- ▶ How effective were the four new videos, their content, and the program overall at presenting realistic material that teachers can use in the classroom?

### 4. Impact on Children

- ▶ What is the impact of the training on learner's socio-emotional, reading, and mathematics learning outcomes?

## 4. Evaluation Design & Data Collection

The purpose of this report is to present the findings from the baseline data collection, which is part of a larger quasi-experimental<sup>5</sup> impact evaluation designed to assess impacts of the TFET program on teachers and their learners. First, this section will present the study design as envisioned from the start of the project; second, we present the target sample for schools and learners, and its rationale; third, we present details on the baseline data collection preparation and field work process; and fourth, we outline the upcoming research activities for the midline and endline phases of the study.

### 4.1 Evaluation Design Overview

NORC is combining the use of quantitative, qualitative and secondary research methods. The study is collecting data at baseline and endline through learner development and learning assessments, learner interviews, teacher and head teacher interviews, and classroom observations. Directed at providing implementers with information to feed into ongoing adaptations of the TFET implementation model, NORC reviewed TFET program training materials (prior to baseline) and observed training sessions (at time of baseline) and will also conduct classroom observations and focus group discussions with teachers at midline data collection.

### Impact Assessment Design

An impact evaluation serves to rigorously assess the causal effect of a specific intervention on a set of outcomes. The goal is to isolate the effect that an intervention may have from any other coexisting factors that might modify outcomes of interest as well, and in that way differentiate causality from mere correlation. The approach therefore allows for the attribution of changes in relevant outcome(s) to a specific intervention by answering the counterfactual question “What would have happened to activity participants in the absence of the intervention?” Ideally, this is done by observing the same program participants both with and without the intervention at the same point in time. Of course, this is not possible; at any given time, a participant either receives the intervention or not. Therefore, we can never directly observe the counterfactual and instead need to create a comparison group to serve as the counterfactual.

Identifying a credible comparison group is a critical aspect of an impact evaluation and there are several approaches to do so. Given that all 51 schools that matched the selection criteria in the Greater Accra and Central Region were selected for the training cohort, this ruled out the possibility of a

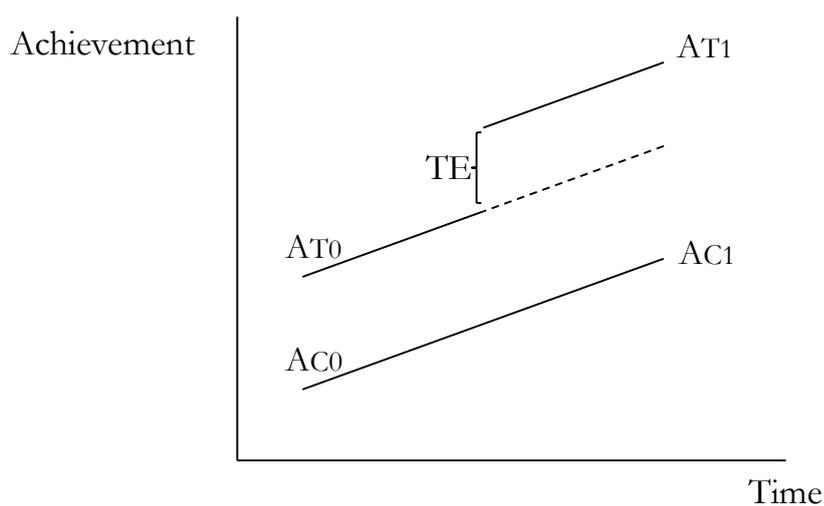
<sup>5</sup> In a quasi-experimental design, schools are not randomly assigned to the treatment or comparison group and statistical methods are used to create an appropriate counter-factual.

design using random assignment of treatment and comparison schools. Instead, our impact evaluation is based on quasi-experimental methods where a comparison group is formed by statistical methods, rather than by random assignment. We first used matching techniques to form a comparison group of schools from the surrounding area that best match the schools receiving treatment. The details used to construct the comparison group of schools are described in detail in Annex I.

Once the comparison group of schools is defined, our empirical strategy relies on the use both propensity score matching (PSM) and Difference-in-Differences (DiD) approaches, as this further improves our ability to control for selection bias. The possibility of selection bias arises because the assignment to treatment was not random and differences in outcomes between the treatment and comparison groups may depend on characteristics related to the probability of being in the treatment group rather than due to the effect of the treatment itself. Propensity score matching attempts to minimize selection bias by creating a comparison groups that is comparable to the treatment group on all observed covariates.

In the (DiD) analysis specifically, we will compare changes in outcomes of interest – early literacy, early numeracy and social-emotional development - between baseline and endline in treatment schools to changes between baseline and endline in those same outcomes in comparison schools. A graphical representation of the proposed methodology is depicted in Figure 1.

**Figure 1.** Difference-in-difference estimator



Where:

$AT_0$  is the average test score for a given grade at baseline in the treatment group

$AC_0$  is the average test score for a given grade at baseline in the control group

$AT_1$  is the average test score for a given grade at endline in the treatment group

$AC_1$  is the average test score for a given grade at endline in the control group

TE is the treatment effect for the corresponding grade

Clearly both groups of schools do not need to be identical at baseline, given that the comparison relies on the relative changes and not the levels. The DiD approach assumes that in absence of the treatment the two groups of schools/learners in the schools would have evolved in the same way (parallel trends in the graph, dotted line).

A simple representation of the measure of impact under this quasi-experimental design is the interaction effect of treatment and time, or the double-difference estimate which is:

$$\text{Estimate of impact} = (YT_1 - YT_0) - (YC_1 - YC_0)$$

where, Y = outcome of interest; T = treatment group; C = comparison group; 0 = baseline or beginning of study; and 1 = end of study. This equation can be estimated using a simple regression approach.

Due to the timing of the first TFET Master training, our approach will vary across research questions. We explain the timing issues and the approaches we will adopt in more detail below.

**Table 1:** Baseline training and fieldwork schedule

Date	Implementation activity	Evaluation activity
12-16 February 2018		Assessor training week 1
19-23 February 2018	Master training	Assessor training week 2 Observation of master training
26 February – 22 March 2018		Fieldwork
12-16 March 2018	Step down training	Observation of step-down training

Ideally, the baseline data collection would have occurred prior to teachers even being aware of and invited to the TFET training. However, given the starting date of the evaluation contract and timelines of the Ghanaian Institutional Review Board (IRB) to provide ethics clearance, the earliest that baseline data collection could proceed was in the week immediately following the first week of TFET master training. Table 1 sets out the timeline of both implementation and evaluation activities around the baseline data collection.

**Impact on children:** While 16% of the learners had a teacher who reported attending the master training, prior to the baseline data collection, these teachers had between zero and a maximum of 16 school days to implement the techniques covered in training. Therefore, it is very unlikely that there is an impact of the TFET training on learner outcomes measured in the baseline data collection. To evaluate the impact of TFET on learner outcomes, we will therefore use a combination of matching and differencing.

**Impact on teachers:** The TFET training could feasibly have an immediate impact on some of the outcomes we measure, particularly attitudes towards teaching and discipline. To assess the impact of the TFET program, we will therefore not use a difference-in-differences approach but will rather employ a propensity score matching strategy. We have a rich set of school, learner and teacher level variables from the baseline to create an appropriate comparison group through matching. The variables used in matching will be those that could not be influenced by having attended the master training or having received step-down training. Annex III provides a summary of the main teacher, school and learner variables that will be used in matching. There is considerable overlap, or common support, between the treatment and control groups on these variables making them appropriate for our propensity score matching approach.

### Process Evaluation Design

Alongside the quantitative data collection to inform the impact assessment, the evaluation team is conducting several qualitative research activities to observe and understand the design and implementation of the TFET training program.

Prior to the teacher training weeks, NORC assessed the TFET training material with a particular focus on coherence, depth, consistency and feasibility while taking the local context into account. We gave due consideration to pace, assessment and sequencing.

For the rollout of training, NORC's Education and Training Expert attended the master training session and assessed the skills and the approach to training used by the master trainers. At the end of the master training session, NORC expert administered a short questionnaire to trainees, as part of the TFET evaluation. The questionnaire probes the respondents' confidence in understanding the key pedagogical concepts underpinning the TFET program. Second, the same expert observed the step-down training in two schools, focusing particularly on fidelity to master training and consistency across trainers.

Approximately six months after training implementation, NORC plans to engage with teachers in focus group discussions (FGDs) and observe their teaching techniques in the classroom. Both activities are in pursuit of understanding how trainees perceived the training and were able to incorporate it into their daily work in schools. More information about the FGDs and classroom observations is detailed in Section 4.4 Midline and Endline Data Collection.

## 4.2 School and Learner Sample

---

We provide the details of the calculations used to arrive at the target sample sizes for the impact evaluation of learner outcomes in Annex I. As with any quantitative analysis based on survey data, the

required sample size is determined by a mathematical calculation that depends on a number of factors. These include features of the study design, properties of the data and outcome variables, and the desired precision of the analysis. To reach the desired statistical precision, NORC set a target sample of 40 schools in receiving the intervention and 40 comparison schools, with 10 learners in KG2 and 10 learners in P2. Then, the total target learner sample size across the treatment and control groups is 1,600 learners. In all cases, learners were randomly selected from those present on the day of the assessment. All learners were included unless they refused to participate or severe disabilities precluded them from participating.

### 4.3 Baseline Data Collection

The baseline data collection for the TFET impact assessment occurred over five weeks in February and March 2018. The field teams conducted advanced visits or telephone calls to head teachers of the 80 treatment and control schools targeted for assessments and interviews to explain the study and the purpose of the data collection visit. Subsequently, six enumeration teams conducted face-to-face computer-assisted interviews with up to 10 KG2 learners, 10 P2 learners, 2 teachers, and the head teacher at each sample school. They also completed a school inventory to document the physical school environment.

The total number of completed schools, teacher interviews and learner assessments as compared to the targeted number for treatment and control groups are shown in Table 2. A total of 80 low-cost private schools across eight districts in two regions in Ghana were surveyed. Two schools were replaced; one declined to participate and the other was deemed temporarily non-operational. In total, there were two learners replaced, and both were due to a language barrier (French-language speakers).

**Table 2.** Summary of Baseline Data Collection by Treatment Status

	Achieved Sample: Treatment	Achieved Sample: Comparison	Total Sample	Target Sample
<b>Schools</b>	40	40	80	80
<b>KG2 Learners</b>	339	372	711	800
<b>P2 Learners</b>	357	363	720	800
<b>KG2 Teachers</b>	39	39	78	80
<b>P2 Teachers</b>	39	40	79	80
<b>Head Teachers</b>	40	40	80	80
<b>Material Inventory</b>	40	40	80	80

## Selection of Individual Respondents

The respondent selection process started with teachers in KG2 and P2 grades. If only one teacher for the grade was available, he or she was automatically selected for interview. If there was more than one teacher in the grade, the teacher selection was randomized using either the flip of a coin or a roll of a dice to determine whom to interview.

The study sample of KG2 learners came from the classroom of the selected KG2 teacher; likewise for the P2 learners and the P2 teacher. The process of selecting the 10 learners from the classrooms for interview was the following: first, any children who were blind, deaf<sup>6</sup>, or whose parents refused consent to participate were not included; second, the remaining learners were placed in a line and each learner one-by-one chose a colored bead from a bag. The bag contained 10 green beads (selection for interview), 2 purple beads (replacements in case a sampled child cannot complete the interview), and enough clear beads for the other children.

The head teacher was interviewed at each sample school if he or she was available. In other cases, the school proprietor was interviewed instead.

## Instrument Design

The design of the baseline assessments and survey questionnaires was led by NORC, with input from IDPF and Sesame Workshop. The questionnaires are designed to measure the outcomes of interest to test the research hypotheses, as well as to collect additional information on school, teacher, and learner characteristics to be used for matching and/or control variables in order to improve the precision of the analysis. In addition to the assessments and interviews, a brief school inventory was designed to collect school-level data on the physical environment in and around the school.

The instruments were submitted to the NORC and the Ghanaian Institutional Review Board in November 2017 for ethics clearance. The KG2 learner and teacher survey tools were pre-tested with two learners and one teacher in December 2017 to assess questionnaire flow, respondent comprehension, and estimated time to administer. The P2 learner survey tool was closely adapted from the 2015 Ghanaian national study, hence, did not require a pre-test.

The main sections of the **KG2 assessment and interview** were adapted from the International Development & Early Learning Assessment<sup>7</sup> (IDELA), which aims to measure four domains—motor development, emergent language and literacy, emergent numeracy/problem-solving, and socio-emotional skills. The IDELA tool has been used in 40 countries, including Ghana. The tool was

---

<sup>6</sup> The IDELA and EGRA/EGMA assessments are not designed for children who have severe visual or hearing impairment.

<sup>7</sup> Save the Children, 2018. IDELA. <https://idela-network.org/about/>

successfully adapted for the Ghanaian context by IPA in a study of pre-schools in Accra. The assessment and interview required approximately 30-35 minutes to complete with a child.

The **P2 learner assessment and interview** included Early Grade Reading Assessment (EGRA) modules and Early Grade Mathematics Assessment<sup>8</sup> (EGMA) modules that were developed for inclusion in a 2015 national study of P2 learners in Ghanaian government schools, and is similar to other mathematics or reading exams that a child may take during the course of the school year. The national study had serious floor effects with very high percentages of learners scoring zero on the sub-tasks. We therefore, included some emergent literacy and numeracy outcomes from IDELA and remove harder parts of the EGRA/EGMA assessments. We also use adapted IDELA questions to assess social emotional learning. The interview time was approximately 30-35 minutes.

The **teacher questionnaire** collected data on teachers' education, experience and demographics, their classroom practices, the support and supervision they receive, beliefs around pupil-centered learning, gender equality and discipline, job satisfaction, motivation and burnout. Many of the questions in the teacher questionnaire were taken from a New York University/IPA study of pre-schools in the Greater Accra Region. The average time to administer was 30 minutes.

The **head teacher questionnaire** gathered information from head teachers regarding their instructional leadership including their education background and training, coaching, support, and monitoring of teachers in school, beliefs around pupil-centered learning, gender equality and discipline. During the head teacher interview, we also asked to see teacher and learner attendance records. The average time to administer was 40 minutes.

The **school materials inventory** was a brief observational tool that recorded key aspects of early childhood and primary school classroom learning environments, such as details of space, facilities, and availability of learning materials.

The KG2 and P2 learner assessments were translated into the common regional languages of Ga, Asante, Twi, and Fante. The EGRA and EGMA modules were available already in all three languages, and the IDELA was available in Ga and Asante Twi. The limited number of additional questions was translated by the local data collection firm. The teacher and head teacher surveys were administered only in English.

NORC contracted with a survey programmer to develop CAPI (computer-assisted personal interviewing) versions of the survey instruments for learners, teachers, and head teachers, allowing enumerators to use tablets to administer the survey and record responses. CAPI minimizes data entry

---

<sup>8</sup> University Research Co., LLC, 2017. EdData.  
<https://globalreadingnetwork.net/eddata/ghana-egra-and-egma-instruments-2015>

errors and improves real-time quality control, compared to paper and pencil. The KG2, teacher, and head teacher instruments were administered using the Nfield CAPI tablet application<sup>9</sup> and the data was uploaded directly to the NORC server for download and quality reviews. The P2 learner survey was programmed and administered in Tangerine<sup>10</sup> – the most common software application used for EGRA and EGMA. Multiple rounds of single-person beta-testing of the instruments were conducted by NORC and our local subcontractor, Kantar TNS RMS. The instruments were refined and improved during interviewer training and finalized before the start of fieldwork.

## Training and Pilot

The enumerator training and pilot tests were held over eight days from February 12-21, 2018 in Accra. There was six days of classroom review and two days for pilot tests. Thirty enumerators, four supervisors and six quality control officers were in full attendance. Enumerators were hired from a pool of experienced field staff in a database maintained by Kantar TNS in Ghana. The hired enumerators had an average experience of 4 years in survey administration with a minimum of tertiary education qualification. Also, most of the enumerators had experience with administering learner assessments. The majority of enumerators were hired from the study regions, and therefore were fluent in the local dialects used for the survey.

