Developing and Testing Supply- and Demand-side Interventions to Improve Kindergarten Educational Quality in Ghana

SIEF Final Report

Original submission: October 2017
Revised submission: December 2017

Prepared by:

Sharon Wolf[a], Edward Tsinigo[b], Jere R. Behrman[c], J. Lawrence Aber[d], Alix Bonarget[b]

[a] Graduate School of Education, University of Pennsylvania

[b] Innovations for Poverty Action

[c] Departments of Economics and Sociology, University of Pennsylvania

[d] Global TIES for Children, Institute for Human Development and Social Change, New York University

Acknowledgements

The project represented in this report reflects contributions from many organizations and individuals. We would like to thank the dedicated staff and thought partners at Innovations from Poverty Action, including Loic Watine, Bridget Gyamfi, Amma Aboagye, Kwabena Asamonah, Henry Atimone, Maham Farhat, Shams Muzaffar, and Renaud Comba, as well as our talented data collection supervisors and enumerators. We also thank the Ghana Education Service national and district offices for their support of this project, including their implementation of key elements of the programs. The National Nursery Teacher Training Center was also a key partner in the development and implementation of the teacher training program, and Sabre Charitable Trusts helped develop the teacher training curriculum. Finally, we would like to thank the QP4G Steering Committee, which provided important feedback and guidance throughout the development of this project. We also thank the UBS Optimus Foundation and the World Bank Early Learning Partnership for additional direct support for this project. In addition, we thank NYU Abu Dhabi Research Institute for core support to the Global TIES for Children Center, which provides intramural support to the fourth author for research activities that directly contributed to the project. Most of all, we thank the dedicated teachers, children, and families who participated in the study and made this all possible.

Executive summary

The Quality Preschool for Ghana (QP4G) project aims to build capacity and support for the implementation of the 2004 kindergarten (KG) curriculum and to enhance the quality of KG education in Ghana. The goal of the project is to develop and evaluate an affordable and scalable model of transformational teacher training to provide high-quality early childhood education (ECE) services to children and to test the added benefits of engaging parents via an awareness campaign designed to align parental expectations with these practices. The programs were designed to improve classroom quality and the development of children's school readiness skills in the Greater Accra Region of Ghana. *Preschool* in this study refers to the 2 years of pre-primary education in Ghana called *kindergarten*.

The training for kindergarten teachers and head teachers began with a five-day course at the start of the school year, followed by two refresher trainings (2 days at the start of term 2, 1 day at the start of term 3). The program offered experiential training for teachers and included ongoing monitoring and support from district education officers. The parental-awareness intervention consisted of three meetings (one per term) held during PTA meetings. At each meeting, district coordinators screened a short video followed by discussion. The themes were (1) play-based learning, (2) parents' role in child learning, and (3) encouraging parent-teacher and parent-school communication.

Researchers randomly assigned 240 schools in six of the most disadvantaged districts of the Greater Accra Region to one of three conditions: teacher training (TT; 82 schools), teacher training plus parental-awareness meetings (TTPA; 79 schools), and a control condition (79 schools). In each school, 15 kindergarten pupils were randomly sampled. Data were collected through interviewer-administered in-person interviews with teachers, phone interviews with caregivers, classroom observations, and direct child assessments.

The implementation and first-year evaluation of the QP4G intervention occurred between September 2015 and June 2016. A second follow-up was conducted the following academic year (June 2017) in order to understand the lasting impacts of the program on children and on teaching quality.

Results for the midline assessment were reported in the July 2017 midline report. Since the report, a new analytic strategy was undertaken to account for sample attrition. Thus, we now report both the updated midline impact estimates as well as endline impact estimates. At midline, moderate-sized impacts were found on some dimensions of professional well-being

(reduced burnout in TT and TTPA, reduced turnover in TT), classroom quality (increased emotional support/behavior management in TT and TTPA, support for student expression in TT), and small impacts on multiple domains of children's school readiness (in TT), including early literacy, early numeracy, and social-emotional development. One year later, impacts on some elements of teaching practice in both TT and TTPA persisted while others reversed, and impacts on children's social-emotional skills in the TT condition persisted.

The findings provide evidence that the QP4G curriculum is an effective and potentially scalable way to train teachers on the KG-specific pedagogy that is specified in the national curriculum, and that impacts on some outcomes are sustained the year after program implementation ended. The results also suggest that the parental-awareness meetings were not an effective way to engage parents in their child's education, and alternative approaches to engage parents need to be explored.

Table of Contents

List of Tables

List of Figures

Abbreviations

COE	College of Education
CRDD	Curriculum Research and Development Division
ECCD	Early Childhood Care and Development
ECD	Early Childhood Development
ECDCAG	Early Childhood Development Care Association of Ghana
ECE	Early Childhood Education
EMIS	Educational Management Information System
GES	Ghana Education Service
ICERDA	International Conference on Education Research for Development in Africa
IDELA	International Development Early Learning Assessment
IPA	Innovations for Poverty Action
IRB	Institutional Review Board
KG	Kindergarten
MOE	Ministry of Education
NCTE	National Council for Tertiary Education
NNTTC	National Nursery Teacher Training Center
PRINCOF	National Conference of Principals of Colleges of Education
QP4G	Quality Preschool for Ghana
RCT	Randomized Control Trial
TED	Teacher Education Division
TIPPS	Teacher Instructional Practices and Processes System
TT	Teacher Training
TTPA	Teacher Training and Parent Awareness
UNICEF	United Nations International Children's Fund

Introduction

The adoption of the 2004 National Early Childhood Care and Development (ECCD) Policy highlighted access to quality kindergarten education as central to improving early childhood development and learning, and a promising way to prevent developmental delays and foster early learning despite adversity. In 2007, Ghana's government became one of the first in the sub-Saharan African region to include 2 years of Kindergarten (KG) education as part of the basic education system.

The Ghana KG Implementation plan, which was written in 2012, adopted a phased approach to scaling-up quality kindergarten education in Ghana. In this plan, the first priority for the government is to train teachers in the KG curriculum and KG-specific pedagogy. However, the plan has not been effectively implemented, in large part due to the lack of affordable options to effectively improve teaching practices at scale. This is the need that the Quality Preschool for Ghana (QP4G) project aimed to address: testing a scalable model to improve KG teaching practices in Ghana.

The QP4G study, supported by the UBS Optimus Foundation and the World Bank SIEF and ELP, was launched in 2015. It was designed as a randomized evaluation to assess the impacts of an eight-day kindergarten teacher training program, coupled with monitoring and coaching, on the implementation of the KG-specific pedagogy, classroom quality and children's learning and development. A second treatment arm assessed the added value of a parental-awareness program regarding quality preschool education to the teacher training. The study also tested whether the program would have different impacts in the public and private sectors. A total of 240 schools (108 public and 132 private) were randomly selected in six of the most disadvantaged districts[1] of the Greater Accra Region to be part of the evaluation.

A baseline report was submitted in February 2016, and a second report submitted in March 2017 summarizing the impact findings from midline, at the end of the implementation academic year. This report presents a re-analysis of the midline results, as well as the final results of the evaluation based on the endline data collection[2]. Thus, data collection for the impact results was done between June 2015 and June 2017. The first Section provides the project context and describes the interventions. Section 2 details the empirical strategy adopted for the evaluation.

[1] For consistency with previous reports, districts as used in this report refer to Municipals.
[2] With support from the World Bank's Early Learning Partnership, we were able to follow the full sample of teachers and children from the baseline and midline samples.

The third and fourth Sections present the findings and discussion of the evaluation, respectively. The final Section focuses on the implications of the findings and the next steps.

Section I: Program Context and Interventions

1.1 Background of kindergarten education in Ghana

1.1.1 Governance, policy and teacher training

The Ministry of Education (MOE) is responsible for initiating and formulating educational policies, developing and revising the national curriculum, overseeing teacher professional certification and implementing national educational policy through its agencies, such as the Ghana Education Service (GES). GES is in charge of pre-tertiary education: it coordinates early childhood education (ECE) activities in school through its Basic Education Division and liaises with Development partners in promoting ECE development in Ghana.

Ghana is one of the few African countries to have developed a national early childhood development policy aimed to promote the development of children from birth to 8 years old and to coordinate stakeholder activities in the sector. The policy, promulgated in 2004, establishes institutional roles and responsibilities for public and private partners and develops an implementation strategy. Among other things, the National Early Childhood Care and Development (ECCD) policy highlight access to quality KG education as central to improving early childhood development and learning, and a promising way to prevent developmental delays and foster early learning despite adversity. In 2007, Ghana became the first sub-Saharan African country to expand the Free Universal and Compulsory Basic Education (FCUBE) to kindergarten, stating that all children are to receive two years compulsory ECE at the ages of four (KG1) and five (KG2) before entering primary school.

The main providers of pre-service kindergarten teacher training are the Colleges of Education (COE), seven of them offering diplomas in basic education focusing on kindergarten. The University of Education, Winneba also offers degrees in early childhood education and the University of Cape Coast proposes specific courses on early childhood education as part of their degree in education. In-service training is provided by the National Nursery Teacher Training Center (NNTTC), a center that is partially funded by GES, and which offers an eight-week in-service training program. There are many NGOs that also offer training throughout the country.

1.1.2 Access to early childhood education

Following the launch of the National ECCD Policy in 2004, the preschool sector in Ghana expanded rapidly. From 2004 to 2010, the number of kindergartens in Ghana doubled, this increase is mostly accounted for in the public sector (Educational Management Information System [EMIS], 2015; MOE, 2016.). However, recent years have witnessed a shift in this

dynamic: while kindergartens in the public sector increased by 6.6% between 2010 and 2015, the number of private sector kindergarten increased by 42.7% over the same period (EMIS, 2015). These increases are largely influenced by the greater involvement of private entities such as parents, churches, and enterprises, in funding and managing schools.

