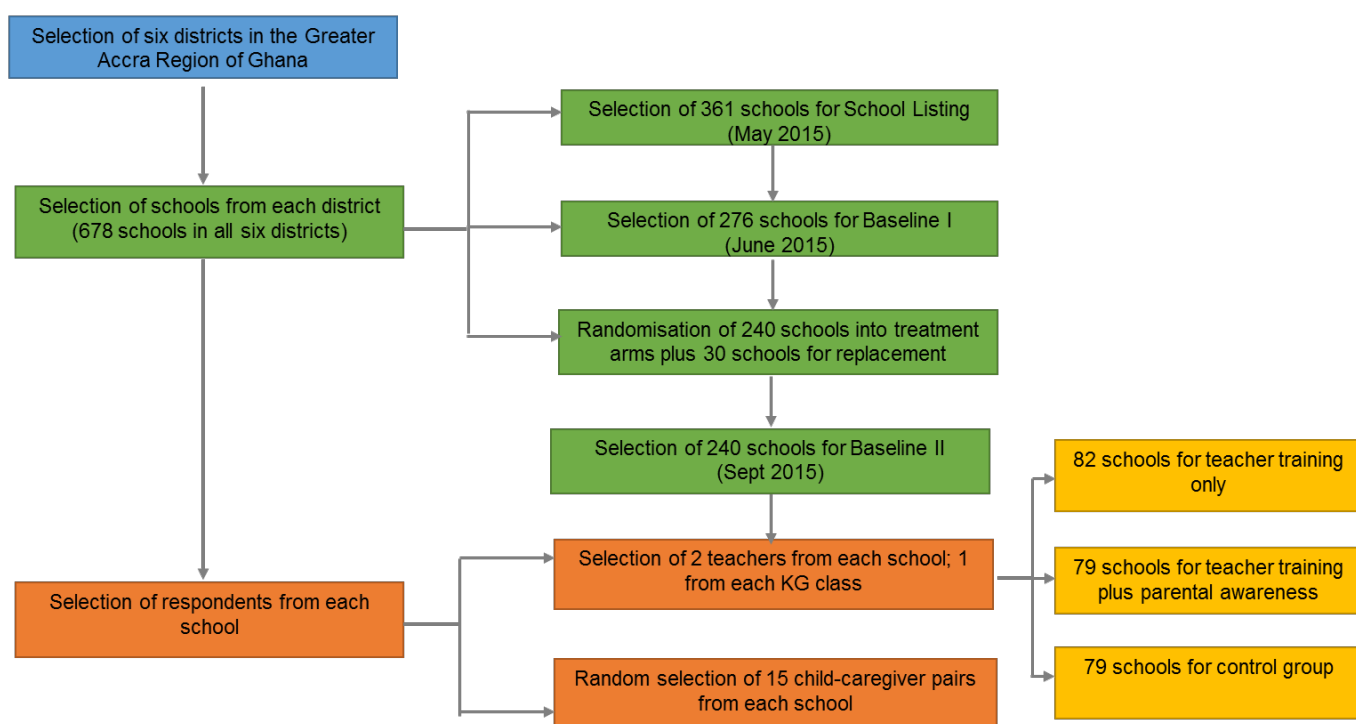


Quality for Preschool in Ghana Study (QP4G) 2015-2017

Note on Sampling Design and Response Rates

27th April 2016

Figure 1 Overview of sampling procedure



Research Design

This impact evaluation applies a cluster-randomized design. Eligible schools were randomly selected to participate in the study. The eligible population was schools with KG 1 and KG 2 classrooms (the two years of universal preprimary education) in six districts in the Greater Accra Region. In these six districts we have sampled 240 schools; 108 public schools and 132 private schools in total.

The unit of randomization for this randomized control trial (RCT) is *schools*, whereby eligible schools (stratified by public and private sector schools) are randomly assigned to: (1) in-service teacher-training program only; (2) in-service teacher-training program plus parental awareness program; or (3) control (current standard operating) condition.

District and School Selection

The sampling frame for this study was based on data in the Education Management Information System (EMIS) from the Ghana Education Service. This data was verified in a 'school listing exercise' conducted in May 2015.