The training was facilitated by Cally Ardington (NORC), Alicia Menendez (NORC), Ruth Essuman (Kantar TNS), and Jonathan Addie (Kantar TNS). The classroom training included:

- Project background and purpose of the study
- Explanation for and signing of the pledge of confidentiality
- Roles and responsibilities of each team member
- Tablet, Nfield application, and Tangerine application instructions
- Administering informed consent forms
- Overview and a detailed question-by-questions review of the questionnaires and assessments
- Group and paired role play and mock interviews
- Inter-rater reliability testing
- Sensitization to the context and concerns of low fee private schools by Raphael Akomeah from IDP Schools Rising

---

<sup>9</sup> NIPO. Online, CAPI and CATI survey solutions for MR professionals. <https://www.nipo.com>

<sup>10</sup> Tangerine. Mobile Assessment Made Easy. <http://www.tangerinecentral.org/>

Two pilot tests were conducted with the full team during the training phase. Pilot tests give trainees the opportunity to conduct interviews and assessments with real respondents and test the electronic applications for any remaining bugs. One pilot test was conducted on the fourth day of training in two schools in the Madina community of Accra and the second test on the seventh day of training in two schools in the community of Labadi. Conducting two pilots—one of them mid-training, and both with full days for debriefing afterward—provided sufficient time for re-training and more practice in key concepts that were not well administered during pilots. Data from the completed pilot questionnaires were used to test-run the data download and data quality review process developed by NORC.

## Data Collection

Data collection took place from February 26<sup>th</sup> to March 22<sup>nd</sup>, 2018. The survey team comprised 29 field staff in total—6 teams of 4 enumerators; 3 supervisors overseeing two teams each; and two mobilizers. The Greater Accra region was assigned two teams, and four teams worked in the Central region. The supervisors ensured strict observation of school entry protocols and protocols for conducting the learner assessments and other surveys. The mobilizers were responsible for visiting the schools in advance for the team visit, in order to inform the schools of the upcoming work and to deliver the parental consent forms and other official letters.

The Quality Control (QC) was implemented at three levels. The first layer of QC was conducted by supervisors and coordinators who monitored the surveys, especially in the first three days of fieldwork, to ensure strict adherence to interviewing standards and protocols. Supervisors accompanied enumerators at interviews (“accompaniments”) and occasionally conducted “spot checks” (e.g. briefly listened to a part of an interview) to ensure compliance to survey protocols. The second layer of QC was conducted by a trained, independent QC team. Six QC officers were deployed over the survey period with each team member assigned to a survey team. Overall, the QC team conducted 12% accompaniment and 4% spot-checks on the achieved sample. Finally, the third layer of QC was conducted by NORC’s office. NORC’s data analyst conducted a thorough data quality review of the incoming data to identify and resolve irregularities and inconsistencies in the data.

## Challenges

As with any survey data collection processes, there were a number of challenges and unexpected circumstances that arose. Overall, these challenges were typical of survey data collection in similar contexts, and do not raise serious concerns about the quality of the data. The main challenges, resolutions, and implications are summarized as follows:

- Several last minute adjustments needed to be made to the sample due to take up of the Sesame Workshop TFET training. In the process of IDPF confirming attendance at the TFET training,

several treatment schools indicated that they were not interested in attending. Together with IDPF, we decided to increase the pool of schools invited to training in order to ensure that we hit our target of 40 treatment schools. On the first day of training, several schools, that had assured IDPF they would be attending, were not present at the training. We therefore had to replace these treatment schools with schools that were attending the training but were not initially part of the evaluation. Further details around the sampling of schools are provided in Annex I. We show in Annex III that there is considerable overlap in the rich set of school, teacher and learner characteristics in our final sample of treatment and control schools. We are satisfied that there is sufficient overlap in the characteristics of treatment and control observations to support a matching approach.

- At a handful of schools, there was confusion among parents regarding the parental consent forms. According to correct procedures, parents were to sign and return the forms if they did *not* want their children to participate. The field teams found in a small number of schools that parents had signed and returned the form, thinking that they were agreeing to participate. Teachers and other school staff called the parents to clarify the situation, obtained consent, and children were included in the sampling process after all.
- The sampling quota of 10 learners per classroom could not be achieved in some schools due to low enrollment, high absenteeism and/or the non-existence of a KG2 class.
- IDPF informed the evaluation team during field work that some sample schools may have participated in the Quality Preschool Education for Ghana (QP4G) project in Ghana, which has a similar approach to TFET training. This news came after about half of the schools were visited. For the remaining school visits, the teams collected information from the head teacher if they were part of the QP4G project.

#### 4.4 Midline and Endline Data Collection

---

Looking forward, the evaluation team will complete several qualitative research activities during the midline phase to continue observing the training implementation process, the feedback from the trainees, and the carryover of the techniques taught in the TFET trainings to the early grade classrooms. The endline phase, scheduled for 2020, will closely resemble baseline data collection, with an added element of classroom observations.

##### Midline Data Collection

NORC will hold 8 focus group discussions with trained teachers in November/December 2018. The FGDs are focused on the training, materials and implementation challenges and successes. They will elicit feedback on positive and negative aspects of the training sessions, and the content and

usefulness of the new videos. We will also ask for suggestions from teachers for improving the training and videos.

Second, we will conduct in-depth classroom observations with four TFET trained teachers and two teachers in a control school around three months after the second week of training. We will select one KG2 and one P2 teacher in each of the three schools. The classroom observations in the two treatment schools will focus on uptake of new techniques in the classroom, the ability of teachers to implement the techniques and use the materials and the challenges faced by teachers in implementing TFET methods. The observations in the control school will provide a picture of 'business as usual' practice. These classroom observations will be conducted by an expert in pedagogy and classroom practices and will include both low and higher inference items. In addition to the classroom observations, the expert will conduct brief semi-structured interviews with each of the six teachers to further probe the teacher's understanding of the key pedagogic principles of the TFET program.

Third, we will conduct classroom observations in 10 treatment and 10 control schools in May/June 2019. These classroom observations will be semi-structured and will be conducted by fieldworkers trained by an education expert in the use of the semi-structured instrument. These data will provide an early indication of the impact of the TFET training on teacher practice.

## **Endline Data Collection**

Endline data collection, scheduled for June of 2020, will target the same schools and classrooms as in the baseline data collection. The learner assessments will be very similar to those conducted at baseline. Several additions to the endline protocols for teachers and head teacher interviews are listed below:

- In treatment schools, the teacher questionnaire will include questions about their experience implementing the TFET pedagogic approach, use of materials and any challenges faced.
- In treatment schools, we will ask the head teacher about perceived changes in their school's instructional environments after IDPF/Sesame Workshop trainings.
- We will provide each P2 teacher with the list of the 10 learners who were randomly selected for learner assessments in their class. We will ask the teacher to select the top two and bottom two performing learners on literacy and numeracy. A comparison of the teacher's ranking against rankings from the learner assessment data that we collect at endline will provide a measure of the teacher's assessment ability.
- Using a classroom observation tool designed to capture elements specific to the preferred pedagogy of the TFET program we will conduct observations in a subset of surveyed classrooms.

## 5. Baseline Results

This section first presents basic descriptions of the survey data findings and then presents findings from multivariate analysis of the quantitative data. The results of the different surveys are woven together to provide a fuller, more nuanced understanding of each topic of interest in this study. The sections are as follows: basic demographic characteristics of learners and teachers, attitudes towards teaching and learning, discipline practices, gender in the classroom, teacher motivation and satisfaction, KG2 emergent literacy and numeracy and social-emotional skills, P2 early literacy and numeracy and social-emotional skills, school and home environment.

Findings are descriptive of the study sample and therefore, data from schools in treatment and control schools are pooled together.

### 5.1 Basic Demographic Characteristics Of Learners And Teachers

Table 3 shows the basic demographic characteristics of the learner sample separately for KG2 and P2 learners. The sample is roughly equally split into boys and girls for both grades. The average KG2 learner is 5.8 years old and the average P2 learners is 7.9 years old. If learners enter school at the correct age then they should be five or six years old in KG2 and seven or eight years old in P2. Around 12% of learners in KG2 are a year behind for their grade and 7% are two or more years behind. This climbs slightly between KG2 and P2 with 17% of P2 learners one year behind and 8% two or more years behind. The majority of the assessments were conducted in Fante (56% for KG2 and 70% for P2) with a substantial portion of KG2 assessments occurring in English (36%).

**Table 3.** Basic demographic characteristics of learner sample

	KG2 learners	P2 learners
<b>Girls</b>	49.1%	51.1%
<b>Boys</b>	50.6%	48.9%
<b>Age</b>	5.8 years	7.9 years
<b>One year behind for grade</b>	12.1%	17.1%
<b>Two or more years behind for grade</b>	7.3%	8.3%
<b>Assessment language:</b>		
<b>Fante</b>	55.8%	70.0%
<b>Ga</b>	1.7%	9.4%
<b>Twi</b>	6.5%	20.6%
<b>English</b>	36.0%	
<b>Observations</b>	<b>711</b>	<b>720</b>

In each school, we interviewed one KG2 teacher, one P2 teacher and either the head teacher or the proprietor if the head teacher was unavailable. In one school there was no KG2 class, in another no P2 class and in a third the KG2 and P2 classes were taught by one teacher. We therefore have 79 KG2

and 79 P2 teachers in our sample (with one teacher in both). In 66% of schools, we interviewed the head teacher and in the remaining 34% we interviewed the proprietor. In 11% of schools, the head teacher was also the proprietor.

The basic characteristics of KG2 teachers, P2 teachers and the head teacher/proprietor are presented in Table 4. There are stark gender differences between the three groups of teachers. Eighty-six (86) percent of KG2 teachers are female in contrast to 61% of P2 teachers and only 36% of head teachers/proprietors<sup>11</sup>. On average, teachers are between 25 and 27 years olds while head teachers are 48 years old. The vast majority (89%) of head-teachers have some form of post-secondary education. Teachers, however, appear largely to not be qualified with only 23% having some form of post-secondary education<sup>12</sup>. KG2 teachers have, on average, 4.3 years teaching experience while P2 teachers have an average of 3 years of teaching experience. The average KG2 and P2 teachers have been working in the current school for 2.9 and 1.5 years respectively. Head teachers/proprietors have much longer tenure at this school – an average of 7.4 years.

**Table 4.** Basic demographic characteristics of teacher and head-teacher/proprietor sample

	KG2 teacher	P2 teacher	Head teacher/proprietor
<b>Female</b>	86.1%	60.8%	36.3%
<b>Age</b>	27.1 years	24.7 years	47.9 years
<b>Post-secondary education</b>	22.8%	22.8%	88.8%
<b>Years of teaching experience</b>	4.32 years	3.02 years	
<b>Years of experience in this school</b>	2.86 years	1.53 years	7.43 years
<b>Observations</b>	<b>79</b>	<b>79</b>	<b>80</b>

## 5.2 Attitudes towards Teaching and Learning

At its core, the TFET program promotes a paradigm shift from ‘teacher-centered’ to ‘pupil-centered’ pedagogy. As such it is important to understand the extent to which teachers’ underlying beliefs support or are in conflict with ‘pupil-centered’ approaches. We are interested not only in whether the TFET program shifts these beliefs but also the extent to which prior beliefs hinder or facilitate the uptake of the key TFET pedagogical principles.

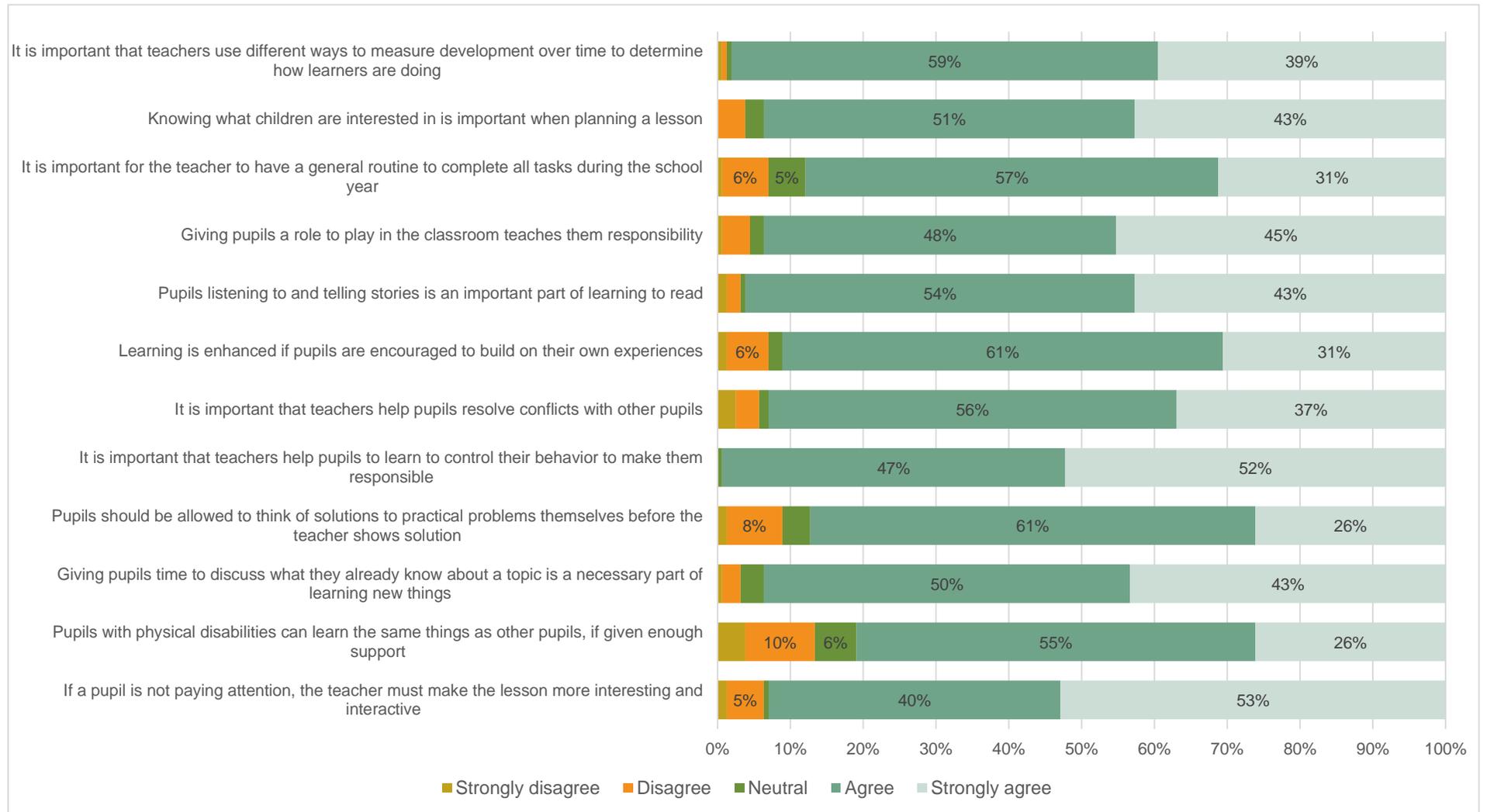
<sup>11</sup> According to the World Bank Development Indicators, in Ghana in 2017, 83% of pre-primary teachers and 41% of primary school teachers were female. The gender breakdown in our sample is similar to that nationally for KG2 but we have a much higher percentage of female teachers in P2 than the national statistics.

<sup>12</sup> The 2014/2015 Ministry of Education report on Basic Statistics and Planning Parameters documents very high percentages of untrained teachers in private schools (Ministry of Education 2015). In public schools in 2014/2015, 62% of kindergarten and 75% of primary teachers were trained. This is in contrast to only 5% and 8% of kindergarten and primary teachers in private schools.