Ghana has consistently displayed higher enrollment levels in basic school than the regional average. In 2015/2016, net enrollment in kindergarten was 80%, slightly lower than previous levels due to rapid growth in the population of 4 to 5 years old (MOE, 2016.)

1.1.3 Quality of early childhood education in Ghana

The ECE sector in Ghana is beset with low staff qualifications and in some cases untrained staff, especially in preschool. The 2012 Government Kindergarten Situational Report concluded that the 2004 curriculum established is sound, but that teacher behavior has not yet adapted to reflect the new pedagogy. This is partly because only half of the teachers in kindergarten received formal training[3] (EMIS, 2015) and also that most teacher training institutions are yet to integrate and promote adapted teaching practices. Consistent with these observations, a 2013 exploratory study by Innovations for Poverty Action (IPA) in Accra revealed that the quality of classroom instruction in preschools was generally low and developmentally inappropriate across the public and private sectors (Bidwell et al., 2014). In addition, the study revealed that parents subjectively assess preschool quality based on the existence of classroom materials and school infrastructure.

These findings suggest that raising the quality of early childhood education requires a two-pronged approach to train preschool educators, ensuring age-appropriate developmental instruction, and engaging parents and communities in children's learning. In addition, a major concern for policymakers remains inequalities in early childhood education quality and lack of standards for assessing early childhood education quality across the public and private sectors. Understanding differences in the public and private sectors in Ghana ECE and their implications for system-level change is therefore critical.

1.2 Intervention and Research Design

1.2.1 Study design

The schools were randomly assigned to each of the three treatment arms: (1) teacher training (TT; 82 schools), (2) teacher training plus parental-awareness program (TTPA; 79 schools), and

[3] This proportion was 61.7% in public schools and 5.1% in private schools in 2014/2015.

(3) control group (79 schools). In addition, treatment schools were then randomly assigned to receive weekly text messages for teachers (N = 80 schools) and picture-based paper flyers for parents (N = 40 schools). The research design is shown in Figure 1~~Figure 1~~.

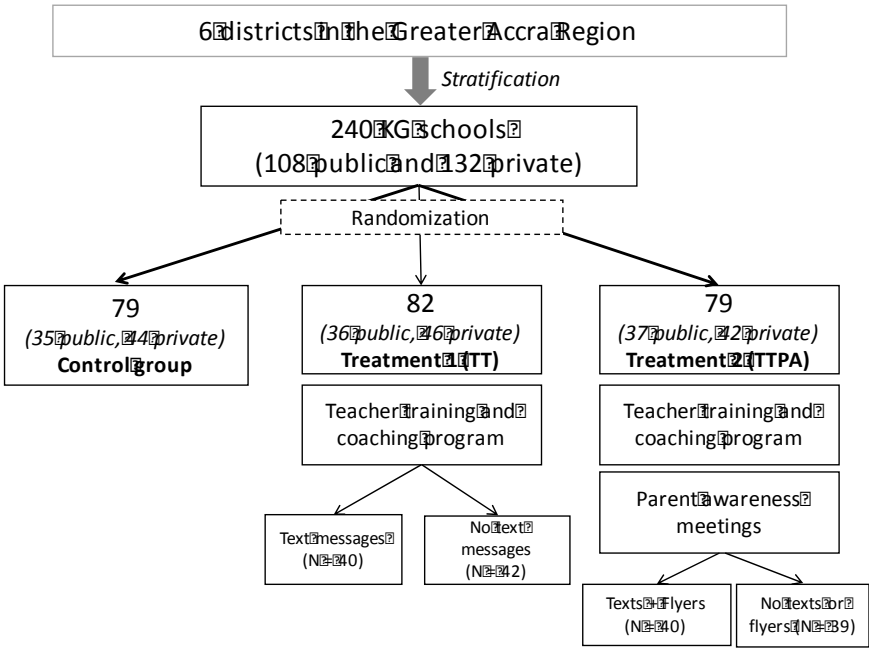


Figure 1. Treatment arms

1.2.2 Description of interventions

The project tests two complementary interventions, targeting both supply-side (in-service teacher training) and demand-side (parental-awareness program) approaches to improving educational quality. While training alone might change teacher behavior, this change might be curtailed by parental requests that emphasize alternative and less effective instructional methodologies. Therefore, to ensure teacher behavioral changes can positively impact child learning outcomes, both teacher training, and parental orientation may be needed to ensure alignment of mindsets and actions on appropriate instructional quality. The control group received no intervention and served as a comparison to the intervention groups.

1.2.2.1 In-service teacher training

The first intervention evaluated is a scalable, cost-effective, in-service teacher training program for Kindergarten teachers in Ghana, aimed to improve the quality of Kindergarten education through the supply-side of education. The training curriculum was eight days total and focused on (a) How Children Learn, (b) Classroom Management, (c) Literacy, (d) Numeracy, and (e) Planning and Assessment. The teacher training was delivered in three stages: (i) an initial five-

day training during the first term, for teachers, head teachers and district staff in participating districts, (ii) a two-day refresher training in the second term that reflects topics identified through information collected from monitoring teachers (to provide remediation targeted to the needs of the teachers) and (iii) a one-day final refresher training in the third term to reflect on progress, identify opportunities for continued growth in the coming year and develop final resources for the term.

NNTTC, the primary training center for ECE certification in Accra, was in charge of delivering the training. NNTTC recently developed a 5-day in-service training module in collaboration with UNICEF, which this project built on, enhancing several components and increasing the time for experiential training. Training materials were developed by consultants from the Sabre Trust in consultation with various government, civil society and development partners. IPA worked with the NNTTC to review the training curriculum, ensure an understanding of the information involved and provide support where needed in the development of posters and visual aids for the training. The training was facilitated by the training staff at the NNTTC. The training team was made up of 9-course tutors including the Officer-In-Charge. Eight of the tutors are staff of NNTTC and 1 visiting tutor was brought in from the International Child Resource Institute (ICRI). The participants were randomly assigned to 3 different classrooms with 3 tutors per class to deliver the content of the training. The facilitators were assigned based on their area of expertise.

The NNTTC also trained existing District Support Focal Person (DSFP) to provide post-training coaching visits and conduct monitoring. The DSFPs were tasked to conduct monitoring visits to the participating schools. The DFSPs are a team based at the district education office in charge of overseeing the activities of the project or whose role relates to an aspect of the QP4G project implementation. They DSFPs are district staff with positions ranging from Deputy District Directors in charge of Supervision, District Early Childhood Education Coordinator, Basic Schools Coordinator and Circuit Supervisors. Some other DSFP occupied the position of District PRO and the Language Coordinator. The initial coaching visits and monitoring training focused on the purpose, expectations, visits and post-visits, reports, understanding expectations of teachers, lesson cycle, pedagogical approaches, and content. They were also oriented on providing feedback to the teachers, using a monitoring checklist that was provided. The content of the subsequent (refresher) training was influenced by the feedback and comments gathered from the DSFPs School Monitoring Report. The refresher training focused on a) the establishment of a norm for effective rating of teachers during DSFP monitoring visits; b)

strengthening DSFPs understanding of the seven steps to take when conducting monitoring visits; c) the effective use of the feedback form to record classroom activity and d) identifying teacher actions and related student actions in order to connect teacher and student behavior.

1.2.2.2 Parental-awareness intervention

The parental-awareness intervention was developed with the objective of enhancing parental awareness and knowledge of early childhood development and quality kindergarten education and to increase parent-school communication. The intervention was built around three video screening sessions organized by the DSFPs from the district offices during special Parent-Teacher Association meetings over the course of the year. The videos focused on six key messages about the importance of creating a child-friendly classroom that encouraged the use of child-centered, play-based activities for the purposes of learning. In addition, the videos highlighted the roles that parents can play in their children's education and school. The parental intervention was designed to provide parents with a visualization of the expectations of a quality KG classroom in order to inform the ways in which they engage with proprietors, head teachers, and teachers. The parental-awareness intervention involves the screening of these videos followed by discussion, led by district coordinators. The discussions centered on (a) play-based learning, (b) parents' role in child learning, and (c) encouraging parent-teacher and parent-school communication. Notably, the DSFPs were trained to enable meaningful dialogue about the main messages in the videos and to provide tips, suggested activities, and positive reinforcement to parents.

1.3 Theory of change

In Ghana, the majority of KG teachers in low-cost private schools are untrained. At the same time, many parents assume private schools are better quality and assess kindergarten quality based on material infrastructure and whether they perceive schools to do "serious lessons," including repetition of letters and numbers, rather than developmentally appropriate instructional quality and classroom interactions (Bidwell et al., 2014). While training teachers in the KG-specific pedagogy (which is play-based and child-centered) might change teacher behavior, this change might be curtailed by parental requests that emphasize more teacher-driven, academically rigorous instructional methodologies.

The *theory of change* was that children's school-readiness outcomes would be enhanced through improved classroom quality (measured through teacher-child interactions) and teacher professional well-being (measured through teachers' levels of motivation, burnout, and job

satisfaction). For schools that also received the parental-awareness training, it was anticipated that parents would increase their involvement in their children's education and show more support for teachers' implementation of the approaches taught in the training, which would thus enhance the impacts on classroom, teacher, and child outcomes (see Figure 2~~Figure 2~~).