Sample selection was done in multiple stages as shown in Figure 1. The first stage involved purposive selection of six districts within the region based on two criteria: (a) most

disadvantaged (using UNICEF's District League Table scores, out of sixteen total districts); and (b) close proximity to Accra Metropolitan for travel for the training of the KG teachers. The six selected municipals were La Nkwantanang-Madina Municipal, Ga Central Municipal, Ledzokuku-Krowor Municipal, Adentan Municipal, Ga South Municipal and Ga East Municipal.

The second stage involved the selection of public and private schools from each of the selected districts in the Accra region. We found 678 public and private schools (schools with kindergarten) in the EMIS database. Of these 361 schools were sampled randomly (stratified by district and school type) for the school listing exercise, done in May 2015. This was made up of 118 public schools and 243 private schools.

The sampling method used for the school listing exercise was based on two approaches depending on the type of school. For the public schools, the full universe of public schools (i.e., 118) were included in the school listing exercise. However, private schools were randomly sampled using probability proportional to the size of the private schools in each district. Specifically, the private schools were sampled in each district proportionate to the total number of district private schools relative to the total number of private schools. In so doing, one school from the Ga South Municipal was removed and added to Ga Central so that all districts have a number of private schools divisible by three. This approach yielded 122 private schools. Additionally, 20 private schools were randomly selected from each of the districts (i.e., based on the remaining list of private schools in each district following from the first selection) to serve as replacement lists. The replacement list was necessary given the potential refusals from the private schools. There were no replacement lists for the public schools since all public schools would automatically qualify for participation.

The third stage involved selecting the final sample for the evaluation using the sampling frame obtained through the listing exercise. A total of 240 schools were randomly selected, distributed by district and sector as shown in Table 1. Schools were randomized into treatment groups after the first round of baseline data collection was completed.

Table 1. Distribution of Evaluation Districts

District	Private Schools	Public Schools	Total
La Nkwantanang-Madina	10	15	25
Ga Central	25	8	33
Ledzokuku-Krowor	33	21	54
Adenta	22	10	32
Ga East	20	18	38
Ga South	22	36	58
Total	132	108	240

Data Collection

Baseline data collection occurred in two rounds, and data was collected from multiple sources and respondents. Baseline I consisted of interviews with school head teachers and school proprietors (for private schools) and was conducted in June 2015. This was done before the summer holidays and the start of the academic year (2015-2016). It was expected that school ownership and senior management would not change between academic years and that this would give us a more clarity about the structure and composition of schools in

our sample. Baseline II was done in Sept-Nov 2015 and consisted of collecting the following data: (a) direct assessments of children's school readiness at school entry, (b) surveys of teacher well-being and demographics, (c) video recordings for classroom observations of teachers, and (d) caregiver surveys.

The video recordings conducted during Baseline II were subsequently analysed/coded by a team of trained researchers to create a dataset for classroom observations (at the KG teacher level). **Due to issues of confidentiality, these video recordings are not included in the submission of data to the Microdata catalogue.**

The sampling method and processes used for the Baseline I were based on the procedure used in the school listing exercise. However, schools that refused to participate in and/or were not covered during the school listing exercise were excluded: 276 schools were selected for Baseline I data collection. An additional 36 schools were selected because we expected some refusals.

240 schools were eventually randomized into two treatment arms (i.e., teacher training only and teacher training plus a parental awareness intervention) and a control group. An additional 30 schools were selected to be on a "reserve list" and were used to replace schools in the original list that declined to participate in the survey. Out of the initial 240 schools randomized, 27 were replaced because of refusals and lack of KG classes in the schools. School replacement was done on a case-by-case basis. The distribution of the final 240 schools is shown in Table 2.

Table 2 Distribution of Sampled Schools

Treatment Type	School Type		Total
	Public	Private	
Teacher training only	36	46	82
Teacher training plus parental awareness intervention	37	42	79
Control	35	44	79
Total	108	132	240

Source: Sampled list of 240 Schools

The survey respondents were sampled using different sampling techniques:

- KG teachers:** The research team sampled two KG teachers from each school; one from KG1 and KG2. KG teachers were sampled using purposive sampling method. In schools where there were more than two KG classes, the KG teachers from the "A" stream were selected. For the treatment schools, all KG teachers were invited to participate in the teacher training program.
- KG child-caregiver pair:** The research team sampled KG children and their respective caregivers using simple random sampling method. Fifteen KG children-caregivers pair were sampled from each school. For schools with less than 15 KG children (8 from KG1, 7 from KG2 where possible), all KG children were included in the survey. KG children were selected from the same class as the selected KG teacher. The survey team used the class register to randomly select KG children who were present on the day of the school visit. Sampling was not stratified by

gender or age. The caregivers of these selected child respondents were invited to participate in the survey.