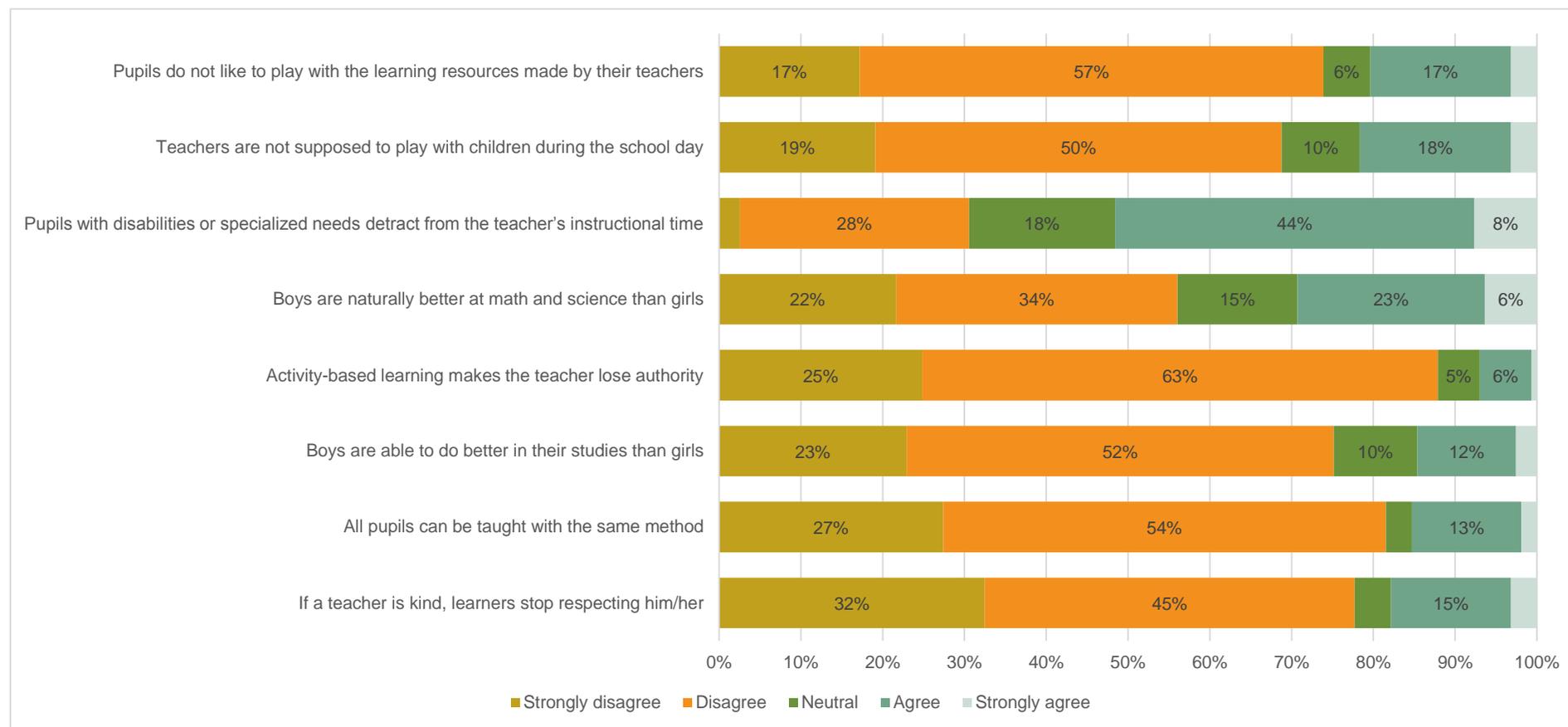
The baseline teacher and head teacher questionnaires asked respondents to indicate on a five-point scale how strongly they agreed or disagreed with 20 statements about teaching. In Figure 2 below, we show the teachers' responses to positively worded learner-centered statements. Figure 3 shows how strongly they agreed or disagreed to statements antithetical to learner-centered learning. For example, the statement, "If a teacher is kind, the learners stop respecting him/her" is in conflict with learner-centered pedagogy and we consider the 32.5% of teachers who strongly disagreed with this statement to be offering a learner-centered response.

There is considerable range across statements in the percentage of teachers who strongly align with the pupil-centered approach. For example, over half (53%) of teachers strongly agree that "if a pupil is not paying attention, the teacher must make the lesson more interesting and interactive" but only one in five strongly disagree that "teachers are not supposed to play with children during the school day."

**Figure 2.** Responses to positively worded statements around learner-centered pedagogy

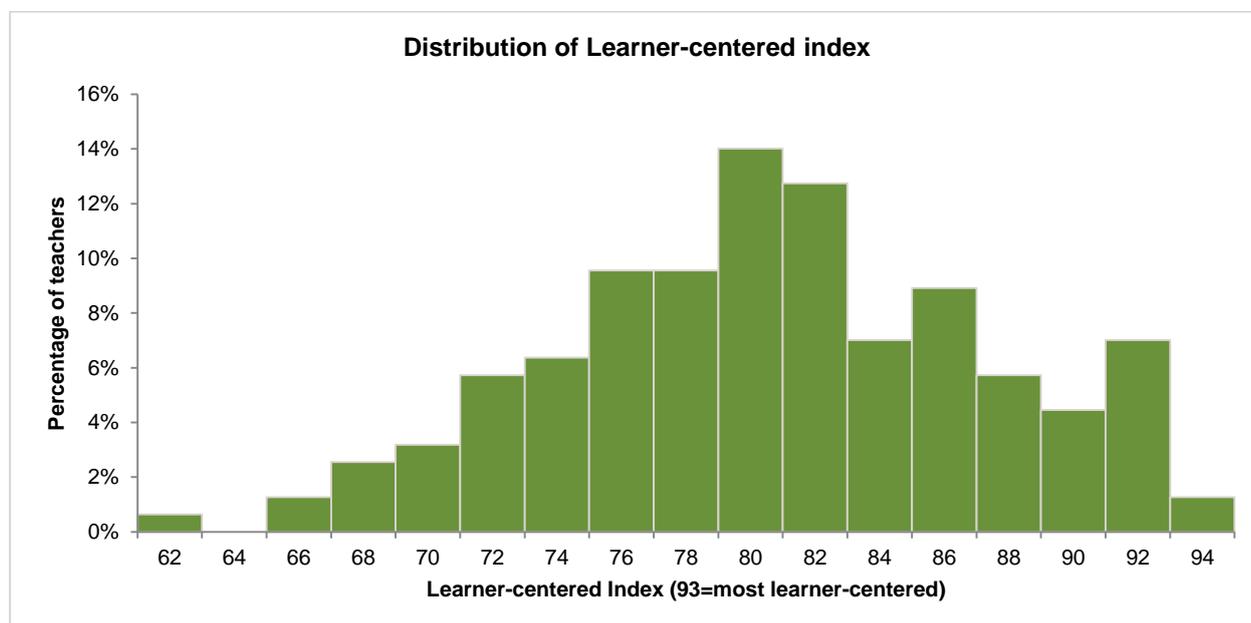


**Figure 3.** Responses to negatively worded statements around learner-centered pedagogy



When we combine the strongly agree and agree responses (or strongly disagree and disagree for non-learner-centered statements), the majority of teachers' responses align with learner-centered beliefs for all statements except "Pupils with disabilities or special needs detract from the teacher's instructional time." Recent research indicates that, despite reform efforts, classrooms in Ghana remain "rigid, chalk-and-talk, teacher-centered, lecture-driven pedagogy, with memorization and rote learning, whereby pupils assume inactive, passive roles."<sup>13</sup> Given this, the seeming alignment with learner-centered pedagogy evident in Figures 2 and 3 likely reflects some desirability bias. Even though each statement is supported by the majority of teachers, there is still considerable variation in responses across the statements allowing us to discriminate between more and less learner-centered teachers. We illustrate this graphically in Figure 4, where we show the distribution of an index of learner-centered beliefs that summarizes the teachers' responses across all of the statements described in Figures 2 and 3 in one index. For statements aligned with pupil-centered beliefs, scores ranged from 5 for strongly agree to 1 for strongly disagree. Negatively worded statements were inversely scored. The index is then the sum of scores across the 20 statements. There is considerable variability across teachers on this index of learner-centered beliefs, as evident in Figure 3.

**Figure 4.** Distribution of Learner-Centered Index

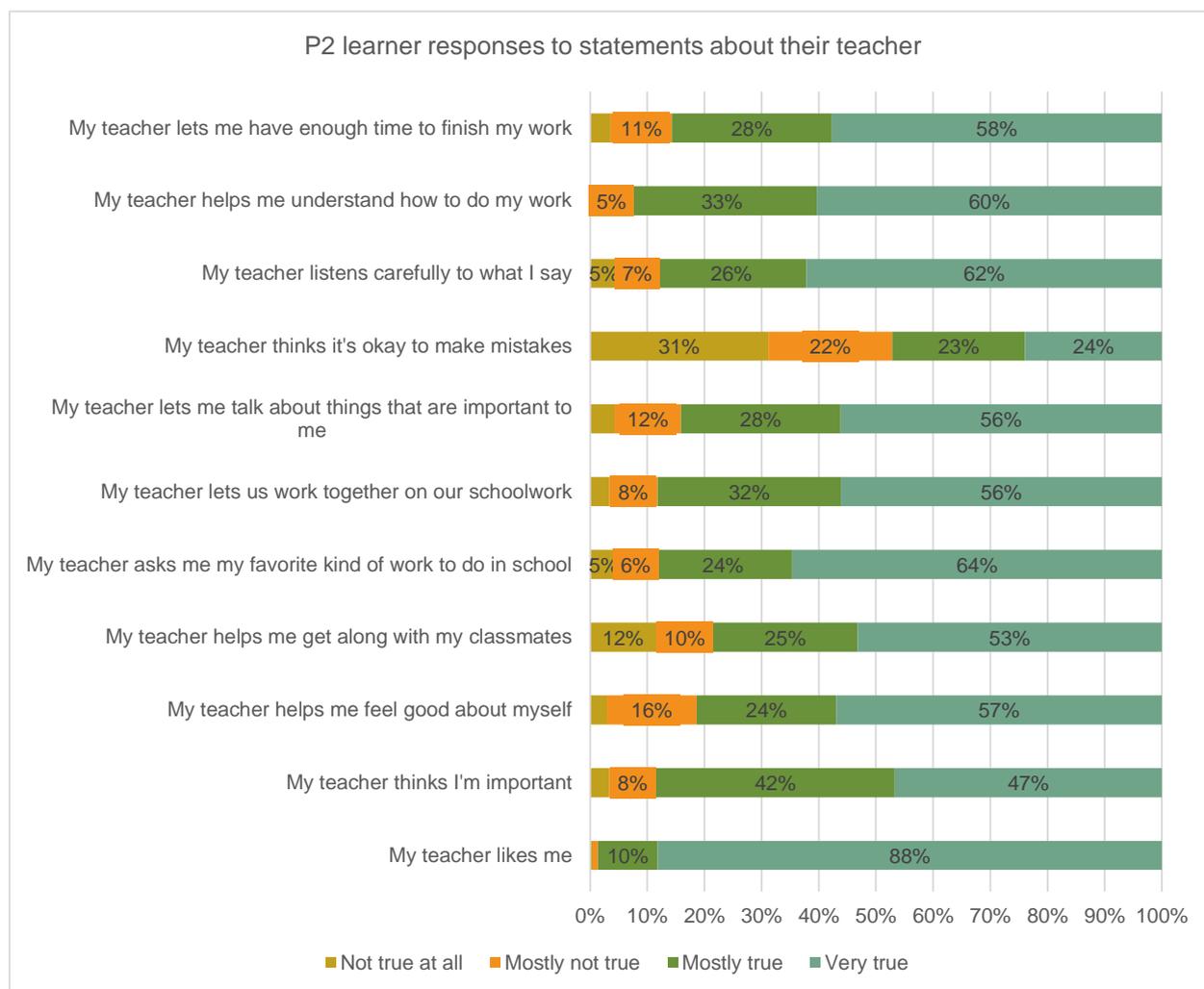


We also asked the P2 learners 11 questions about their teachers in an attempt to probe the extent to which the learners experienced their teachers as being pupil-centered. Questions included "My teacher thinks it's OK to make mistakes" and "My teacher lets me talk about things that are important to me."

<sup>13</sup> Nketsia, W., Saloviita, T., Gyimah, E.K., 2016. Teacher Educators' Views on Inclusive Education and Teacher Preparation in Ghana. *International Journal of Whole Schooling*. <https://files.eric.ed.gov/fulltext/EJ1118090.pdf>

The responses to the 11 questions are summarized in Figure 5 below. Perhaps unsurprisingly for young children, a high percentage of P2 learners consistently expressed beliefs that their teachers liked them and engaged in behaviors that helped them learn. The notable exception is the statement “My teacher thinks it’s okay to make mistakes,” where slightly less than a quarter of the learners (23.8%) thought this was very true of their teacher.

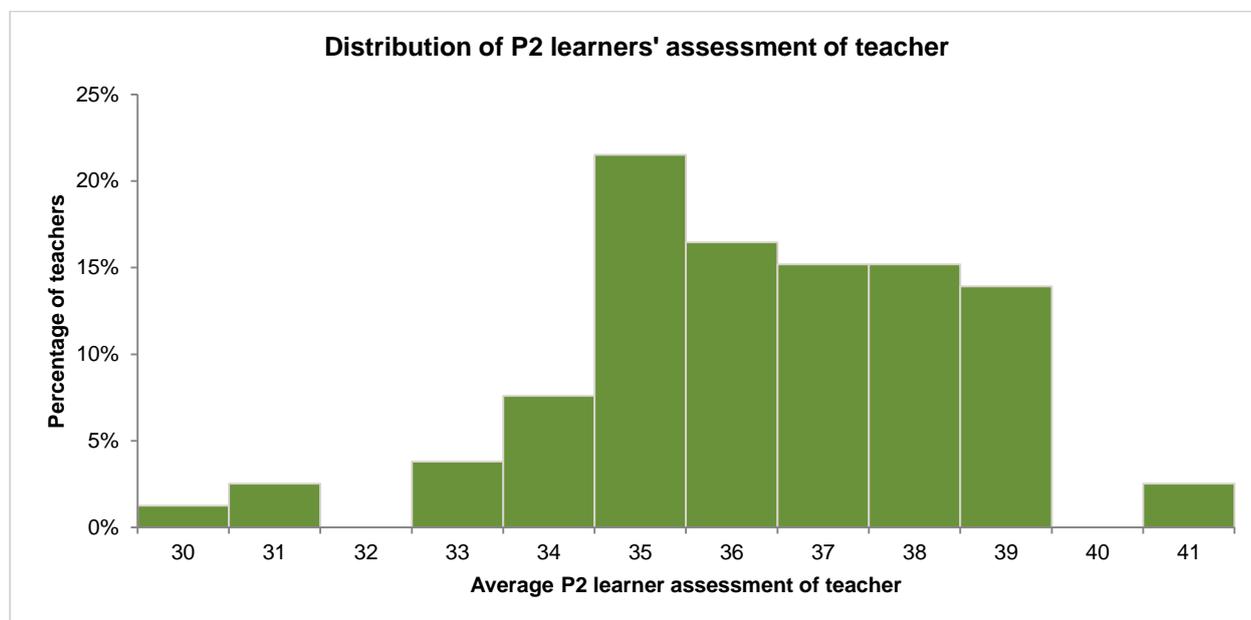
**Figure 5.** P2 learners’ assessment of their teacher



We summarize the learners’ assessments about their teacher in an index. We gave each response a numerical value (i.e., 1 = Not true at all, 2 = Mostly not true, 3 = Mostly true, 4 = Very true), added the responses for each learner across all statements, and found the average across all P2 learners with the same teacher (i.e. in each school). Figure 6 shows the distribution of scores for the resulting index. The possible scores range from a minimum of 11 to a maximum of 44. The average score is 36.8, indicating that the P2 learners were more likely to agree than disagree with statements suggesting their teachers were learner-centered. Similar to the index of learner-centered beliefs shown in Figure 4

above, there is variation in how the teachers are perceived by their learners, enabling us to identify more and less learner-centered teachers.