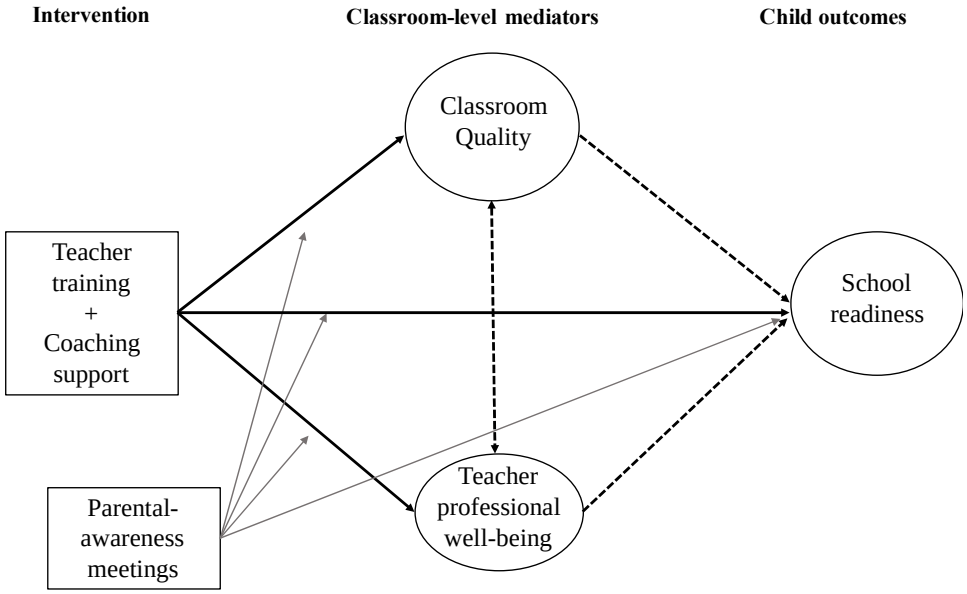


Figure 2. Theory of change. *Notes.* Solid lines represent causal relationships which are tested in this report. Dashed lines represent non-causal relationship, which are not tested.

1.4 Geographic location

The QP4G study was conducted in six districts of the 16 districts of the Greater Accra Region, namely La Nkwantanang-Madina, Ga Central, Ledzokuku-Krowor, Adenta, Ga East and Ga South (Figure 3~~Figure 3~~). These districts were identified as among the nine most disadvantaged in the region using the ranking provided in United Nations International Children's Fund's (UNICEF) District League Table, 2015. Three districts were, however, eliminated due to their distance to Accra, the training location.

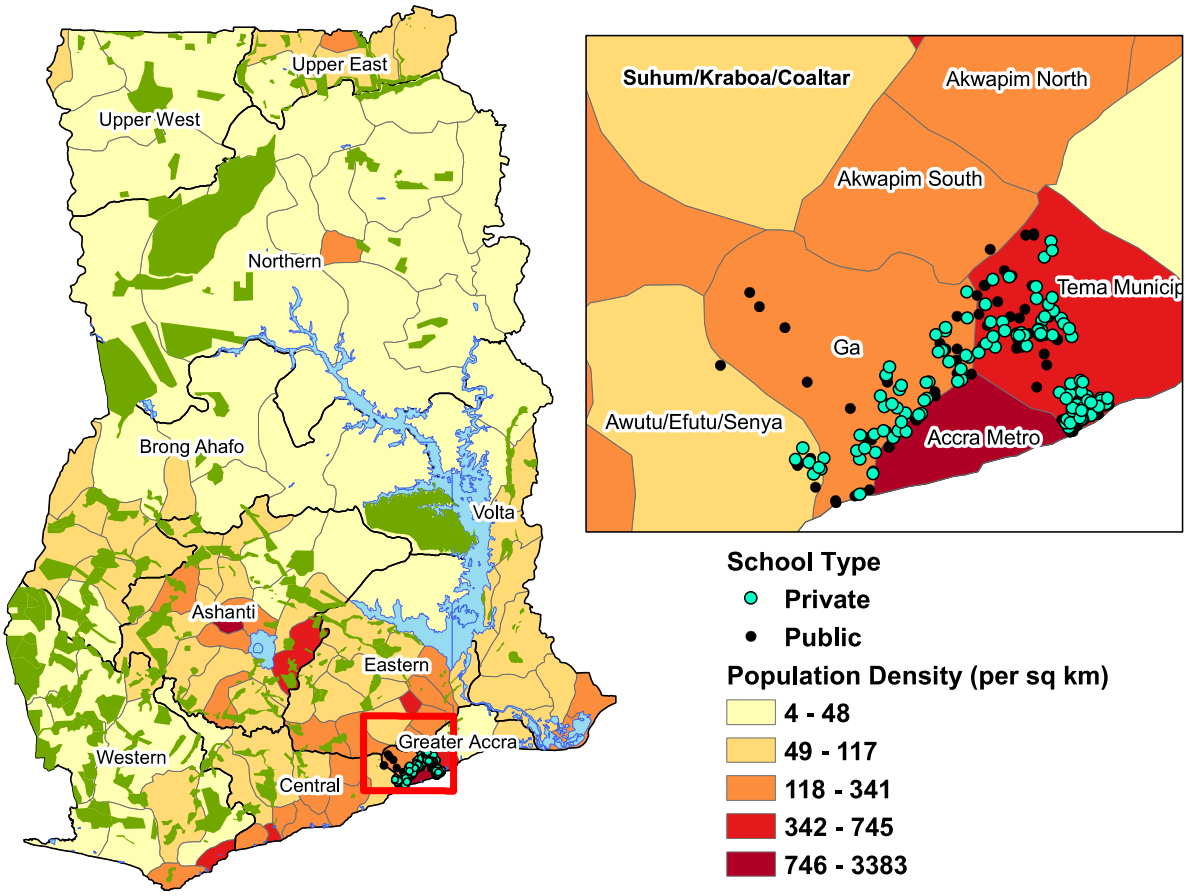


Figure 3. QP4G study areas in the Greater Accra Region

1.5 Intervention implementation and monitoring

1.5.1 Intervention implementation

The in-service teacher training program and parental-awareness intervention were implemented during the 2015-2016 academic year. The implementation activities started with the in-service teacher training program. IPA worked with NNTTC to develop additional resources, familiarize all trainers on the course content and make logistical preparations for the training, as well as to facilitate logistics and administrative work involved in collecting attendance, distributing materials and overseeing the quality of implementation. The implementation activities were conducted within one academic year, i.e., 2015/2016 academic year (Table 1~~Table 1~~). The main part of the implementation activities was conducted during the first term of the 2015/2016 academic year. Refresher programs were conducted during the second and third terms of the 2015/2016 academic year.

Table 1. QP4G Implementation Timelines

Implementation Activities	Period	
	From	**To**
5 days teacher training program - Term I	17th Sept.	9th Oct. 2015
1st parental-awareness meeting - Term I	18th Nov.	10th Dec. 2015
2 days of refresher teacher training program - Term II	7th Jan.	14th Jan. 2016
2nd parental-awareness meeting - Term II	8th Feb.	5th April 2016
1 day of refresher teacher training program - Term III	16th May	18th May 2016
3rd parental-awareness meeting - Term III	5th June	20th July 2016

The majority of schools assigned to the treatment conditions had two KG teachers, though the range was from one to five. All KG teachers were invited to participate in the training. If there were more than two KG teachers in the school, all teachers were invited to participate in the training and two teachers were randomly sampled per school for the evaluation (one from KG1 and one from KG2). Some schools only had one KG teacher, and in this case, the one teacher was sampled. In addition, school head teachers, and district staff were invited to attend the training. Overall, training received overwhelmingly positive feedback from participants. The attendance rates for the various training sessions are shown in Table 2~~Table 2~~. Following training, a randomly selected cohort of teachers also received text messages with tips that reflect teaching from the training sessions.

Table 2. Attendance rates during the teacher training program

	Invited				**Attended**				
	Teachers	*Head Teacher*	*District Staff*	*Total*	*Teachers*	*Head Teachers*	*District Staff*	*Total*	*Attendance Rate*
Term 1	393	183	17	591	337	124	15	476	81%
Term 2	395	0	44	388	315	0	29	344	88%
Term 3	317	161	45	523	254	93	23	370	71%
Average	*368*	*114.67*	*35.33*	*500.67*	*302*	*72.33*	*22.33*	*396.67*	*80%*

The parental-awareness intervention was conducted by the DSFPs. The parental-awareness intervention was held at school PTA meetings and was open to all parents with KG children. It involved video screening followed by discussion, led by the DSFPs with focus on a) play-based learning, b) parents' role in child learning, and c) encouraging parent-teacher and parent-school communication. The parental-awareness intervention was attended mostly by women between

the ages of 22 and 30 and meetings averaged between 26 and 30 participants (Table 3~~Table 3~~). The parent meetings were intended to be interactive and provide a platform for parents to provide inputs and ideas. They included the use of multimedia, and ended with the distribution of snacks. Many of the district staff expected to see increased participation from parents over the course of the year as more of them hear about the nature of the meetings.