The research team sought informed consent from the school head teacher, caregivers, as well as child respondents.

Survey coverage and response rates

Baseline I

Out of the 276 schools that were selected for the Baseline I, 269 schools were surveyed (remember that potential replacement schools were also surveyed during Baseline I). This represents a response rate of 97%. The response rate for Baseline I is shown in Table 1. It must, however, be emphasized that there were incomplete surveys in some of the schools, especially for the private schools. Incomplete surveys mean that only one of the surveys (instead of the two) was administered.

Table 3 Coverage Rate for Private and Public Schools

Categories	Target	Actual	% of Target
Proprietors	150	135	90%
Head teachers	276	265	96%
Public schools	115	113	98%
Private schools (proprietor/head)	161	156	97%
Total number of schools surveyed	276	269	97%

The following explains the response rate for the Baseline I:

- The slightly lower response rate was because the proprietor and head teacher surveys coincided with the Basic Education Certificate Examination (BECE). This led to the postponement of appointments to first week in July by some proprietors/head teachers. The major lesson learned is that future surveys should be conducted at least two weeks to the start of major school activities such as the BECE.
- The coverage rate for the public schools was 98% due to the non-availability of the head teachers of two schools. The head teachers had expressed interest in the study but had invited us to conduct the survey on a later date.
- The coverage rate for private schools was 97% due to the non-availability of the respondents. Some proprietors had travelled outside the country and could not even be contacted by their heads to decide whether to participate or not. As a result, 5 private schools were not surveyed.
- 12 selected schools refused to participate in the study due to reasons of bad experience from research institutions/GES; perceived no direct benefit of the study to

the school or themselves; and the possible interference of the survey with school activities.

- 16 schools were surveyed as a replacement for the 5 non-available schools/respondents and 12 refused schools.

The key challenge experienced during data collection was the difficulties in meeting the school heads/proprietors due to the Basic Education Certificate Examination. Some school heads reported that they had to prepare their students before and during the examination period. No cases of administrative/logistical challenges were reported.

Baseline II

All the surveys/assessment [with the exception of the Caregiver Survey] reported more than 90% response rate.

- The response rate for the KG Teacher Survey was 92.5%. Reasons for non-response were the refusal of some teachers, transfer of KG teachers without replacement [eight teachers were transferred without replacement], lack of KG teachers and the combined nature of KG1 and KG2 classes (Table 4). One of the teachers that refused belonged to teacher training and parental awareness treatment arm. The others belonged to the control group. Twenty-one additional teachers were surveyed but were excluded from the 444 completed surveys. These comprised of eight teachers who were transferred from their respective schools and 13 teachers from [nine schools], which refused to continue with the QP4G Survey.

Table 4 Response Rates for KG Teacher Survey

Survey Outcome	Target	Actual	% of Target
Completed	480	443	92.5
Partially completed	-	1	
Refused	-	2	
Sample loss due to combined KG class and lack of teachers*	-	34	
Total	480	480	100
<i>*Note: 20 schools were combined. An additional 14 schools were 'temporarily' combined due to lack of teachers</i>			

Source: SurveyCTO data.

- The response rate for the Child Assessment was 95.4%. Non-responses were due to refusals from children and/or caregivers, insufficient KG children, and absenteeism (Table 5). Ten caregivers from a control school refused to allow their children to participate in the assessment. The other two refusals were from KG children from a control and a treatment school. It is interesting to note that some children who refused initially later agreed to be assessed. Children mostly refused initially because of fear and discomfort with the assessors. However, upon seeing their peers go through the assessment successful, they were motivated to do it.

Table 5 Response Rates for Child Assessment

Survey Outcome	Target	Actual	% of Target
Completed	3600	3425	95.4
Partially completed	-	10	
Refused (from children and/or caregivers)	-	12	
Sample loss due to insufficient KG children and absenteeism	-	153	
Total	3600	3600	

Source: SurveyCTO data.