**Figure 6.** Distribution of learners' perception of their teacher

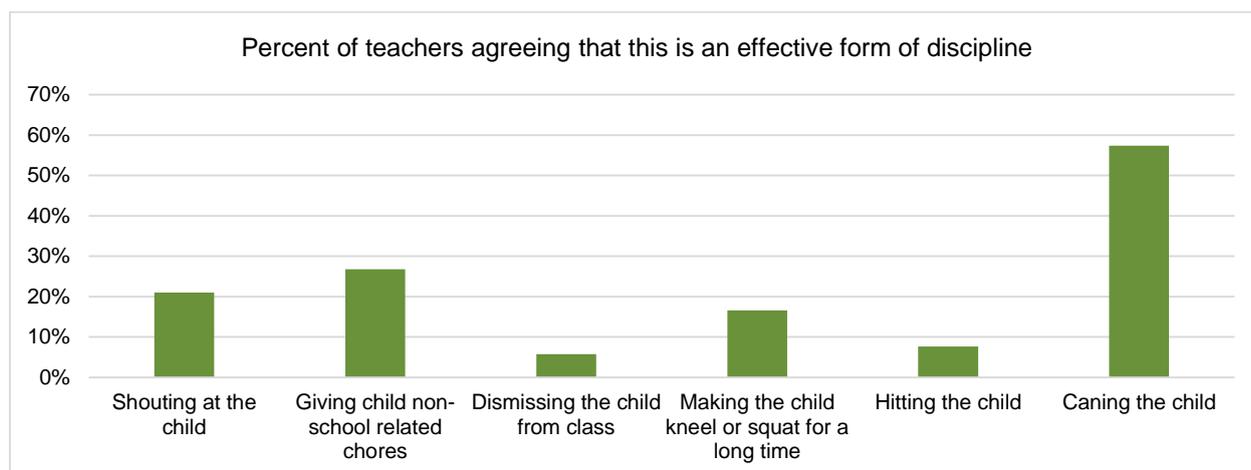


### 5.3 Discipline Practices

Module 3 “Positive discipline and classroom management” of the TFET program focuses on positive discipline. While corporal punishment is widespread, Phase I research found particularly high levels of caning. For ethical reasons, we cannot directly ask whether teachers employ corporal punishment<sup>14</sup>. We therefore asked head teachers and teachers about whether they considered various discipline strategies to be effective. Figure 7 below shows the proportion of teachers who considered each form of discipline to be effective. Dismissing the child from class was the strategy with the least agreement (5.7%) from teachers as an effective strategy, while the majority (57.3%) believes caning is effective.

<sup>14</sup> Corporal punishment is against the law in Ghana. If our field team directly asks a teacher whether they use corporal punishment and they say yes, there is an ethical dilemma. On the one hand they have assured the teacher that their responses are confidential. On the other hand, they would be morally obliged to report the teacher to the authorities.

**Figure 7.** Percentage of teachers agreeing that various forms of discipline are effective



Mirroring results from an evaluation that NORC is conducting in Uganda, caning appears to have much higher social acceptance than other forms of corporal punishment such as hitting the child<sup>15</sup>.

Using data from the P2 questionnaire, we can see the extent to which teachers’ disciplinary practices are reflected in learners expressing fear of coming to school because of the disciplinary strategies teachers employ. Overall, 51% of P2 learners say that children at their school are afraid to come to school for fear of punishment. We took the percentage of P2 learner respondents at each school who said children feared coming to school because of punishment and connected this with the school’s P2 teacher responses on corporal punishment. However, as Figure 8 below shows, children appear only slightly more fearful when their teacher expresses support for corporal punishment. For most types of punishment, the percentage of learners claiming children are afraid to come to school is approximately 4 points higher when the teacher expresses support for the punishment. A potential explanation for why these differences are not greater could be social desirability bias; if some teachers who employ these strategies are reluctant to admit it, their learners would be fearful despite the fact that we do not categorize their teacher as being supportive of corporal punishment.

<sup>15</sup> NORC 2017. Uganda Performance and Impact Evaluation for Literacy Achievement and Retention Activity (LARA): Baseline Report.

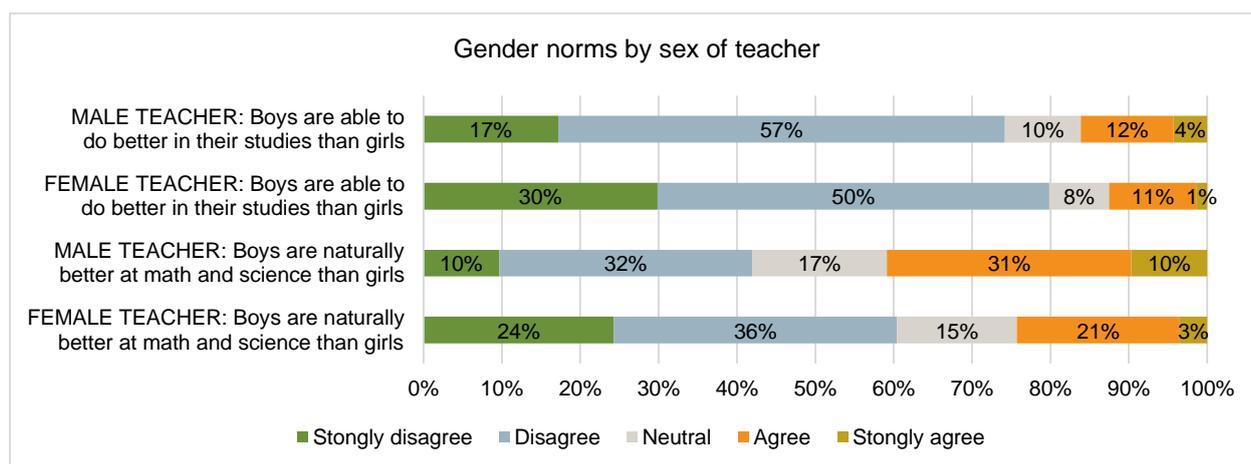
**Figure 8.** Teachers' beliefs about effective disciplinary strategies and learners being afraid of punishment at school



### 5.4 Gender in the Classroom

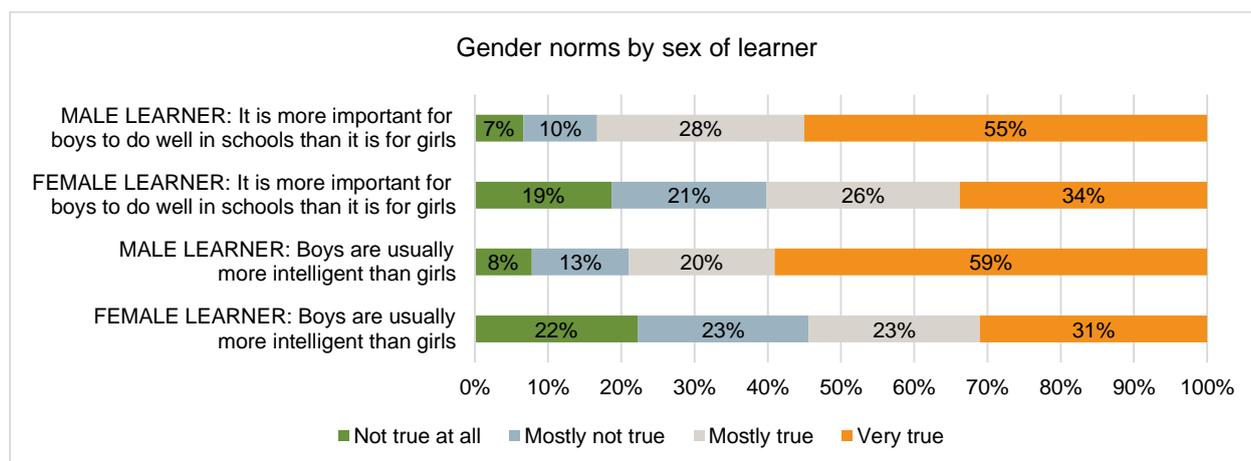
We asked teachers and P2 learners for their opinions towards gender norms using statements. As before, teachers could strongly agree, agree, be neutral, disagree or strongly disagree and learners could indicate that the statement was very true, mostly true, mostly not true or not true at all. Figures 9 and 10 below summarize the teacher and learner responses separately by gender.

**Figure 9.** Gender norms by sex of teacher



Overall, 14% of teachers believe that boys are able to do better than girls in their studies and 31% of teachers believe that boys are naturally better at math and science than girls. Male teachers tend to hold more inequitable gender views than female teachers. For example, 41% of male teachers believe that boys are naturally better at math and science compared to 24% of female teachers.

**Figure 10.** Gender norms by sex of learner



Turning next to the P2 learners’ views, we see much stronger evidence of gender bias, particularly amongst boys. More than half (55%) of boys think that the statement “it is more important for boys to do well in school than it is for girls” is very true. In contrast, one third of girls think this statement is very true. When we combine the very true and true categories, we find that 83% of boys and 60% of girls agree with this statement. Results are similar for the statement ‘boys are usually more intelligent than girls’.

While teachers appear to exhibit less gender bias than learners, social desirability bias is likely to have a much stronger impact on teachers’ reported views than on the views of learners as these tend to be less constrained by social norms and more candid. Interestingly, although the learners hold strongly biased gender views, when asked whether teachers at their school treat boys and girls differently, only 10% of P2 learners responded affirmatively.

## 5.5 Teacher Motivation and Satisfaction

Our teacher survey collected information on the work conditions and morale of the teachers in our sample. This allows us to characterize the teachers according to their contractual and classroom conditions, and perform analyses looking at the relationship between these conditions and professional motivation and job satisfaction.

In Figure 11 below, we show the distribution of teachers by position type. Approximately 46% of teachers have a temporary position, a figure that is higher for P2 teachers (52%) than for KG2 teachers (40%).

**Figure 11.** Teacher contract status by grade

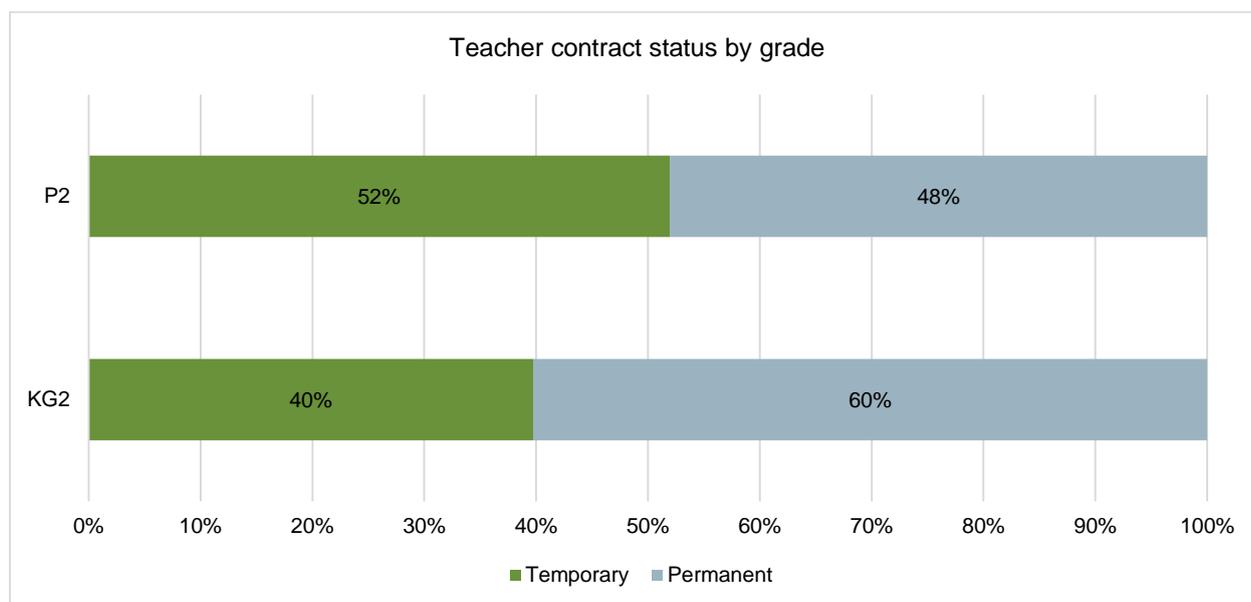


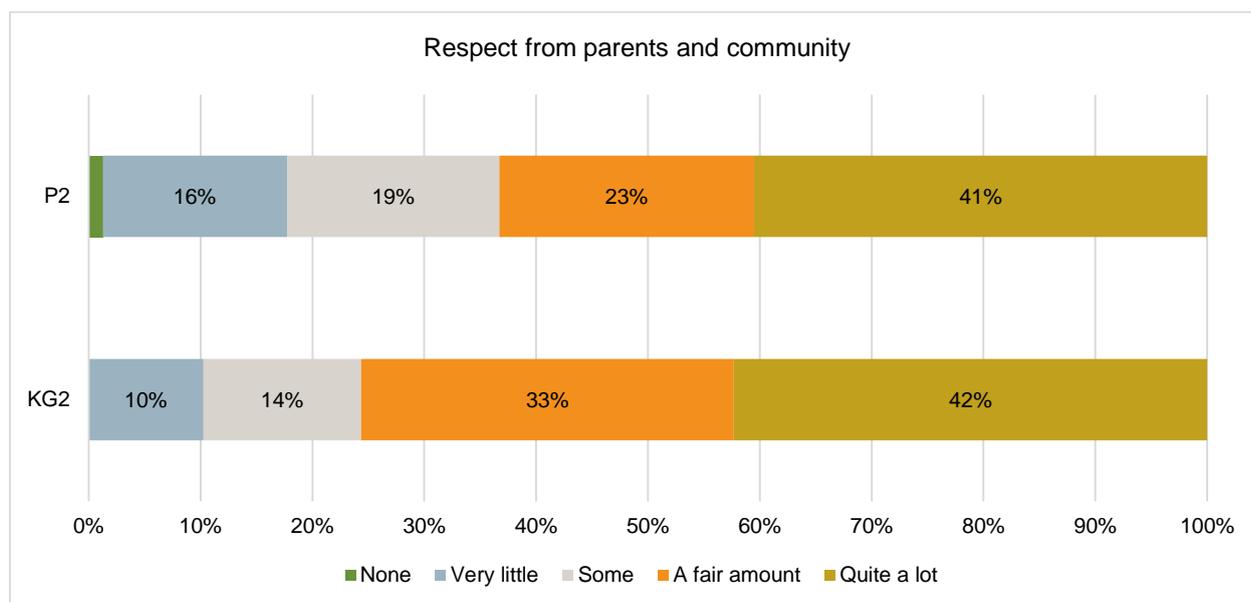
Figure 12 shows the distribution of teachers' salaries by grade. Although KG2 teachers are more likely to have a permanent position, they tend to earn lower salaries.

**Figure 12.** Distribution of teacher salaries by grade



In general, the teachers in our sample report receiving relatively high levels of respect from parents and community leaders. Just under 70% of teachers report receiving either “quite a lot” or “a fair amount” of respect, a figure that is higher for KG2 teachers compared to P2 (Figure 13). Only one P2 teacher in the entire sample reported feeling they received no respect from parents and community leaders.