Table 3. Attendance rates for the parental intervention

District	Schools	Expected # of Parent/school	Total Expected	Term 1	Term 2	Term 3
Adenta	10	40	400	240	307	208
Ga Central	10	40	400	206	343	258
Ga East	12	40	480	276	267	255
Ga South	21	40	840	453	678	758
Ledzokuku-Krowor	18	40	720	353	582	424
La Nkwantanang-Madina	9	40	360	170	152	143
Total	*80*	*240*	*3200*	*1698*	*2329*	*2046*

1.5.2 Intervention monitoring

Following the delivery of the teacher training, the DSFPs were also trained to ensure they could adequately monitor and coach teachers, and administer the parental-awareness intervention. The DSFPs were dispatched to the participating schools to conduct coaching visits and monitoring. In addition to the general monitoring guidelines, DSFPs were particularly entreated to adhere to specific guidelines during their monitoring visits as follows; a) to spend at least an hour in observation and about 20 - 30 minutes in giving feedback and offering advice for steps to improving the lessons; and b) once observations and coaching were completed, to leave one of the three copies of the monitoring checklist with the head teacher after speaking to the head teacher about their observations. This was meant to ensure accountability between the district and the school, showing that the DSFP had visited and ensured the head teacher was aware of the progress of their teachers and could better support the teachers in their roles. Overall, districts monitored over 90% of schools and reached selected schools for parental engagement activities in the first semester. On average, about 447 monitoring visits were conducted during the implementation phase (Table 4~~Table 4~~). The highest number of monitoring visits was conducted during the third term of the program. Through monitoring feedback given by district staff, decisions were made about how best to plan the refresher training programs to address

the biggest challenges that were observed: classroom management and the development and use of learning centers.

Table 4. Monitoring visits conducted by DSFPs

Term	Number of Monitoring Visits
Term 1	300
Term 2	502
Term 3	540

Notes. This table presents the total number of visits conducted across the treatment groups at each term. Since each teacher was supposed to receive 2 visits per term, the total visits should have been close to 700 for each term. The table shows that the total number of planned visits were not conducted, but that each term district officials did conduct more visits than in the previous term.

Another way this information can be considered is through teachers' experiences with the training measured via the average number of days teachers attended the training and received coaching visits as an overview of implementation. Of the teachers selected to participate in the program at baseline, teachers in the TT arm (N = 154) on average attended 6.4 (SD = 2.1) days of the 8 total days of training, with an average of 4.6 days (SD = 1.4) of the primary 5-day training, 1.2 days (SD = 1.0) of the first two-day refresher training, and 0.7 days (SD = 0.5) of the final one-day refresher. In the TTPA arm (N = 149), on average teachers attended 6.4 (SD = 2.4) days of the 8 total days of training, with an average of 4.3 days (SD = 1.6) of the primary 5-day training, 1.4 days of the first refresher (SD = 0.9) of the first two-day refresher training, and 0.8 days (SD = 0.4) of the final one-day refresher. Notably, a few teachers in the control group (N = 141) attended the training, with an average of 0.1 (SD = 0.7) total days. For the coaching visits, teachers in the TT arm received an average of 3.7 coaching visits (SD = 2.2) over the year, and teachers in the TTPA arm received 4.0 (SD = 2.1) visits.

Section 2: Evaluation Design and Data

2.1 Study Design

The study employed a cluster-randomized design. In six districts of the Greater Accra Region, 108 public schools and 132 private schools were selected for the evaluation, and randomly assigned to one of the following groups: (1) in-service teacher-training program only; (2) in-service teacher-training program plus parental-awareness program; or (3) control (current standard) condition. Randomization was stratified by two levels: districts and public and private sector. In addition, treatment schools were then randomly assigned to receive reinforcement messages (weekly text messages for teachers and picture-based paper flyers for parents).

2.2 Research questions

The current study is an impact evaluation to test the efficacy of the QP4G programs with kindergarten classrooms in both private and public schools to improve (a) teacher well-being, (b) the quality of KG teacher practices and interactions with children, and (c) children's school readiness and learning skills. Additional goals of this evaluation are to test the added value of combining a scalable (low-cost) parental-awareness intervention with teacher in-service training; to compare implementation challenges in public and private schools, and to examine several important sources of potential heterogeneity of impact, primarily impacts in public versus private schools. Specifically, we test two hypotheses as our primary analysis and three additional research questions as our secondary analyses:

Research Question 1: What are the impacts of the QP4G teacher training program on teacher professional well-being, classroom quality, and children's school readiness relative to a control group?

Research Question 2: What are the impacts of the QP4G teacher training paired with parental-awareness meetings on teacher professional well-being, classroom quality, and children's school readiness relative to (i) a control group and (ii) the QP4G teacher training only?

Research Question 3: Do reinforcements treatment school teachers (in the form of text messages) and parents (in the form of paper flyers) strengthen the impact of each program on teacher professional well-being, classroom quality, and children's school readiness?

Research Question 4: Do impacts of the intervention vary by key child characteristics (gender,

grade level, baseline school readiness skills) and school sector (public vs. private)?

Research Question 5: Are there any sustained impacts of the QP4G teacher training and QP4G training provided with parental-awareness meetings one year after implementation has ended?

2.3 Sample and sampling procedure

2.3.1 Sample selection

Sample selection for the QP4G study involved a broad three-stage sampling design comprising the selection of the districts, schools, and respondents.

a. **Selection of districts**. The first stage involved purposive selection of six (6) districts within Greater Accra Region based on two criteria: (a) disadvantaged districts in terms of infrastructure (using UNICEF's District League Table scores); and (b) close proximity to Accra for travel for the training of the KG teachers. The six "disadvantaged" districts in the Greater Accra Region were based on 2015 UNICEF's District League Table created in collaboration with the Ministry of Local Government and Rural Development. The six selected districts were La Nkwantanang-Madina, Ga Central, Ledzokuku-Krowor, Adentan, Ga South and Ga East.

b. **Selection of basic schools with kindergarten**. The second stage involved a stratified random sampling of public and private schools with a kindergarten unit (i.e., KG1 or KG2 or both) from each of the selected districts in the Greater Accra Region. The sampling frame for the schools was obtained from the EMIS database of the Ghana Education Service. Two-hundred and forty schools were randomly selected, stratifying by district and by public and private status. Since there were only 108 public schools, all were included in the sample. One hundred and thirty-two private schools were randomly selected in proportion to the total number of private schools in each district relative to the six districts. In each district, 20 additional private schools were randomly sampled to serve as "reserve" schools in the event that one of the original schools sampled refused or was not eligible to participate in the study. Eventually, eleven schools were replaced from the original 240 at baseline. Examples of needing to use reserve schools were: refusal to participate, a discovery that a school did not have a kindergarten program, and that a school listed in the EMIS dataset no longer existed. In the spring midline assessment, 2 schools had closed and 3 schools dropped out of the study. At endline,

all the schools from the midline round remained in the sample. Thus, the sample for the impact analysis was 235 schools.

c. **Selection of respondents**. The final stage involved the selection of the respondents. The QP4G study recognizes the central role played by diverse stakeholders in the task of delivering high-quality education to improve learning results for kindergarten children in Ghana. Recognizing the role of such stakeholders, data for the various outcomes were collected from KG teachers, KG children, caregivers of selected KG children, head teachers and school proprietors (for the private schools only). The subjects were sampled using different sampling techniques:

a. The majority of schools had two KG teachers, though the range was from one to five. All KG teachers were invited to participate in the training. If there were more than two teachers in a school, two were randomly sampled per school for the evaluation (one from KG1 and one from KG2). Thirty-five schools only had one KG teacher, and in this case the one teacher was sampled. The final sample included 444 teachers.

b. Fifteen child-caregiver pairs from each school (8 from KG 1 and 7 from KG 2) were sampled randomly. The whole universe of KG children was included in the survey for schools with less than 15 children. The KG children were sampled from the particular KG class from which a KG teacher was selected. A total of 3435 KG children were recruited into the QP4G study. Due to the difficulty in obtaining contact information on the caregivers of selected children and the incorrect / non-working phone numbers given by schools/caregivers, only 2,159 caregivers, representing 63% of the number of KG children recruited, were included in the study at baseline. At midline, the number of caregivers reached increased to 79% as we collected more contact information on caregivers from school heads and teachers. This, however, fell to 75% at endline due largely to non-working phone numbers.

c. Head teachers and/or school proprietors of selected schools were automatically included in the study. Two-hundred and forty head teachers and 156 proprietors were included in the QP4G study.

2.3.2 Survey attrition

Survey attrition may be a major concern when conducting school and household surveys. This is because if those who are followed-up are different in some way from those who are not, this could potentially bias the estimates. Available evidence suggests that Ghana's educational system experiences high rates of teacher mobility and turnover (Osei, 2006). To minimize attrition rates, several methods were adopted. This included collecting detailed contact information on KG teachers and caregivers; providing incentives (e.g., GHC 5 airtime, a branded desktop calendar to schools, QP4G training curriculum[4]) to respondents; and providing baseline and midline briefs to school heads and teachers. Of the 444 teachers who were recruited for the study at baseline, 68% (n=302) stayed from baseline to endline. For the KG children, 77% (n = 2,657) stayed from baseline to endline. This showed an average of about 11.5% attrition rates for the KG children and almost 16% for the KG teachers for each wave of data collection of the QP4G study. Appendix A Table 1 and Appendix A Figure 1 present the sample flow chart and sample attrition across all waves of the study.

2.4 Analytical strategy

2.4.1 Missing data imputation

We used multiple imputation (with Stata's "ice" command) to handle missing data on all missing variables, using three rounds of data collection (baseline and follow up, as well as a second follow up of data). While the data are not missing completely at random (MCAR), if variables that strongly predict attrition are incorporated into the missing data strategy, the plausibility of a missing at random (MAR) assumption increases (Young & Johnson, 2015). In other words, in estimating multiple chains of models by including a large set of covariates, including those that predict differential attrition, assumptions of MAR have been shown to be robust.

We conducted the imputation in two steps. First, using a rich set of teacher demographic and background variables, as well as outcome scores for professional well-being and classroom quality across all three waves, we imputed 20 teacher-level datasets. Second, we randomly selected ten of these teacher datasets. We merged each individual dataset with child outcome data and basic child demographic characteristics from all three waves of data. For each of the 10 teacher datasets, we imputed 10 child datasets, resulting in 100 child-level datasets.