- c. The response rate for the Caregiver Survey was 60.0% (Table 6). Reasons for non-completion were refusals from caregivers to participate in the study, sample loss due to incorrect and unreachable numbers, and lack of or incomplete information on caregivers. Out of the 12 refusals, four each of the caregivers were from each of the three treatment arms. Note that ten caregivers from a control school refused to participate in the study and consent to their children's participation in the study. These caregivers were not part of the refusals in Table 8, as they did not provide any name or contact numbers.

Table 6 Response Rates for Caregiver Survey

Survey Outcome	Target	Actual	% of Target
Completed	3600	2,144	60.0
Partially completed	-	15	
Refused	-	12	
Sample loss due to either out of scope ¹ and/or incorrect/unreachable contacts ²	-	470	
Caregivers with no information		422	
Caregivers with names but no contact numbers		537	
Total	3600	3600	

Source: Varied. ¹ Out of scope: The supposed caregiver contacted was not related to the sampled KG child in any way. ² Sample loss: The caregivers' numbers were incorrect and/or unreachable.

The response rates for the different components of the Baseline II Study are largely dependent on five factors. These are the willingness of schools to participate, number of KG children admitted to the school [as of the time of the data collection], available information on caregivers for selected children, available spaces for the KG unit, and the number of KG teachers in the school. The following presents the research team's observations.

- a. Of the 240 schools randomised into treatment and control groups, 23 schools were replaced with reserve schools due to refusals and lack of KG classes (in addition to the replacements from Baseline I). Most of the schools in this list refused because of lack of interest in the survey and the perceived distrust for research organizations. Some other schools were removed because they do not have a KG unit. Table 7 presents the distribution of the schools that were removed from the original list.

Table 7 Distribution of Removed Schools from the Original List

Treatment Type	School Type		Total
	Public	Private	
Teacher training only	2	6	8

Teacher training plus parental awareness intervention	2	5	7
Control	3	5	8
Total	7	16	23

- b. Twenty schools were identified as having combined KG classes. A combined KG class depicts a situation in which a school combines both KG1 and KG2 children to form a distinct KG class [and mostly managed by one teacher and/or an attendant] due to lack of teachers to manage the separate classes and/or lack of classroom facilities to accommodate the two classes. The survey team surveyed and/or videotaped only one teacher in the combined classes. These KG teachers can be identified in the dataset with the teacher code 31. This situation had led to KG teacher sample loss by 20 teachers. Notwithstanding, the combined status of such classes/schools could change any time if the schools hire additional teachers or acquire additional classroom facilities. Similarly, due to lack of or inadequate KG teachers in 14 schools, 14 KG teachers handled both KG1 and KG2 classes. These teachers were treated as a special case of combined KG. They can be identified in the dataset with the teacher code 41. The sample loss from this was 14 KG teachers.
- c. Less than 15 KG children per school were assessed in 30 schools. The average number of KG children assessed in those schools was ten, with a range of 4 to 14 children. The reasons were limited number of KG children in the schools, refusal of some caregivers to consent to their children's participation, and absenteeism. The total sample loss resulting from this was 165 child-caregiver pairs.
- d. We obtained caregiver information from 226 schools, representing 94.2% of the 240 schools. Only 89 schools had information on all 15 caregivers. The average number of caregivers reached per school was 10, with a range of 1 to 14. Fourteen schools had no caregiver information because the schools and/or caregivers refused to provide the information or could not provide the information to the research team. Only two of such schools were in the two treatment arms. In terms of actual caregivers, 3,178 caregivers' information was obtained from the schools. In Table 8 we use 3,600 as the target based on the research design. Note that the number of caregivers listed is dependent on the total number of KG children assessed, as the sample comprised child-caregiver pairs. Out of the 3,435 children assessed, we had complete information [with or without contacts] on 73.4% caregivers.

Table 8 Outcome of Caregiver Listing

Caregiver Listing	Survey Target	Actual	% of Target
Schools	240	226	94.2
Caregivers			
Total caregivers listed	3600	3178	88.3
Caregivers with contact numbers	3600	2641	73.4
Caregivers without contact numbers	-	537	
Total caregivers not listed	-	422	

Source: Tracking sheet, SurveyCTO sampling forms and Sampled KG children Form