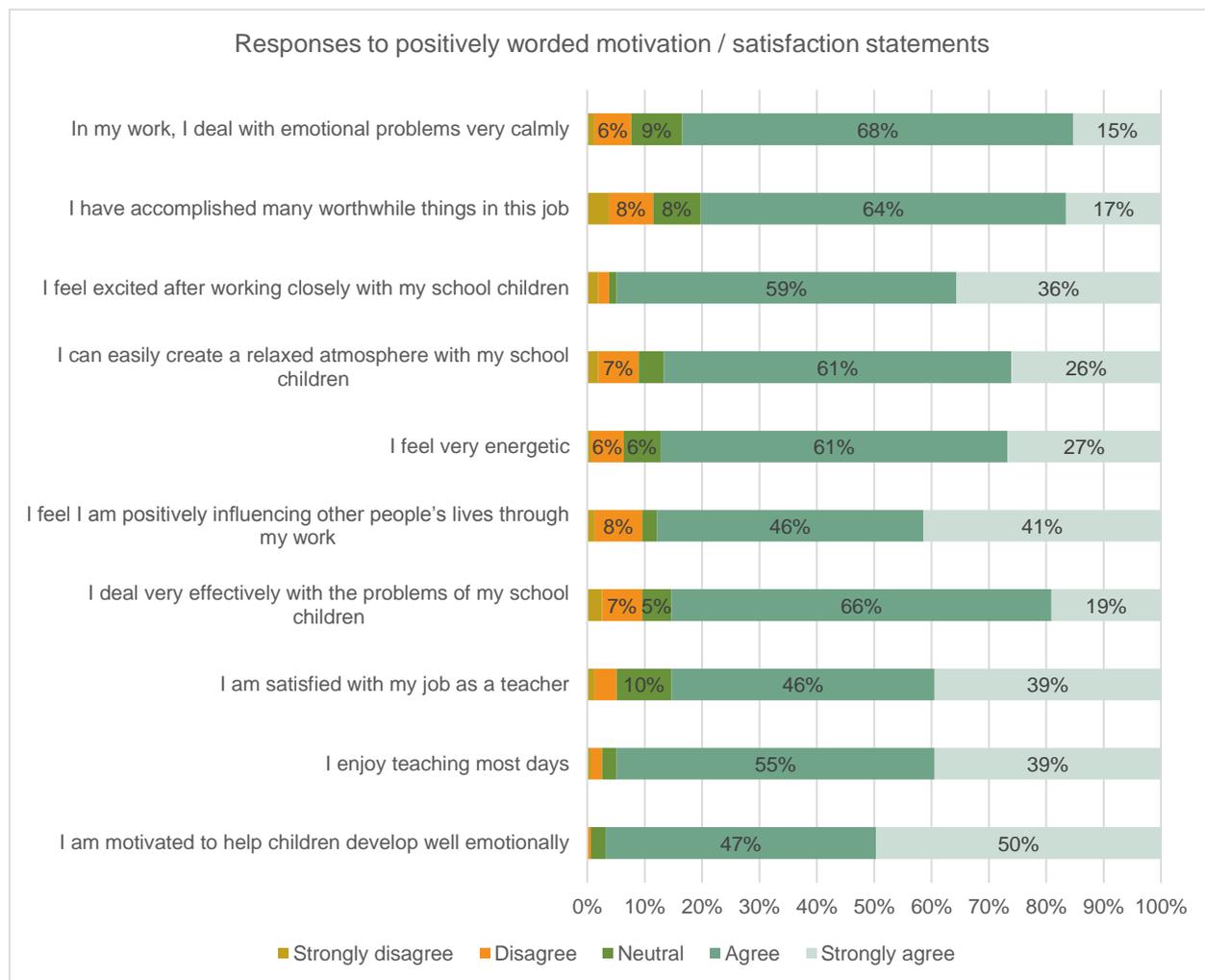
**Figure 13.** Respect from parents and community leaders for teachers



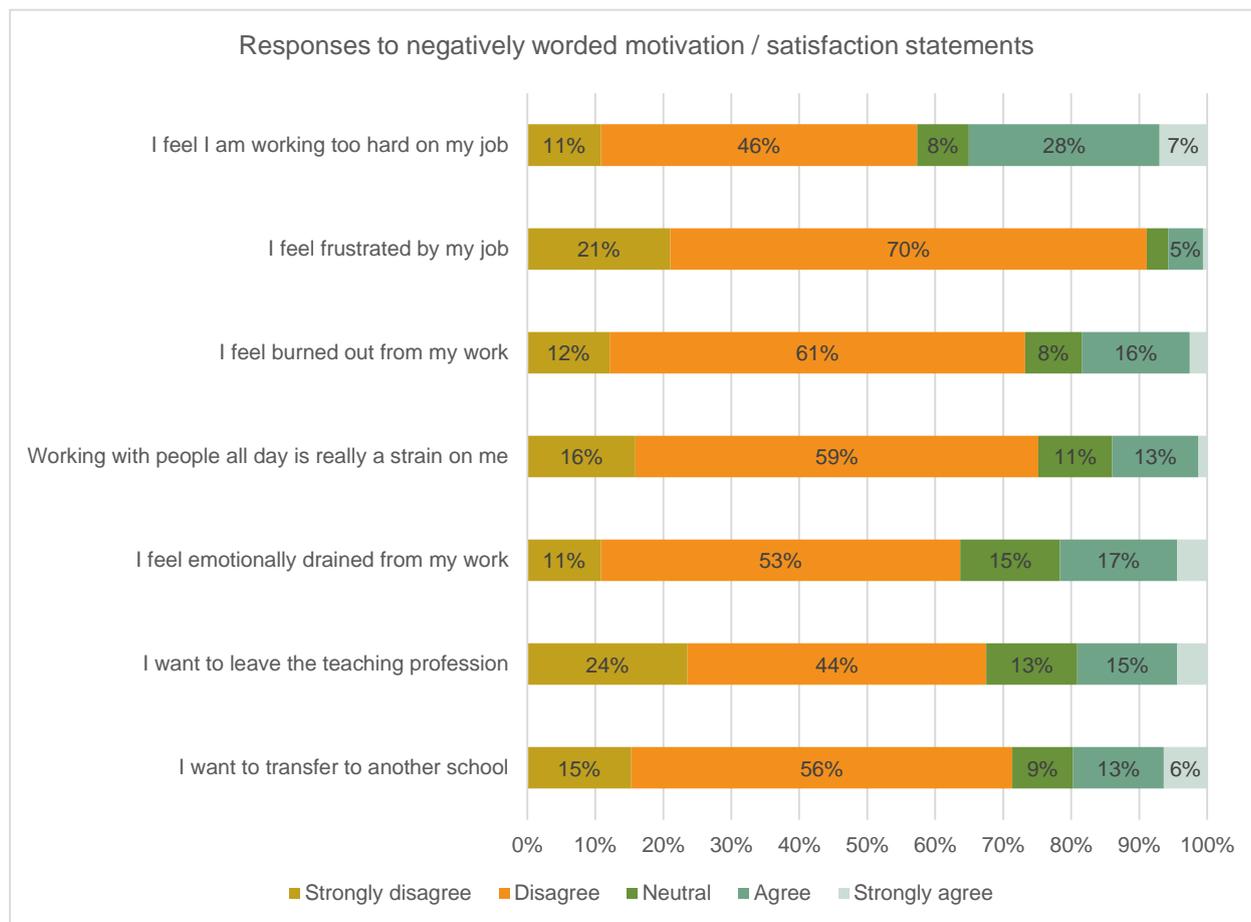
As suggested by the levels of respect teachers report receiving, self-reported motivation among teachers is moderately high. Along a list of 17 statements, teachers were asked the extent to which they agreed or disagreed with each on a scale of 1 (Strongly Disagree) to 5 (Strongly Agree). Figure 14 summarizes the responses for positively worded statements while Figure 15 shows the distribution of response for negatively worded statements.

Starting with the positively worded statements, there is considerable variation in the percentage of teachers expressing a very positive outlook with the strongly agree response ranging from 15% to 50%. However, when we combine the “agree” and “strongly agree” responses, teachers tend to overwhelmingly have a positive outlook. Looking at the negatively worded statements, we see that teachers largely disagree with these statements which signal a positive outlook. The highest agreement is for “working too hard” for which 35% of teachers would be considered to have a negative outlook (i.e. they agree or strongly agree with the statement).

**Figure 14.** Responses to positively worded motivation / satisfaction statements

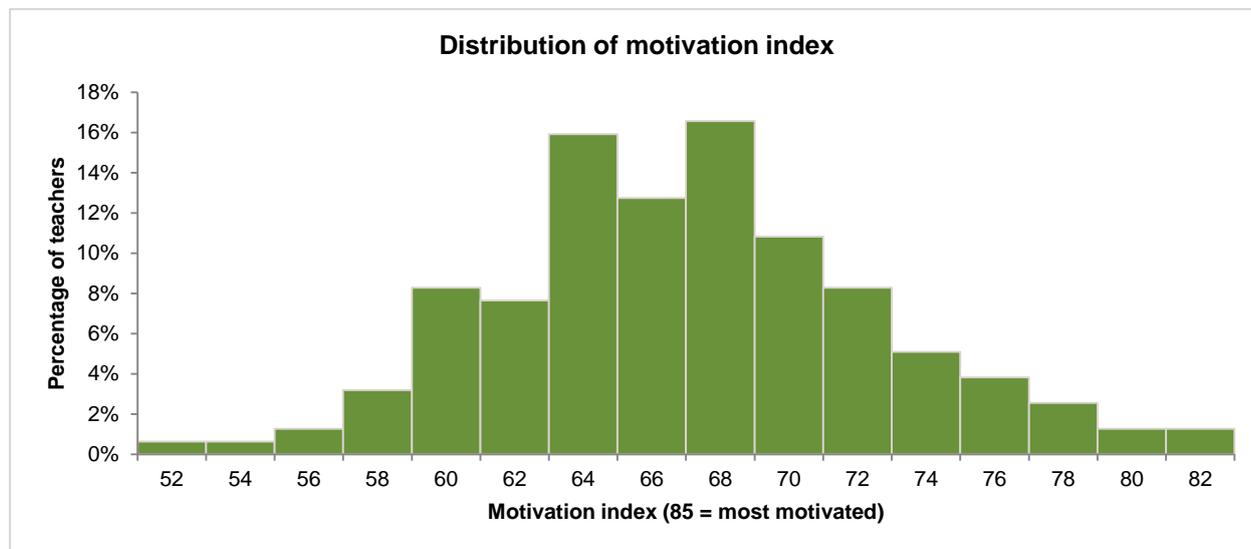


**Figure 15.** Responses to negatively worded motivation / satisfaction statements



Similar to the process for creating the learner-centered teaching index, we took teachers’ responses to the 17 job satisfaction questions and created a teacher motivation index, which ranges from a minimum possible of 17 to a maximum possible of 85. While overall motivation appears high, there is a fair degree of variation in responses evident in Figure 16 below. The teachers in the sample have index values that range from 52 to 82, with a mean of approximately 67.

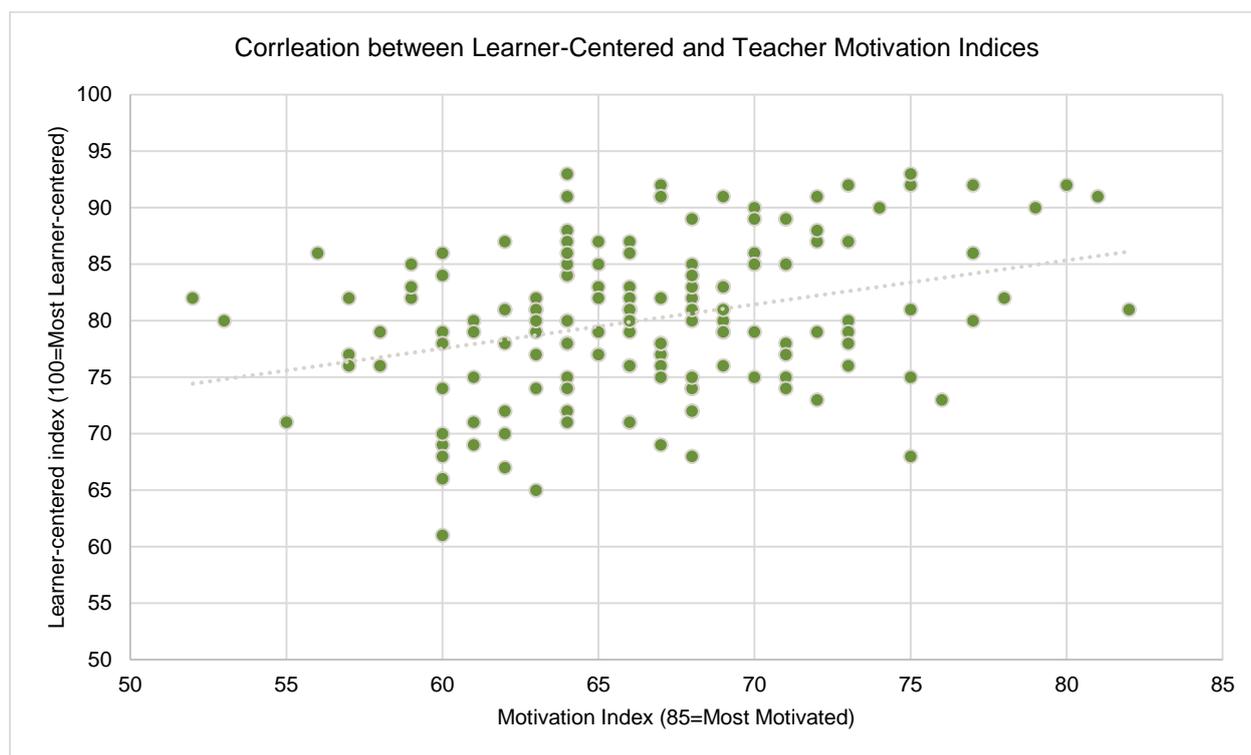
**Figure 16.** Distribution of motivation index



Of particular interest is the relationship between a teacher’s level of motivation and their learner-centered teaching practices. As shown in the scatterplot in Figure 17 below, there appears to be a correlation between the Teacher Motivation Index and the Learner-Centered Index based on teachers’ responses. The relationship between the two indices is highly statistically significant ( $p < 0.01$ ), and our estimates suggest that an increase of 10 units in the motivation index is associated with a nearly 4-unit increase in the learner-centered teaching index.

However, we are careful not to draw conclusions regarding causality in the relationship - it could be the case that higher job satisfaction and motivation lead to more learner-centered teaching practices, but it could also be the case that employing more learner-centered teaching practices leads to higher job satisfaction, or that the two are linked via some other mechanism, such as professional training. Still, the relationship is interesting and suggests that environments where teachers are demotivated will also see lower-quality teaching in the form of classes that are less learner-centered.

**Figure 17.** Correlation between learner-centered and teacher motivation indices



## 5.6 Teacher and Child Attendance

The TFET program aims to increase the professional pride, motivation and job satisfaction of teachers. One of the ways in which they may manifest is in lower teacher turnover and absenteeism. More child-friendly classroom and learner-centered teaching practices may also increase children’s motivation to attend school.

School attendance is clearly an issue with 48% of the P2 learners reporting that they were absent for at least one day in the preceding week. For the 61 schools that had learner attendance records<sup>16</sup>, P2 learners were at school on 69% of the teaching days in the previous school term. Attendance for KG2 learners was at 65%.

For the 80% of schools that had teacher attendance registers, KG2 teachers were at school on 81% of the teaching days in the previous school term while the average attendance rate amongst P2 teachers was at 90%.