[4] This was provided to all schools two years after the in-service teacher training program.

2.4.2 Impact estimates

All impacts are assessed within hierarchical linear models, with children nested in teachers and teachers nested in schools. We use the baseline schools from which teachers and children were sampled, regardless of within-sample mobility, which provides a more conservative estimate of impacts.

Separate models were fitted to estimate main intervention impacts on (a) teacher professional well-being (i.e., motivation, burnout, and job satisfaction), (b) classroom quality factors (i.e., fidelity checklist, facilitating deeper learning, emotional support and behavior management, and supporting student expression), and (c) children's school readiness (i.e., total IDELA score). As a post-hoc test, we estimated impacts on each of the four individual domains of children's school readiness (i.e., early literacy, early numeracy, social-emotional development, and executive function) to assess if impacts on child outcomes were driven by any particular domain. The equations for the three-level model were as follows:

Level 1 (Child-level) Model:

$$Y_{ijk} = B_{0jk} + B_{1jk}'X_{ijk} + e_{ijk}$$

Where $\mathbf{X_{ijk}}$ is the vector of child covariates (gender, age, baseline score, student mobility dummy variables).

Level 2 (Classroom-level) Model:

$$B_{0jk} = \gamma 00k + u0jk$$

Where $\mathbf{B_{0jk}}$ is the classroom-level random intercept.

Level 3 (School-level) Model:

$$\gamma_{00k} = \pi_{000} + \pi_{001}TT_k + \pi_{002}TTPA_k + \pi_{003}'Z_k + v_{00k}$$

Where $\boldsymbol{\gamma_{00k}}$ is the school-level random intercept; $\mathbf{Z_k}$ is the vector of school-level covariates (district dummies, private or public status, within sample mobility dummies); and $\mathbf{TT_k}$ is an indicator for schools assigned to the teacher training condition, and $\mathbf{TTPA_k}$ is an indicator for schools assigned to the teacher training plus parental-awareness condition.

Third, as a secondary analysis, we examined whether intervention impacts were moderated by child characteristics (gender, child baseline scores, and grade level [KG1 and KG2]) and by school sector (private or public). Moderation of impacts by child covariates was tested by adding a cross-level interaction term between each treatment condition (at level 3) and child characteristic (at level 1). Moderation by sector was calculated with an interaction term (at level 3) between school sector (*1* = private, *0* = public) and treatment status.

Each model was estimated at midline (a re-estimation from the SIEF midline report) and endline. All impact estimates on teacher and classroom-level data were computed on the 20 teacher-level datasets described in Section 2.4.1 (using Stata's "mi estimate" command). All impact estimates on child-level outcomes were computed on the 100 child-level datasets also described in Section 2.4.1.

2.5 Measures

A number of instruments were used and developed for the purpose of the study. Information was collected at the school level, the teacher and classroom levels, the child level and the household level. Below we describe the dependent variables used in the impact analysis. The mean (*M*), standard deviation (*SD*), and Cronbach's alpha (representing internal consistency for the items making up a particular scale; *α*) are presented for each construct.

2.5.1 Teacher professional well-being

Teachers answered a survey in English. Items were selected from existing scales and were pilot tested. First, we conducted five cognitive interviews with teachers to assess whether they understood each question, both consistently across constructs and in the way the item was intended (Collins, 2003). Next, we piloted the survey by administering it to 20 teachers and then assessed the distribution of responses for each item. From both of these exercises, we concluded that all items were suitable for use in this sample. Notably, all items have been used in previous research with teachers in sub-Saharan Africa (Wolf, Aber et al., 2015; Wolf, Torrente et al., 2015). Factors were derived through exploratory factor analyses conducted with the baseline data. All outcomes were measured at baseline and follow-up.

2.5.2 Motivation

Teacher's motivation was measured using five items adapted from Bennell and Akyeampong (2007) as reported in Torrente et al. (2012). Items were answered on the following scale: *1 = false*, *2 = mostly false, 3 = sometimes, 4 = mostly true, 5 = true.* Sample items included "I am motivated to help children develop well socially (i.e., behave well, get along with peers, cooperate)" and "I am motivated to help children learn math" ($M = 4.6$, $SD = 0.59$, $\alpha = 0.77$).

2.5.3 Burnout

Teacher burnout was measured using 11 items from the Maslach Burnout Inventory (Maslach, et al., 1996). Items asked teachers to use a scale from *1* ("never") to *7* ("every day") to indicate, for instance, how often they have felt "emotionally drained from my work," "fatigued when I get

up in the morning and have to face another day on the job", and "burned out from my work" (*M* = 2.03, *SD* = 0.90, *α* = 0.75).

2.5.4 Job satisfaction

Teacher's job satisfaction was measured using six items adapted from Bennell and Akyeampong (2007) as reported in Torrente et al. (2012). Items were answered on the following scale: *1 = true*, *2 = somewhat true, 3 = somewhat false, 4 = false.* Sample items include "I am satisfied with my job at this school", "I want to transfer to another school", and "Other teachers are satisfied with their decision to be a teacher in this school." Responses to each item were coded so that higher scores indicated higher job satisfaction (*M* = 3.09, *SD* = 0.69, *α* = 0.73).

2.5.5 Turnover

Teacher turnover (*1* = yes, *0* = no) was indicated if the teacher had left his or her position by follow-up data collection in the third term. If the teacher was absent, confirmation was obtained from the school administration that the teacher had left his or her position at the school. Approximately one-quarter of teachers (*N* = 107) had left their position by follow-up.

2.5.6 Classroom outcomes

All teachers were videotaped teaching a lesson in their classrooms for 30–60 minutes in May or June of 2016. Videos were coded with two instruments: an implementation fidelity checklist and a tool to assess the quality of teacher-child interactions. Both were assessed at follow-up only and collected for treatment and control schools.

2.5.7 Fidelity of implementation

We created a checklist of 15 activities that were explicitly covered in the teacher training related to behavior management and instructional practice. Each practice was coded as either present in the video (a score of 1) or absent in the video (a score of 0). Items included: "Teacher praises children for positive behavior", "Teacher threatens children with or used a cane on children at least once (reverse coded)"; "Teacher explicitly reminds children of the class rules"; "Teacher uses a signal to gain children's attention (e.g., drum beat, song, bell); "Children are seated in a way that children can see each other's faces (e.g., in a circle, or tables together in groups)"; "Teacher uses one or multiple songs to facilitate learning at some point in the lesson"; and "There is an activity that facilitated the lesson objectives that involved manipulation of materials" (*M* = 3.51, *SD* = 2.22).

2.5.8 Teacher-child interaction quality

All videos were coded using the Teacher Instructional Practices and Processes System (TIPPS; Seidman et al., 2013; Seidman et al., 2017). The TIPPS is a classroom observation tool for assessing classroom quality that focuses on the nature of teacher-child interactions; it was created for use in low- and middle-income countries. We used the TIPPS-Early Childhood Education version and made minor adaptations for use in Ghana (e.g., referring to pupils as children, as is common in Ghanaian kindergarten settings). More information about the assessment tool can be obtained by referring to Seidman et al. (2013, 2017).

The TIPPS is made up of 19 items. We dropped four items due to lack of variability in their scores across classrooms. We then randomly split the sample in half and conducted an exploratory factor analysis with one half, and confirmed the final model on the second half. Based on the results, we grouped the remaining 15 items into three factors: *facilitating deeper learning* (3 items: connects lesson to teaching objectives; provides specific, high-quality feedback; and uses scaffolding; α = 0.42), *emotional support and behavior management* (7 items: positive climate; negative climate; sensitivity and responsiveness; tone of voice; positive behavior management; provides consistent routines; and student engagement in class activities; α = 0.83), and *supporting student expression* (4 items: considers student ideas and interests; encourages students to reason and problem solve; connects lesson to students' daily lives; and models complex language; α = 0.63). See Wolf et al. (2017) for details on the analysis and concurrent validity of the three factors in this sample.

Reliability. Video coders were trained and had to achieve the pre-specified levels of reliability in order to pass the training. Raters were recruited in Ghana, had a bachelor's or master's degree, and attended a five-day training session on the instrument. Each rater had to meet or exceed three TIPPS calibration criteria within three attempts to be certified as a TIPPS observer. TIPPS calibration criteria not only look at agreement but also the degree of deviation from master codes - both are important aspects given that there are only four scale points and that understanding of the concept is critical for precise coding (see Seidman et al., 2013 for details on calibration cut-offs). Collectively, these three criteria enhance the likelihood of achieving acceptable levels of inter-rater reliability. Raters who achieved calibration were also required to participate in 30-minute weekly refresher sessions led by TIPPS trainers that included a review of different manual concepts, short practice videos, and time for questions and discussion.

To assess inter-rater reliability, 15% of videos collected at baseline were coded by three raters. We calculated the ICC of the final scores to assess how the partition of the variance in scores

breaks down into differences in individual raters and shared variance across raters. On average across items, 71.1% of the variance was shared across raters.

2.5.9 Child school-readiness outcomes

Children's school readiness was directly assessed in four domains: early literacy, early numeracy, social-emotional function, and executive function. The instrument used was the International Development and Early Learning Assessment (IDELA), developed by Save the Children (Pisani et al., 2015). The tool was translated into three local languages (Twi, Ewe, and Ga). Surveys were translated and then back-translated by a different person to check for accuracy. Any discrepancies were discussed and addressed. Finally, after being trained on the instrument, a group of surveyors read and discussed the translated version in their respective local language and made additional changes as a group.