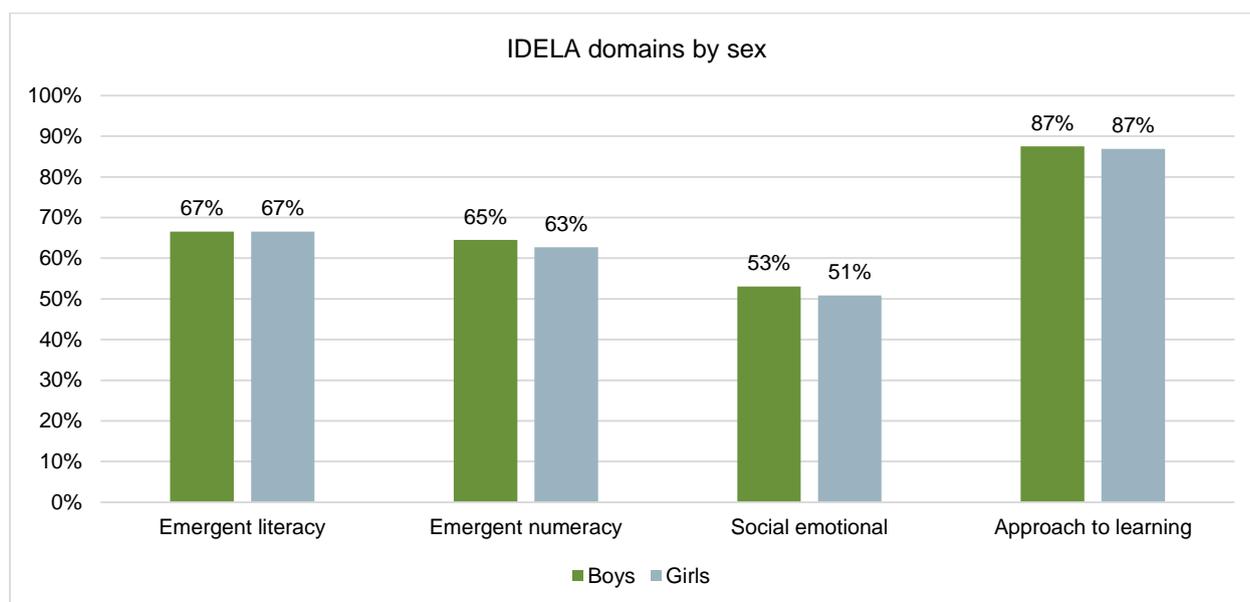
<sup>16</sup> Seventy-six (76) of the 80 schools said they had attendance records but four did not or would not show us the records. Of the 72 records that our fieldworkers were shown, the figures in 7 and 9 schools don’t make sense for KG2 and P2 learners respectively (i.e. attendance is greater than 100%).

## 5.7 KG2 Learner Assessments

The KG2 learner assessment used three of the four IDELA core developmental domains, namely emergent language and literacy, emergent numeracy/problem solving and social emotional skills.

Figure 18 summarizes the scores on the three domains by gender.

**Figure 18.** KG2 learner scores on IDELA domains by sex



In addition to the three domains, we also include an approach to learning score which measures persistence and engagement with the various tasks. Boys and girls tend to score very similarly on all domains. On average, the KG2 learners scored 67% for emergent literacy, 64% for emergent numeracy, 52% for social-emotional skills and 87% for their approach to learning.

Each of the domains was made up of a number of sub-tasks. The scores for the individual items making up the three domains are presented in Table 5 below. There is considerable variation in subtask scores within each domain. Within the emergent literacy domain, almost all learners were able to write their name and most learners demonstrated good familiarity with print and ability to identify letters. Phonemic awareness is fairly undeveloped with 28% of KG2 learners unable to identify the first sound of any words. KG2 learners' performance was reasonable on the more procedural subtasks of the emergent numeracy domain such as length and number identification. The more conceptual puzzle and sorting tasks clearly presented challenges with 32% and 41% of KG2 learners scoring zero for these subtasks respectively. Learners' scores were fairly high for the self-awareness and naming of friends subtasks of the social-emotional skills domain. However, they struggled with the empathy subtask with an average score of only 24%.

**Table 5.** KG2 IDELA Task Summary Statistics

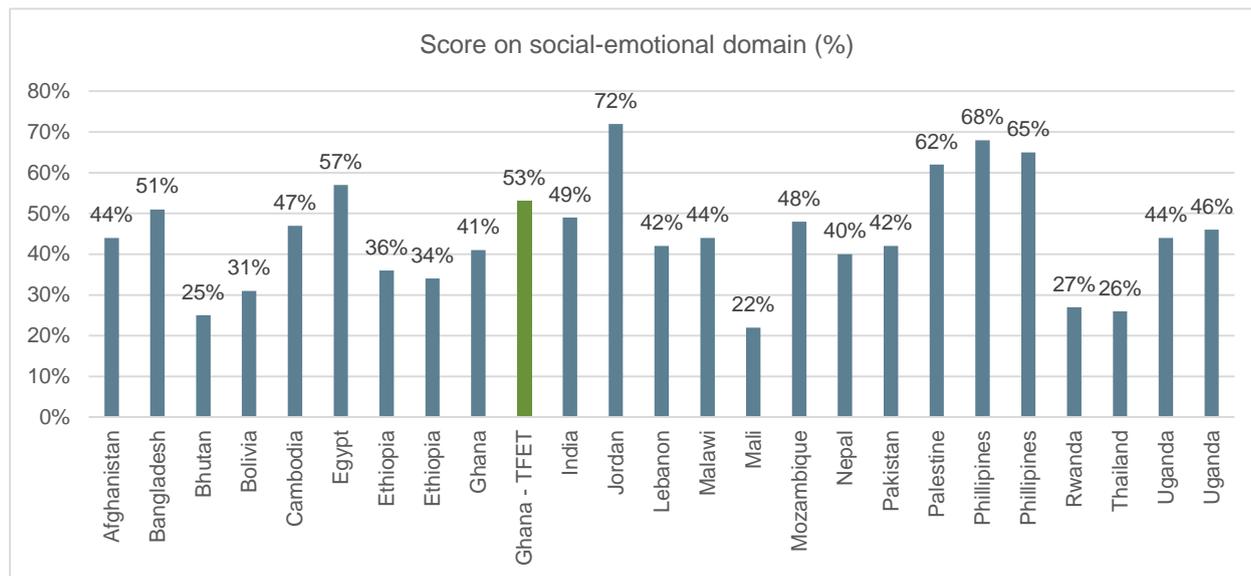
Domain	Subtask	Score (%)
<b>Emergent literacy</b>	Expressive vocabulary	55.8%
	Print orientation	77.0%
	Letter identification	75.8%
	First sounds	41.3%
	Writing	95.6%
	Oral comprehension	54.0%
<b>Emergent numeracy</b>	Length	91.3%
	Sorting	33.9%
	Shapes	65.6%
	Number identification	76.6%
	One to one correspondence	83.5%
	Addition and subtraction	68.4%
	Puzzle	26.2%
<b>Social-emotional skills</b>	Self-awareness	79.7%
	Friends	62.6%
	Emotional awareness	44.9%
	Empathy	24.0%
	Conflict solving	48.6%

There are no benchmarks or norms for the IDELA domains, but we can compare our results to those from other studies that have used the IDELA tool. IDELA scores tend to increase with age so we compare the average scores of the 271 five year olds in our baseline with average scores of five year olds from other studies that have used the IDELA tool. Figures 19 to 21 below show the scores for each of the IDELA domains for all publically available studies using IDELA with a sample of at least 100 five year olds. The other studies are from 21 countries with sample sizes ranging from 104 to 933 five year olds. The results for five years olds in the TFET baseline are shown by the green bar.

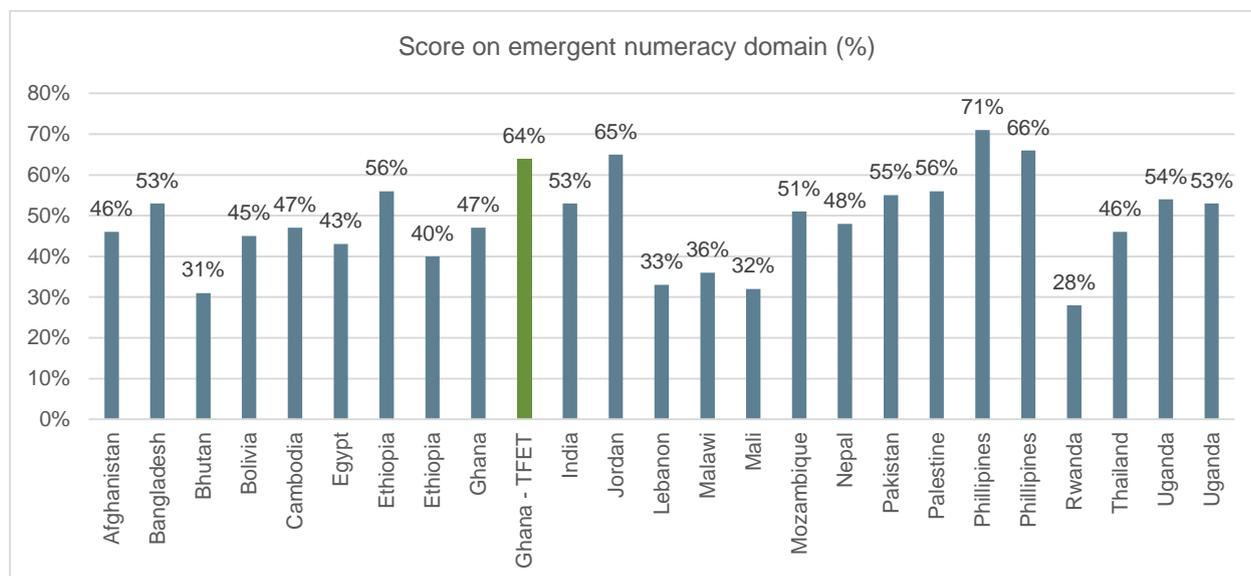
These are not nationally representative studies and the context varies across countries. For example, some of the studies were conducted with refugees. Nevertheless, it is still interesting to compare the results of the five year olds in our sample with those from other studies. The five year olds in our sample have the highest average score for emergent literacy. The learner scores in our sample also compare favorably for emergent literacy and social-emotional skills. Our learners have the fourth highest score for emergent literacy and the sixth highest score for the social-emotional domain.

One of IDELA studies shown in the figures was conducted by New York University/IPA in the Greater Accra region. There were 933 five year olds in their baseline sample. The five year olds in our sample perform better, on average, across all three domains.

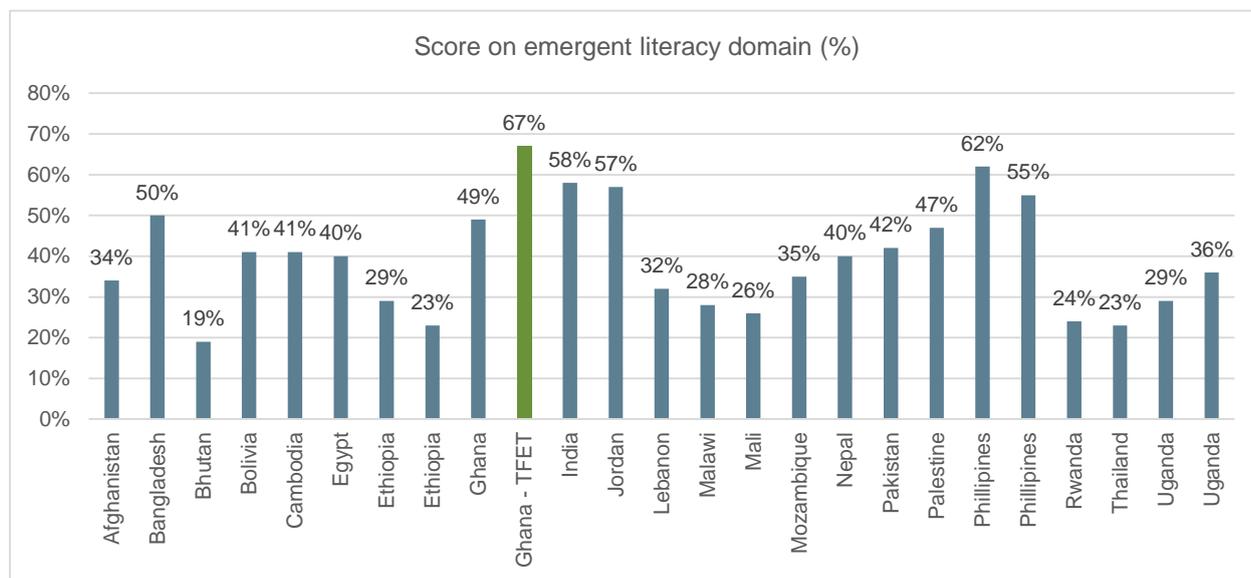
**Figure 19.** Average score on social-emotional domain for children five years of age by country



**Figure 20.** Average score on emergent numeracy domain for children five years of age by country



**Figure 21.** Average score on emergent literacy domain for children five years of age by country



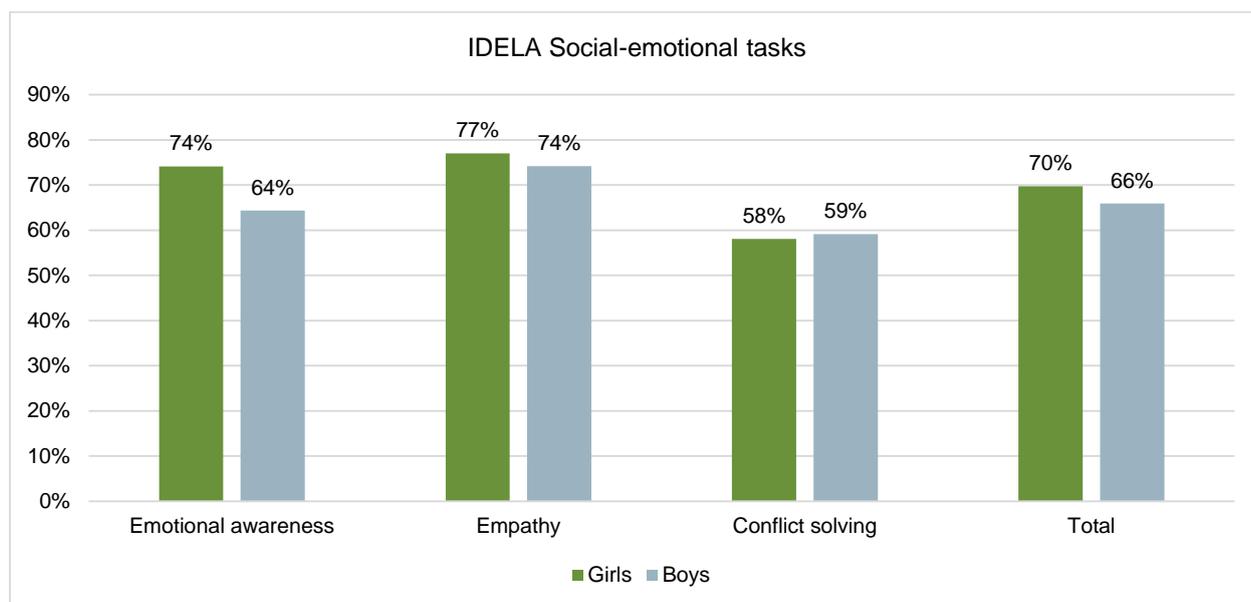
## 5.8 P2 Learner Assessments

### Social-emotional skills

In order to assess the social-emotional skills of the P2 learners, we used a subset of the tasks from the IDELA social-emotional domain. Figure 22 shows the percentage scores for each subtask and a total separately for boys and girls. Girls score 10 percentage points higher than boys on the emotional awareness subtask. There are no significant gender differences between the other two subtasks.

Comparing the scores to those for KG2 learners, we see a clear development in social-emotional scores. On average, KG2 learners scored 45%, 24% and 49% for emotional awareness, empathy and conflict solving. In contrast, P2 learners score 69%, 76% and 59% respectively.

**Figure 22.** Social-emotional skills of P2 learners



### Early Literacy

In order to assess early literacy outcomes for P2 learners we used the emergent literacy item from the IDELA tool and the Early Grade Reading Assessment (EGRA) tool. The EGRA is a brief oral reading assessment that tests a range of literacy skills and has been successfully used in over 50 countries. We used a version of this tool previously used in Ghana in 2015 by the National Education Assessment Unit (NEAU) of the Ghana Education Service (GES). The specific subtasks that we included are: listening comprehension, oral passage reading, and reading comprehension in both local languages and English. The listening comprehension subtask assesses the ability of the child to understand a short story that is read to them. The aim is to measure language comprehension. The oral reading passage subtask measures the ability to read with accuracy and speed a short connected text and the reading comprehension subtask includes questions about the passage read.

The national 2015 study of Grade 2 learners documented very poor literacy outcomes in Ghana with many children unable to read even one word. Due to concerns around floor effects, we also included one of the expressive vocabulary items from the IDELA instrument to measure emergent literacy outcomes.

Figure 23 shows the distribution of scores for the emergent literacy item that we included from IDELA that asked children to name items that they would see in a market in any language that they were comfortable with. The task ended when learners had identified 10 different items. Although almost all learners could identify at least one market item, only 21% were able to identify at least 10. On average, learners identified 6.6 market items.

**Figure 23.** P2 learners score on expressive vocabulary task

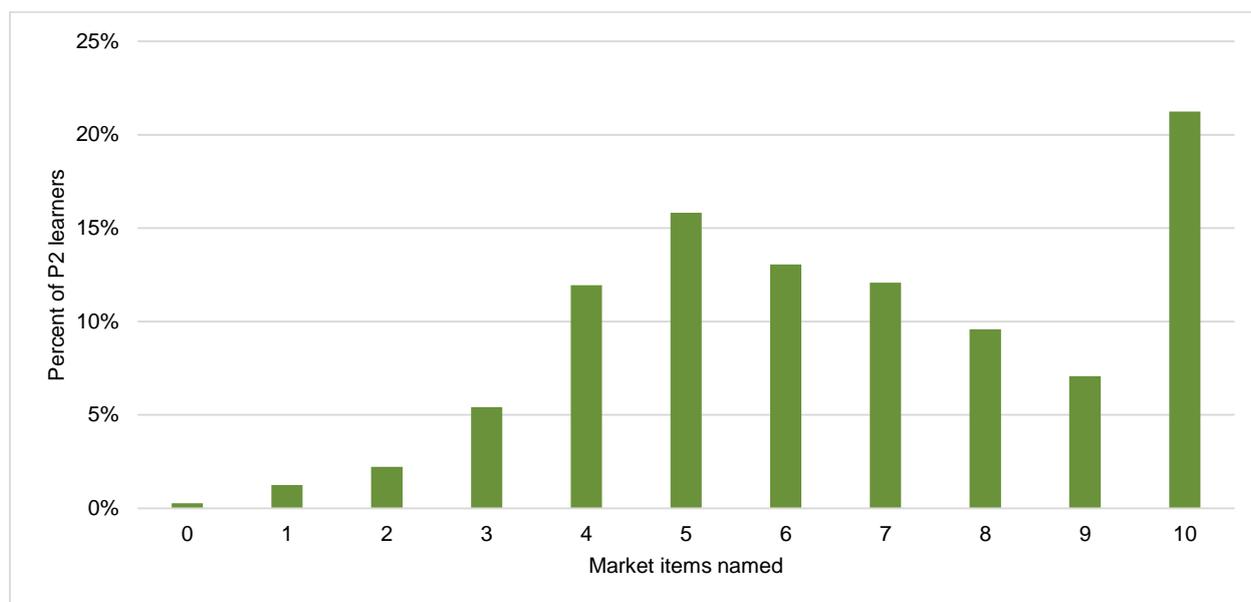
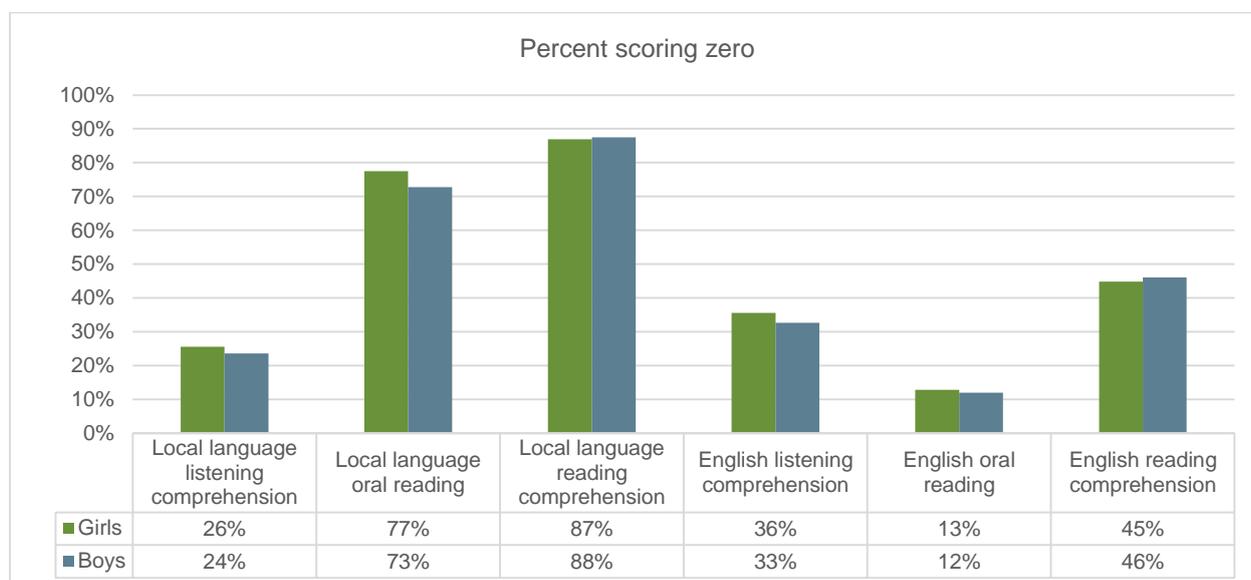


Figure 24 shows the percentage of P2 learners scoring zero for each of the EGRA sub-tasks by gender. There are no significant differences in the percentage of boys and girls scoring zero on any of the EGRA sub-tasks. Starting with the local languages, we see that a quarter of learners cannot answer one comprehension question correctly about a short paragraph read to them and three-quarters are unable to correctly identify one word from the reading passage. Learners were only asked reading comprehension questions related to the section of the passage that they read. Of the 20% of learners who read far enough in the passage to attempt at least one comprehension question, 37% were unable to correctly answer one comprehension question.

**Figure 24.** Percent of P2 learners scoring zero on EGRA subtasks



Turning to the English subtasks in Figure 24, we see that just over a third of learners score zero on the English listening comprehension. Results for oral reading are substantially better with only 12% of P2 learners unable to read one word in English. However, comprehension is poor with 46% of learners unable to correctly answer one English comprehension question.

Table 6 shows the average scores for each EGRA subtask both for the whole sample and for the subset who did not score zero. As seen in Figure 24 above, a very high percentage of learners score zero on each of the sub-tasks. While the overall averages for the EGRA subtasks are important indicators, they are in large part driven by the zero scores. It is also informative to investigate fluency rates and comprehension levels amongst those learners who can actually identify at least one word. For the whole sample, learners read at an average of 5 words per minute in their local language. Among those who can correctly identify at least one word in their local language, learners read at an average of 20 words per minute in local languages. For reading in English, there is a much larger subset who can identify at least one word in English, and they read on average at 38 words a minute.

**Table 6.** P2 EGRA scores

	All learners	Learners scoring greater than 0%
<b>Listening comprehension (% correct)</b>	47.2%	62.6%
<b>Oral reading fluency (correct words per minute)</b>	5.0	19.9
<b>Reading comprehension (% correct)</b>	4.5%	35.0%
<b>English Listening comprehension (% correct)</b>	48.7%	74.0%
<b>English Oral reading fluency (correct words per minute)</b>	33.1	37.8
<b>English Reading comprehension (% correct)</b>	27.1%	50.8%

Table 7 below compares the TFET baseline EGRA results to those from the 2013 National EGRA/EGMA study conducted by the Ghana Education Service and RTI international with the support of USAID. The table also shows the proposed benchmarks and targets that came out of the 2014 Ghana benchmarking exercise. The P2 learners in our sample perform marginally better than the national average for Ghanaian languages but substantially better than the national average for English. For example, in 2013 the average P2 learner in Ghanaian schools was reading at 9.2 English words per minute and could correctly answer 4.2% of the English comprehension. This in contrast to our sample of P2 learners who were reading at 33 English words per minute and scored an average of 12.9% on the English comprehension. While the results for our sample are better than the national average, they are still way off the proposed benchmarks. Only 4%, 1%, 33% and 13% of our P2 learners are reaching the benchmark for Ghanaian language oral reading, Ghanaian language comprehension, English oral reading and English comprehension respectively. Amongst the learners who are able to correctly identify at least one word in their local language, average fluency rates are half the proposed local language benchmark of 40 words per minute.

**Table 7.** Comparison of EGRA scores to 2013 national results and proposed benchmarks

	Ghanaian language Correct Words Per Minute	Ghanaian language comprehension	English Correct Words Per Minute	English comprehension
<b>Average scores:</b>				
<b>2013 National</b>	3.7	3.2%	9.2	4.2%
<b>TFET baseline</b>	5	5.0%	33	12.9%
<b>Proposed benchmark</b>	40	80%	45	80%
<b>Percentage meeting targets:</b>				
<b>Proposed 5 year target for meeting benchmark</b>	13%	5%	35%	8%
<b>2013 % meeting benchmark</b>	2.5%	0.9%	7.0%	1.5%
<b>TFET % meeting benchmark</b>	3.8%	1.3%	32.5%	12.9%
<b>Percentage of zero scores:</b>				
<b>Proposed 5 year target</b>	39.0%	46.0%	25.0%	45.0%
<b>2013 National</b>	78.1%	91.8%	50.4%	90.2%
<b>TFET baseline</b>	75.1%	87.2%	12.4%	45.5%

The results in Figure 24 and Table 6 clearly indicate better listening, reading and comprehension performance in English than in local languages for the P2 learners in our sample. Perhaps some of the reason for the poor performance in local languages is the mismatch between the learner’s home language and the local language used in the assessment. Ghana is a multilingual country and this is reflected in the classroom where learners do not always speak the same mother tongue. This is particularly the case in urban and peri-urban areas where families from different areas of the country live together. Our EGRA assessments were conducted in both English and local language. The local language used in our assessment was the dominant local language of the areas where we collected data. In our sample, just over three-quarters (77%) of learners were assessed in the language they speak most frequently at home. These learners perform significantly better on the listening comprehension - 18% of these learners score zero in contrast to 46% of learners for whom the assessment language is not their home language. Performance on the oral reading and comprehension subtasks is, however, very similar between learners whose home language does and does not match the language of assessment. The advantage of the alignment of home language and language of assessment seems to be limited to listening and to not extend to reading performance.

To some extent, this is not surprising as one of the key reasons that parents choose low fee private schools is a preference for greater exposure to English.<sup>17</sup> Local language instruction is required for

<sup>17</sup> Erling, E., Adinolfi, L., Hultgren, A.K., 2017. Multilingual classrooms: opportunities and challenges for English medium instruction in low and middle income contexts. Education Development Trust.

public schools in Ghana<sup>18</sup> but in low fee private schools instruction tends to be largely in English.<sup>19</sup> Table 8 below shows the languages that KG2 and P2 teachers report using in the classroom. Only one class in our sample is conducted solely in local languages. The vast majority of classes use a mixture of English and local languages.

**Table 8:** Languages used for teaching in KG2 and P2 classrooms

	KG2 classes	P2 classes
English only	7.7%	11.4%
Local language only	0.0%	1.3%
Mixture of English and local language	92.3%	87.3%

### Early Numeracy

In order to assess early numeracy outcomes for P2 learners we used the Early Grade Mathematics Assessment (EGMA) tool respectively. The EGMA is a test of early mathematics knowledge which emphasizes on numbers and operations. It is an oral assessment individually administered to learners by trained assessors. Like in the case of EGRA, we used a version of EGMA that had been used successfully in Ghana in 2015. The EGMA consists of six subtests: number identification, addition and subtraction level 1, missing number, addition and subtraction level 2, and word problems.

Figure 25 shows the percentage of P2 learners scoring zero on each of the EGMA subtasks by gender. Boys are slightly less likely to score zero on all tasks, but the difference is only statistically significant for the basic addition and basic subtraction subtasks. The vast majority of learners were able to correctly answer at least one missing number and one word problem item correctly. While 97% of learners were able to correctly do at least one basic addition, only 62% of learners did not score zero on the more advanced addition subtask. Results were similar for subtraction with the percentage of learners scoring zero rising from 8% for the basic subtraction to 59% for the more advanced subtraction.

<sup>18</sup> Owu-Ewie, C., Eshun, E.S., 2015. The Use of English as Medium of Instruction at the Upper Basic Level (Primary Four to Junior High School) in Ghana: From Theory to Practice. *Journal of Education and Practice*.

<sup>19</sup> Erling, E., Adinolfi, L., Hultgren, A.K., 2017. Multilingual classrooms: opportunities and challenges for English medium instruction in low and middle income contexts. Education Development Trust.

**Figure 25.** Percent of P2 learners scoring zero on EGMA subtasks

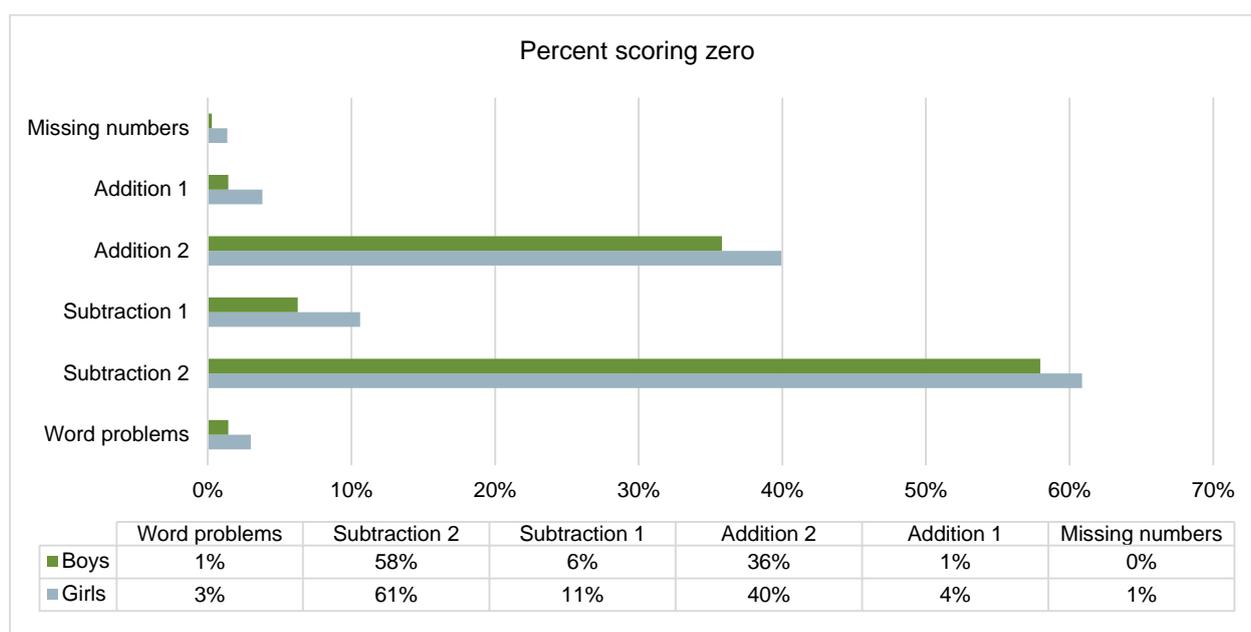


Table 9 shows the average score as a percentage for each of the subtasks for all learners and for the subset of learners who did not score zero. Although not many learners scored zero, average scores on missing numbers, basic addition and subtraction and word problems were low. The percentage correct ranges from 35% to 48% for these subtasks. Among those learners who did not score zero on the more advanced addition and subtraction tasks, average scores were 53% and 44% respectively.

**Table 9.** P2 EGMA Task Summary Statistics

	All learners	Learners scoring greater than zero
<b>Missing numbers (% correct)</b>	39.0%	39.4%
<b>Addition 1 (% correct)</b>	48.0%	49.3%
<b>Addition 1 (correct sums per minute)</b>	9.67	9.93
<b>Addition 2 (% correct)</b>	32.6%	52.6%
<b>Subtraction 1 (% correct)</b>	35.4%	38.7%
<b>Subtraction 1 (correct sums per minute)</b>	7.09	7.75
<b>Subtraction 2 (% correct)</b>	17.9%	44.2%
<b>Word problems (% correct)</b>	45.2%	46.3%

The basic addition and subtraction tasks were timed with learners stopping after 60 seconds. The low scores on these subtasks were driven more by learners working slowly than inaccurately. Of the items attempted within the minute, the average percentage correct was 79% and 69% for addition and subtraction respectively. On average, learners were able to correctly complete 9.7 basic additions and 7.1 basic subtractions in a minute. Children are slowed down by relying on inefficient strategies such as tick marks. On average, children used tick marks or counting on their fingers 55%, 59% and 79% of the time for advanced addition, advanced subtraction and word problems respectively.