Early literacy. The domain of early literacy consists of 38 items grouped into six subtasks and covers constructs of print awareness, letter knowledge, phonological awareness, oral comprehension, emergent writing, and expressive vocabulary. An example subtask on phonological awareness asked children to identify words that begin with the same sound. A sample item is: "Here is my friend mouse. Mouse starts with /m/. What other word starts with /m/? Cow, doll, milk" (α = 0.74).

Early numeracy. The domain of early numeracy consists of 39 items grouped into eight subtasks, and it covers constructs of number knowledge, basic addition and subtraction, one-to-one correspondence, shape identification, sorting abilities based on color and shape, size and length differentiation, and completion of a simple puzzle. An example item assessing shape identification showed the child a picture with six shapes and asked the child to identify the circle (α = 0.72).

Social-emotional development. The domain of social-emotional development consists of 14 items grouped into five subtasks and covers constructs of self-awareness, emotion identification, perspective taking and empathy, friendship, and conflict and problem-solving. An example item of conflict solving involved asking the child to imagine he or she is playing with a toy and another child wants to play with the same toy and asking the child what they would do to resolve that conflict. "Correct" answers in the Ghanaian context as agreed upon by the assessors during training included talking to the child, taking turns, sharing, and getting another toy (α = 0.69).

Executive function. The domain of executive function was assessed with 10 items grouped

into two subtasks focused on working memory (i.e., forward digit span) and impulse control (i.e., head-toes task). For the forward digit span, assessors read aloud five-digit sequences (beginning with two digits and increasing up to six digits), and children were asked to repeat the digit span; their responses were marked as correct or incorrect. For the head-toes task, assessors asked children to touch their toes when the assessor touched his or her head, and vice versa, in a series of five items (α = 0.83).

School-readiness composite. For the primary impact analysis, scores on the four domains measured via direct assessment were combined to create a total "school readiness" score.

Reliability. Interrater reliability on the child outcome measure was assessed. Enumerators were paired and each scored two children together. Cohen's kappa values were calculated for each pair across each item in the entire assessment, and values ranged from 0.67 to 0.97, with an average kappa value of 0.86.

2.5.10 Covariates

We included a select set of covariates to improve the precision of our impact estimates. For all models, these included private-sector status of the school, five district dummies, a dummy variable for if the school was randomly assigned to receive teacher text messages, a dummy for if the school was randomly assigned to receive parent flyers, and a series of five dummy variables accounting for within-sample mobility (e.g., between baseline and follow-up a baseline school split into two separate schools; two schools merged into one school; children or teachers moved to a different school within the sample). For estimating impacts on child outcomes, we also included child gender, age in years, KG level (1, 2, or 3; 3 is a categorical variable if KG1 and KG2 were combined in one classroom), and baseline score for each respective outcome. For estimating impacts on teacher outcomes, we also included teacher gender, age, level of education, years of teaching experience, and baseline score for each respective outcome.

2.6 Data collection timelines

The QP4G study spans two school years to allow estimation of one-year effects and to what extent effects persisted after another year. The study started in January 2015 and ended in July 2017 (Table 5~~Table 5~~). In total, three main rounds of quantitative data collection were conducted over the study period – baseline, midline, and endline – but each round spanned several months as different survey instruments were administered sequentially. In addition, two rounds of qualitative data collection occurred during the quantitative baseline and midline. Before

baseline, a school listing exercise was also conducted to obtain basic information on the schools and finalize the sampling frame.

Table 5. Data Collection Timelines

Phase	Main Data Collection Activities	Period	
		From	To
Listing	School listing exercise	18th May	22nd May 2015
Baseline	Baseline I: Headteacher and proprietor surveys	10th June	25th June 2015
	Qualitative research – focus group discussion with parents	27th July	July 29, 2015
	Baseline II: Teacher survey and classroom observation	9th Sept.	14th Sept. 2015
	Baseline II: Child assessment	28th Sept.	21st Oct. 2015
	Baseline II: Caregiver survey	19th Oct.	3rd Dec. 2015
	Baseline II: Video coding exercise	18th Jan.	8th Feb. 2016
Midline	School survey: Teacher survey and child direct assessment	16th May	11th Jul. 2016
	School attendance records	11th July	25th July 2016
	Video Coding	20th July	12 Aug. 2016
	Caregiver survey	18th Aug.	26th Sept. 2016
	In-depth qualitative interviews with teachers and caregivers	4th Nov.	11th Nov. 2016
Endline	School survey: teacher survey, child assessment, classroom observation, and school attendance records	6th Feb.	31st March 2017
	Video Coding	3rd April 2017	6th May 2017
	Caregiver survey	22nd May	7th July 2017

2.7 Ethical considerations

Ethical review and approval were provided by IPA's Institutional Review Board (IRB), New York University's University Committee on Activities Involving Human Subjects (UCAIHS) and the University of Pennsylvania's IRB. These reviews provide assurance that the research protocols developed in the framework of this study respect ethical principles and guidelines for the protection of participating human subjects, especially for vulnerable participants including children. All the members of the evaluation team completed a recognized human subjects training (either the National Institute of Health's Protecting Human Research Participants or the Collaborative Institutional Training Initiative's Certification in Human Protection).

The anticipated risk due to participating in this study was minimal and sensitive questions were not included in the instruments. Informed consent was formally requested at the beginning of each interview with the teachers and the head teachers to ensure that all respondents were willing to participate in the study and were informed of the risks incurred in doing so. In the case of the kindergarten pupils, verbal assent was obtained before the administration of the assessment and surveyors were trained to carefully pay attention to the child during the assessment in order to identify any sign of distress or sign that the child wanted to stop the test.

Confidentiality of data and of participants is of the highest priority to the study team, while in the field and in the offices. Every effort was made to ensure confidentiality by holding surveys in private settings and the study team members took great care in protecting data in both transport and storage. All digital data including personally identifiable information is encrypted and only accessible to team members approved by the IRBs.

2.8 Data quality control systems

A number of control systems were put in place to guarantee quality data. These included programming of instruments using SurveyCTO, observing and monitoring fieldworks, running high-frequency checks, and conducting audit [discrepancy] checks.

With the exception of the qualitative interview protocols, all the instruments were programmed on SurveyCTO. The programmed instruments were administered using Samsung tablets to save time and enhance accuracy. Enhanced quality controls systems such as automatic skip patterns, relevance, and constraints were integrated into the programming to guarantee data quality. The programming was done by an IPA programmer and the Research Associate.

Observations took the form of accompaniments and spot checks. Survey accompaniments, where an assessor from the country office evaluation team or the field supervision team sat with the surveyor and observed them conducting a survey, were frequently conducted at the initial stage of the data collection to ensure that the questionnaire was appropriately understood and administered by the surveyors. Random spot checks, when supervisors unexpectedly visited a school and observed the team operating, were regularly conducted during the survey period. Both forms of observation or monitoring were followed by immediate feedback and, if needed, in-field refresher training. Observations were made to establish whether the protocols were followed and to assess the performance of the field staff. Specific monitoring protocols, namely, the IDELA Monitoring Form, the Teacher Interviewers Monitoring Form, and the Video Quality Form were designed and used to ensure consistency in reporting and aid in providing feedback.

Daily data quality checks - high-frequency checks (HFCs) – were conducted for all data (excluding the qualitative data) sent to the SurveyCTO server. HFCs were done to establish whether there were violations of data quality indicators - duplicate IDs, missing values, constraints, skip patterns and survey logic or inconsistencies. As per IPA guidelines, 15% or more of all surveys were audited. Each data collector's work was frequently audited. Auditing of surveys was conducted for all administered surveys excluding the child direct observations. The video coding activities were audited by coding randomly selected videos by independent video coders who were not part of the actual video coding exercise. The HFCs and audit discrepancy checks were conducted to identify and correct any mistakes in the data submitted by the field teams, potentially leading to follow-up visits to respondents to clarify specific answers.

As part of the coding of the classroom videos, we peer-coded randomly selected videos. Each video coder had the opportunity to code at least four videos from his/her colleague. Each of these four videos was coded by two different video coders. Thus, each sampled video for the peer coding was actually coded three times by three distinct video coders during the video coding exercise – the first video coder who coded the original video and two other video coders who "peer-coded" the same video. The allocation of the six randomly selected videos per each video coder followed a pairing regime that employed the mathematical permutations rule. Only the original video coding was to be used in the analysis. The two subsequent "peer-coded" videos were for quality control and feedback purposes only. The videos that were "peer-coded" were not included in the sampled videos for the auditing to avoid duplication of data quality measures.

Section 3: Findings

3.1 Teacher professional well-being

3.1.1 Midline

Appendix A Appendix A Table~~Table 2~~ shows the results of analyses estimating the impact of the two treatment conditions on teachers' motivation, burnout, job satisfaction, and turnover. There were no program impacts on either motivation or job satisfaction. The program did impact teacher burnout, reducing burnout in the TT condition ($p < .01$, d_{wt}[5]= -.40) and the TTPA condition ($p < .001$, d_{wt} = -0.55). Additionally, the TT condition impacted teacher turnover, reducing the probability that a teacher would leave the KG classroom by the third term by 43.5% ($p < .05$, OR = 0.30), reducing turnover from 44.3% of teachers to 26.8%.

To test differences between the TT and TTPA conditions, all models were re-run with TT as the reference group. While the full sample was included in this analysis, the interest was in the coefficient estimate of TT vs. the coefficient estimate of TTPA. There were no statistically significant differences between TT and TTPA arms, indicating that across the two treatment conditions impacts were not different.