Table 10 below compares the results of the P2 learners in our sample with national averages from the 2013 EGRA/EGMA study and shows the proposed national benchmarks. On average, the P2 learners in our sample out-performed those in the 2013 national study. They were also less likely to score zero on the subtasks. While the percentage of learners meeting each benchmark is higher for our sample than the national sample, they are very far from the proposed five year targets.

**Table 10.** Comparison of EGMA scores to national averages and proposed benchmarks

	Addition and subtraction level 2	Missing number	Word problems
<b>Average scores:</b>			
<b>2013 National</b>	16.6%	26.2%	40.0%
<b>TFET baseline</b>	25.3%	39.0%	45.0%
<b>Proposed benchmark</b>	80%	70%	80%
<b>Percentage meeting benchmarks:</b>			
<b>Proposed 5 year target for meeting benchmark</b>	16%	16%	55%
<b>2013 % meeting benchmark</b>	3.1%	3.3%	11.0%
<b>TFET % meeting benchmark</b>	7.9%	4.2%	10.0%
<b>Percentage of zero scores:</b>			
<b>Proposed 5 year target</b>	23.0%	4.0%	4.0%
<b>2013 National</b>	46.0%	7.6%	9.0%
<b>TFET baseline</b>	36.1%	0.8%	2.2%

## 5.9 School and Home Environment

Understanding the learners’ school and home environment is important for both statistical and contextual reasons. Firstly, variables describing school infrastructure, household assets and learning support in the home will be useful for matching learners in treatment and comparison schools. These variables also tend to explain some of the differences in learner performance. As such, they are useful in absorbing variation and help in isolating the effect of the TFET program. Secondly, these variables help us describe the conditions and context in which these teachers are teaching. This is important for understanding what hinders or facilitates uptake and success of the TFET principles in schools.

During school visits to conduct the learner assessments and teacher interviews, enumerators also carried out a school inventory, characterizing the school infrastructure, sanitary conditions, and learning environment. Table 11 summarizes the data collected from the inventory, describing the proportion of schools in the sample with different characteristics. For example, 80% of the schools in the sample are located in buildings that were purposely built as schools (as opposed to being located in a home or an unenclosed structure), 78.8% are connected to the national electricity grid, and 61.3% have running water on-site.

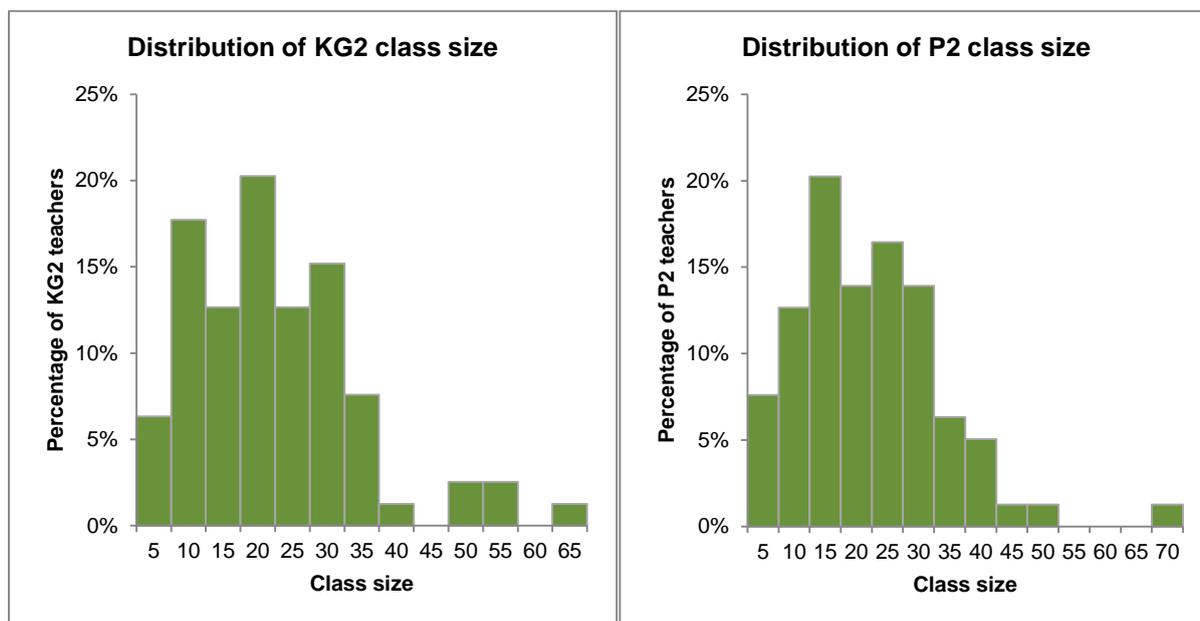
Nearly all schools (98.8%) use traditional seating arrangements with learners seated at desks or tables in rows, while relatively few (30%) appear to have sufficient space for the class to come together with children seated on the floor. This, combined with the proportion of schools where there is sufficient space for learning and play, a daily program on display, and sufficient outdoor space for play suggests that there may still be room for improvement in creating environments conducive to learner-centered learning.

**Table 11.** School Inventory Summary Statistics

	Percent of schools
Main Building Materials of P2 Classroom = Brick	87.5%
P2 Classroom is an Enclosed Space	92.5%
P2 Classroom Has Windows	82.5%
Building is a purpose-built school	80.0%
Connected to National Electricity Grid	78.8%
Tapped Water on-site	61.3%
Learners Wash Hands from Inside/Outside Tap or Hand Pump	48.8%
Learners Have No Option for Hand washing	23.8%
Soap Available for Hand washing	30.0%
Flush Toilet Available to Learners	43.8%
Toilets are Clean and Hygienic	63.6%
No flies/mosquitos in the toilet	71.3%
Site fenced with lockable gate	72.5%
Enough indoor space for learning and play	77.5%
Open space where all the children can sit on the floor	30.0%
Children seated at desks/tables in rows	98.8%
Daily program on display	63.8%
Outdoor space spacious enough for play	53.8%

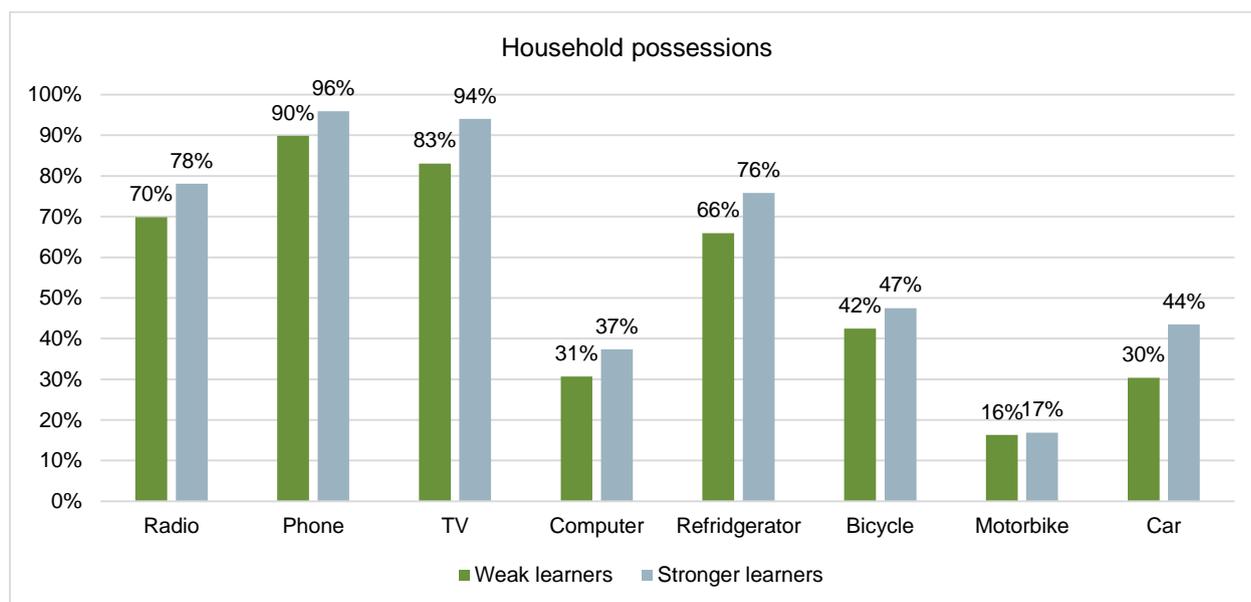
Figure 26 shows the distribution of class size by grade. The average class size for both KG2 and P2 is around 20 learners. While there are some large classes, in the majority of cases the number of learners is small enough for teachers to reasonably incorporate the activity-based and learner-centered methods covered in the TFET training.

**Figure 26.** Class size by grade



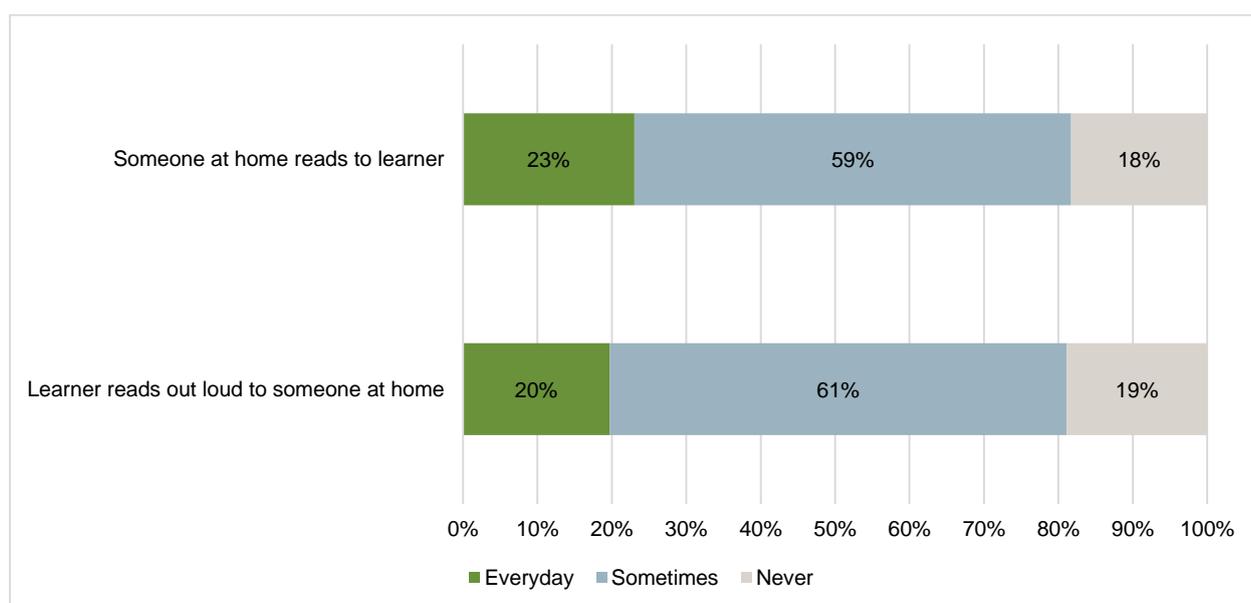
We identify a group of P2 learners with very weak literacy and numeracy outcomes as those who score zero on both the English reading comprehension and the more advanced addition subtasks. Overall, 25% of P2 learners fall into this category. During the P2 assessment, learners were asked about their home environment. In Figure 27 we characterize the ownership of various standard household items, categorizing learners according to those with very weak literacy and numeracy outcomes and those not scoring zero on both the English comprehension and addition 2 subtasks. Weaker learners clearly have lower levels of ownership of nearly all items, suggesting a correlation between wealth and academic performance.

**Figure 27.** Household possessions



We also asked P2 learners about the support they receive at home. Figure 28 shows how often learners read out loud to someone at home and how often a family member reads to the learner at home. In general, only a minority of learners appear to engage in reading activities at home with household members on a daily basis, with just 20% and 23% saying they read out loud to or are read to by a family member, respectively. The vast majority of learners (87%) report that someone at home helps them with their homework. Just over three-quarters (77%) of learners have access to at least some reading materials in the home.

**Figure 28.** Reading support at home



## 6. Final Comments

We present in this report the findings from the baseline collected for the impact evaluation of the TFET Project that includes KG2 and P2 learner assessments, learner interviews, teacher and head teacher interviews, attendance, and a school inventory.

In a context where teacher-dominated instruction predominates, the TFET program promotes a shift to child-centered, activity based learning. To understand the extent to which teachers' beliefs align with such learner-centered approaches, we probed their level of agreement with 20 statements about pedagogical practices. While teachers tend to respond positively to statements that align with learner-centered practice, there was considerable variation across statements and between teachers, allowing us to discriminate between those who are more and less learner centered. Teachers who express more learner-centered beliefs tend to report higher motivation and job satisfaction.

The use of emotional and physical punishment seems common among KG2 and P2 teachers and learners report fear of going to school because of discipline practices. Caning in particular is reported by a majority of teachers as an effective way to create discipline. This clearly indicates the importance of the TFET module that aims to train teachers in the use of positive discipline in order to promote the establishment of a positive and supportive school climate for learning.

The baseline results also show that student absenteeism is an important issue which gets reinforced by teacher's absenteeism. Almost 50% of P2 pupils interviewed reported missing at least one day of school during the school prior week, and attendance records indicate an average of 31 and 35 percent absenteeism in P2 and KG2 in the previous school term.

Our data indicates that strong gender inequitable attitudes prevail in school. Learners hold stronger inequitable gender attitudes than teachers, and males stronger than females.

In the literacy assessment of P2 learners, we find that average oral reading fluency in English of 33 words per minute is below the proposed benchmark for Ghana of 45 words per minute<sup>20</sup>, however it is in the comprehension subtasks (both listening and reading comprehension) that results are more concerning as they show very low levels of understanding.

The analyses of the math assessments show that, in general, learners perform better in numeracy subtasks that are procedural and repetitive than in conceptual ones. KG2 learners show difficulties matching and sorting and putting a simple puzzle together which are activities related to spatial

---

<sup>20</sup> USAID 2014. Proposing Benchmarks for Early Grade Reading and Mathematics in Ghana. [https://pdf.usaid.gov/pdf\\_docs/PA00KS7N.pdf](https://pdf.usaid.gov/pdf_docs/PA00KS7N.pdf)

reasoning and are linked to early math skills. P2 learners tend to solve problems and operations using inefficient strategies such as tick marks or their fingers that preclude them from solving additions or subtractions that are not very simple.

These findings may relate to classroom practices that do not tend to encourage children to explore, experiment and solve problems independently or through working with each other. Some opinions expressed by learners and teachers seem to support this argument. Over half (53%) of our P2 learners indicate that their “teacher does not think that is okay to make mistakes” and teachers are not too enthusiastic about the idea that “pupils should be allowed to think of solutions to practical problems by themselves before the teacher shows the solution” with only 26% strongly agreeing with this statement.

The key focus of the evaluation of the TFET program is to examine the extent to which teachers are able to make the paradigm shift from “teacher-centered” to “learner-centered” pedagogy, and importantly whether this translates to better development and learning outcomes for their pupils. This baseline report has provided a rich picture of a challenging context with multiple opportunities for the program to potentially impact practices and outcomes. Learner outcomes are poor, particularly for tasks with higher cognitive demand such as problem solving and comprehension. Absenteeism is high, corporal punishment widely considered effective, and both learners and teachers exhibit substantial gender bias.

The in-depth qualitative inquiries at midline will provide an opportunity to probe trained teachers’ understanding of the key TFET pedagogical principles, their ability to implement the techniques learnt and the challenges faced in doing so. We will also be able to explore the extent to which prior beliefs on learner-centered approaches hinder or facilitate the uptake of these key principles. Finally, baseline and midline findings and lessons will inform our endline data collection work and analyses.