3.1.2 Endline

At endline, impacts on reduced teacher burnout persisted for the TTPA arm (d_{wt}= -.29, p < .05). Impacts in the TT arm were still negative (d_{wt}= -0.19), but not statistically significant (p = .155). There were no statistically significant treatment impacts on teacher turnover from midline to endline.

When contrasting the TT vs TTPA arms, there were no statistically significant differences between the two conditions for any of the outcomes, including burnout.

[5] d_{WT} represents a standardized mean difference between treatment and control clusters. This was calculated with the following equation from Hedges (2009):

$$d_{WT} = \frac{b}{\sqrt{\hat{\sigma}^2_{BS} + \hat{\sigma}^2_{BC} + \hat{\sigma}^2_{WC}}},$$

where b represents the unstandardized regression coefficient with covariate adjustment (e.g., b = .11), and the three terms of the denominator represent variances at the cluster, school, and child levels, respectively, without covariate adjustment. The rationale behind covariate adjustment for the treatment effect, but not the variances, was to obtain a more precise treatment effect (i.e., adjusted), but standardized based on typical (i.e., unadjusted) variances at each level (L. V. Hedges, personal communication, November 3, 2014). Variance estimates for each level were computed using the pooled estimates across all imputed datasets using Rubin's combining rules. This same approach was utilized to estimate d_{WT} for this and other main effects presently reported.

3.2 Classroom quality

3.2.1 Midline

Appendix A Appendix A Table~~Table 2~~ shows the impact estimates on classroom outcomes. We first addressed the question of fidelity of implementation. We assessed the number of developmentally appropriate practices observed in the classroom using a checklist of 15 instructional practices that were specifically promoted in the teacher training. The program increased the number of activities teachers used in the classroom in both treatment conditions by similar magnitudes ($p < .001$, $d_{wt} = 0.54$ in TT and 0.60 in TTPA). Compared to control group classrooms, where teachers implemented an average of 3.1 activities during the observational assessments, in the TT and TTPA condition, teachers implemented 4.7 and 4.8 activities, respectively.

Next, we assessed impacts on classroom quality based on three domains of teacher-child interactions: facilitating deeper learning (e.g., scaffolding, high-quality feedback), supporting student expression (e.g., considering student ideas during the lesson, encouraging students to reason and problem solve), and emotional support and behavior management (e.g., positive climate, teacher sensitivity and responsiveness to student needs, providing consistent routines). There were no impacts of either treatment condition on levels of facilitating deeper learning. Both treatment conditions increased the level of emotional support and behavior management observed in the classroom ($p < .001$, $d_{wt} = .62$ in the TT condition and 0.64 in the TTPA condition). Finally, the TT condition increased levels of supporting student expression in classrooms ($p < .01$, $d_{wt} = .48$), but there were no statistically significant impacts of the TTPA condition.

To test differences between the TT and TTPA conditions, all models were re-run with TT as the reference. There were no statistically significant differences between TT and TTPA arms.

3.2.2 Endline

At endline, impacts on the fidelity checklist were still significant, though smaller in magnitude ($d_{wt} = .21$, $p < .05$ in the TT condition; $d_{wt} = .22$, $p < .01$ in the TTPA condition).

Surprisingly, there were statistically significant and *negative* impacts at endline on teachers' support of student expression ($d_{wt} = -.40$, $p < .05$ in the TT condition; $d_{wt} = -.72$, $p < .01$ in the TTPA condition). There were no other statistically significant impacts on classroom quality outcomes at endline.

To test differences between the TT and TTPA conditions, all models were re-run with TT as the reference. There were no statistically significant differences between the TT and TTPA arms.

3.3 Child school-readiness outcomes

3.3.1 Midline

Appendix A Appendix A Table ~~Table 3~~ presents the impact estimates of the treatment programs on children's school readiness. We first assessed impacts on the composite score of children's school-readiness skills as our primary outcome of interest. We then conducted post-hoc analyses to assess impacts on each domain of development individually to examine if the findings were driven by any particular developmental domains. The TT program increased children's school readiness ($p < .05$, $d_{wt} = .15$). When broken down by domain, impacts were statistically significant for three of the four domains, including early numeracy ($p < .05$, $d_{wt} = .11$), early literacy ($p < .05$, $d_{wt} = .14$), and social-emotional development ($p < .05$, $d_{wt} = .17$). There were no impacts on children's school readiness when the parental-training program was added to the teacher training (TTPA).

Compared to the TT condition directly, children in the TTPA condition had marginally statistically lower scores on overall school readiness ($b = -.018$, $SE = 0.010$, $p < .10$), and on the domain of early numeracy ($b = -.026$, $SE = 0.011$, $p < .05$).

3.3.2 Endline

At endline, marginally statistically significant impacts of the TT condition were detected on children's overall school readiness ($p < .10$, $d_{wt} = .11$). When broken down by individual domains, impacts on social-emotional development were statistically significant ($p < .05$, $d_{wt} = .14$), and impacts on executive function were marginally statistically significant ($p < .10$, $d_{wt} = .11$).

Compared to the TT condition directly, children in the TTPA condition had statistically lower scores on overall school readiness at endline ($b = -.018$, $SE = .008$, $p < .05$). On individual domains, children in the TTPA condition had marginally statistically significant lower early literacy scores ($b = -.023$, $SE = .014$, $p < .10$) and social-emotional development ($b = -.022$, $SE = .011$, $p < .10$), and statistically significantly lower executive function scores (b = -.022, SE = .012, p < .05).

3.4 Impacts of Reinforcements

Our first additional research question was to assess if added reinforcements to teachers via text message or the flyers to parents on school readiness strengthened the impact of each program (results not shown). We found no impacts on any outcome (results not shown) of the text message reinforcements to teachers or flyers to parents at midline. At endline (one year after reinforcements ended), there were only three significant effects of either the parent or teacher reinforcement, which we conclude was not greater than would occur by chance.

3.5 Moderation by Child Characteristics and Public- and Private-Sector Schools

3.5.1 Midline

Our second research question was concerned with impact variation. We assessed if impacts on school readiness were moderated by three child characteristics: gender, baseline school readiness, and grade level (KG1 and KG2). We found no statistically significant interactions between treatment status for any child characteristics.

We then assessed if program impacts on teacher, classroom, and child outcomes were moderated by school status (i.e., public vs. private sector) (Appendix A Table 4). Of the eight primary outcomes assessed, we found two statistically significant interactions between treatment status and public- or private-sector schools, both in the domain of teacher professional well-being. First, there was a significant interaction between the TT and TTPA conditions and private-sector status in predicting levels of teacher burnout (b = -0.44, SE=.20, p < .05 and b = -0.48, SE=.20, p < .05, respectively). The results indicate that impacts on reduced burnout were larger in private schools.

Second, the interaction term predicting teacher turnover between private-school status and the TT condition was marginally statistically significant (b = -0.18, SE=.11, p < .07, and between private-school status and the TTPA was significant (b = -0.25, SE=.11, p < .05). Figure 3 illustrates the nature of these differences, showing the predicted probability of teacher turnover by treatment condition in private- and public-sector schools separately. The treatment reduced the predicted probability of teacher turnover from 43.5% to 12.3% (TT condition) and to 17.4% (TTPA condition). Notably, in private schools, the treatment reduced turnover to levels similar to the public sector.

3.5.2 Endline

At endline, there were no moderation effects for child gender, baseline scores, or grade level. Regarding impact variation by public vs. private status, there was one statistically significant interaction term between TT treatment status and school sector in predicting teacher burnout, showing (similar to midline) that reductions in burnout at endline were larger in private schools.

Section 4: Discussion

The impact evaluation of the QP4G program found that in the year the interventions were implemented, the in-service teacher training and coaching program significantly improved classroom quality (with moderate-sized impacts on the implementation of developmentally appropriate practices, and moderate-sized impacts on some dimensions of teacher-child interactions), had moderate to large impacts on reducing teacher burnout and teacher turnover, and had small impacts on improving children's school readiness (including literacy, numeracy, and social-emotional skills). Adding the parental-awareness intervention to the teacher training did not enhance these effects, and in fact counter-acted the impacts on child outcomes. One year after the intervention had ended, teachers were still using more developmentally appropriate practices and reported lower levels of burnout, but impacts on support for student expression (one dimension of classroom quality) were negative in both treatment conditions. Impacts on children's school readiness persisted, primarily for social-emotional skills.

The results hold promise for improving the quality of education delivered in Ghana's KG educational system and children's early learning outcomes when it comes to in-service teacher training. The training was built into existing education structures: the teacher trainers and training center, NNTTC, offered the five-day course in its center. In addition, coaches/monitors were district coordinators from the district government. Thus, expanding the full training (main training plus two refresher training, and regular coaching/monitoring) could be continued with national and district government support.

It is quite remarkable that one year later when children had transitioned to a new class, gains from the implementation year were sustained. A critical question concerns children's transition to primary school, where the curriculum and teaching approach changes abruptly to a teacher-driven and stricter classroom environment. Understanding how these changes in school and classroom contexts affect children development and learning, and how they interact with treatment condition, is an important next step for understanding the longer-term implications of the findings of this study.

In addition, it is hopeful to see that one year later, teachers who were trained were still using more of the developmentally appropriate practices as specified in the KG-specific pedagogy. A puzzling finding, however, is the *reduction* in support for student expression compared to the control group at endline. This will require further analysis to understand if teachers in particular types of school environments were more or less likely to reduce this element of their classroom after the training and district support ended (e.g., schools with more or less "enabling"

environments). The findings raise an important issue of how to provide teachers with continued, *annual* professional development with consistent directive, and may point to the critical role the coaching and monitoring played in changing teacher practice.

Regarding the parental-awareness training, this element of the findings raised more questions than it answered. Our conclusion is that parenting probably is a critical element in supporting children's early learning based on decades of research in child development, and recent evidence of the important role of parents in the SSA context (see Ozler et al., 2016). However, the QP4G approach was not the right way to engage parents. Research with peri-urban Ghanaian parents indicates that parents view preschool as a way to prepare children for primary school and place an emphasis on academic learning (Bidwell & Watine, 2014; Kabay et al., 2017). Perhaps parents did not agree with our program's messages and attempted to counter the changes in teachers' practices at home. Follow-up qualitative interviews with teachers in the parental-awareness treatment arm indicate that after the intervention, parents complained more to teachers about their child's behavior and academic problems. Or perhaps the counteracting effect was due to the content of the parental-awareness training itself. The training consisted of screened, staged videos in the local language of two mothers discussing the preschool education of their children, and featured the two different classrooms and teachers that were being discussed. It is possible that these videos did not relate to caregivers' experiences and, as a result, caused them to distance themselves from the schools and their child's education. Alternatively, it is possible that the training was not implemented with fidelity and that parents' experiences varied widely based on the district education coordinator who was implementing the program. Thus, our conjecture is not that parental-awareness training is harmful to children but rather that it must be done carefully by the right personnel and in a way that successfully conveys to parents the advantages of the new approaches. Notably, a recent study in Malawi found that a more intensive, 12-module, group-based parenting-support program administered through child-care centers by teachers and their mentors combined with intensive teacher training was effective in improving early childhood developmental outcomes (Özler et al., 2016), suggesting that parenting programs administered through schools by local (not district) personnel can be effective. However, it is possible that such programs need to have frequent enough meetings for parents to internalize the messages.

To our knowledge, this is one of the first studies in sub-Saharan Africa to assess impacts of a pre-primary school teacher training on observed classroom quality (measured via teacher-child interactions), as well as on multiple domains of children's school readiness (measured by direct

child assessments), including social-emotional and executive function outcomes (but see Özler et al., 2016, for an exception). The effect sizes we observed for teacher and classroom measures (d = .48 to 0.64) are in line with those found in other ECE interventions in the United States (e.g., Morris et al., 2014; Raver et al., 2008) and are consistent with the large effect-sizes found in pedagogical educational impact evaluations in sub-Saharan Africa (Conn, 2017; Yoshikawa et al., 2015). Part of the discrepancy is likely due to the treatment contrast in LMICs compared to the United States. The effect sizes observed for child outcomes (d = .15 to .17) are in line with related ECE interventions in the United States. This suggests that future initiatives should focus on how to translate the large improvements in teaching quality to larger impacts in learning outcomes if early education strategies are to have the dramatic effects required to help all children learn adequately. The QP4G study was designed with national scalability in mind, thus limiting the intensity and cost of the training. The tension between achieving large impacts and creating an intervention that can feasibly be implemented at scale is one that the field must continue to grapple with.

Section 5: Implications for Policy Decision and Next Steps

5.1 Cost-effectiveness analysis

The costs and cost-effectiveness analysis (see Appendix B) presented in this report reveals critical information needed to consider scaling-up the QP4G program. The program cost about $135,506.91 to implement the eight-day in-service teacher training program with monitoring and coaching visits in the 161 schools across six municipalities in the Greater Accra Region. About 67% of these costs were attributed to the training itself. The average total costs per school, teacher, and child were approximately $842, $402, and $16. When considering budgetary expenditures only (excluding resource and time costs of participants), expenditures per school, teacher, and child were $512.35, $244,77, and $9.79, respectively.

We then consider the cost-effectiveness ratios (CERs), with estimated effects sizes of 0.54 and 0.15 by the end of the school year in 2016 and 0.21 and 0.11 a year later for improved implementation of developmentally appropriate pedagogy and school readiness among children, respectively. To compare the QP4G program in terms of cost-effectiveness with other programs, the CERs in Appendix B Table 4 need to be compared with the CERs for the other programs. The QP4G program is more (less) cost-effective than other programs that have CERs greater (less) than those in these tables. Overall, the program would be more cost-effective in increasing school readiness and the implementation of KG-specific pedagogy than any other program with a cost-effectiveness ratio (CER) of a one standard-deviation improvement of more than $107.23 per child and $744.63 per teacher, respectively, per year. The CERs for the QP4G program at first glance may appear more favorable to the implementation of developmentally appropriate pedagogy in schools than the improvement of school readiness because of the higher effect sizes reported for the former. But that inference needs to be heavily qualified because there were not two independent programs for developmentally appropriate pedagogy in schools and for improvement of child primary school readiness. Instead there was a package that affected both of these outcomes. The cost and CER estimates for the QP4G program are sensitive to the underlying assumptions and would increase as reflected in Appendix B Table 3 and Appendix B Table 4 were there to be increases in travel and transportation costs in any scale-up process. Finally, comparisons between the CER (Appendix B Table 4) and BER (Appendix B Appendix Table 2) estimates illustrate how under the assumptions of direct expenditures excluding time costs for teachers or broader direct expenditures with no teacher, coaches, and administrative personnel payments, governmental expenditures would be less

than resource costs, and by how much they therefore would be downward biased estimates of resource costs.

5.2 Policy engagement strategy

5.2.1 QP4G Steering Committee

The primary function of the Steering Committee was to take responsibility for the feasibility of the QP4G study and ensure that it is moving in a direction consistent with the overall plans of the GES and MOE. The Steering Committee monitored and reviewed the project status, as well as provided feedback on various instruments used during the implementation of the project. Furthermore, the Steering Committee provided key insight into ways in which lessons from the study can effectively compliment initiatives being made by the GES and MOE.

A total of seven meetings were organized from June 2015 to July 2017. All identified stakeholders participated in all planned Steering Committee meetings and monitoring visits to treatment schools. The establishment and continuous meeting of the Steering Committee created an open dialogue between government entities and private school associations who seemed to have strong working relationships.

5.2.2 Working with the public sector

IPA established a strong working relationship with the GES and the NNTTC – both of which have shown great interest and ownership in the project, closely monitoring and reviewing the implementation reports to ascertain the best lessons that the education system can adopt from the project.

IPA engaged the Teacher Education Division (TED), the ECE Unit, and the Curriculum Research and Development Division (CRDD) of GES to (i) incorporate some of the study's lessons into the teacher professional development system as a whole; and (ii) support TED to finalize the National ECE Framework being developed for In-Service KG Teacher training (GES through TED has invited IPA to sit on the Committee to draft the ECE Framework.) In addition, the School Supervisor Training has been shared with TED in early May 2017.

NNTTC has included the QP4G training model to their suites of trainings on offer to teachers. Going forward, IPA will be supporting NNTTC to develop a roll out/implementation strategy which get the training to be delivered to both public and private school teachers in Ghana.

Finally, IPA worked closely with the MoE through the National Teaching Council (formerly Teacher Education Division) to develop the National Framework for KG Teacher Training in

Ghana, which is being printed by UNICEF for distribution in UNICEF focused districts. The engagement is on-going and dependent on the time government will finally rollout the Teacher Licensing Policy.

5.2.3 Engaging the private sector

IPA's strategy is to increase its engagement with the private school associations through the promising results from the QP4G program. The increases in teacher retention over the course of the school year suggest that investing in the training would pay off for private schools.

Private school organizations, including the Early Childhood Development Care Association of Ghana (ECDCAG) network, have expressed their willingness to use the QP4G training. IPA participated in the National Congress of the ECDCAG in Takoradi in November 2017 with follow-up on discussions to adopt the training in some private schools in their network.

IPA is also exploring the option of building capacity at the NNTTC and TED to deliver the QP4G teacher training to private sector schools beyond the lifetime of the project - both within and outside of the Greater Accra Region.

5.3.4 Working with donors and large-scale project implementing organizations

Development partners, including FHI360/USAID Learning and IDP Rising Schools/Sesame Street, have invited IPA to share materials, evaluation results, and expertise in order to inform their programs.

The USAID Learning Project which focuses on developing literacy skills in pupils from KG to P3 has taken interest in adopting training materials: the midline results and training materials have been shared with the USAID team and the IPA policy team met the consultant in charge of developing the Learning Project material.

IPA is sitting on a Committee to inform the design of a Sesame Street/IDP Rising Schools program on early grades (KG and lower primary). The first meeting occurred in June 2017.

5.4 Dissemination events

IPA has organized and participated in several conferences and events on Education in Ghana to present the findings of the QP4G study. Several policy briefs and presentations summarizing the results and policy implications of the evaluation were produced and shared during these events. The main dissemination events attended were:

a. *Evidence in Education Summit organized by IPA in partnership with MOE in March 2017*. IPA partnered with the Ghana's Ministry of Education and the Ghana Education

Service to host a summit gathering policymakers and researchers to engage existing rigorous evidence about education in Ghana. The agenda of the first day was to disseminate relevant research and discuss how to apply it to policy; attendees engaged enthusiastically on both goals through research presentations and panels. From an evidence-informed policymaking perspective, a major achievement of the event was in how invested Ghanaian policymakers were in its success. Findings from the QP4G study were presented by Sharon Wolf. The presentation emphasized several lessons for ECE in Ghana: the importance of refresher training, coaching and monitoring, for example, as a way of encouraging teachers to implement lessons from trainings.

b. The midline results were also also presented during the UNICEF/GES Stakeholders Dissemination Forum on KG Formative Assessment on the 7th of Sept. 2017 in Accra.

c. *Dissemination event organized by IPA in October 2017.* IPA organized a dissemination event in October 2017 in order to share the final findings of the study and policy implications for the public and private ECD sector in Ghana. The event sought to mobilize stakeholders and determine new opportunities the QP4G program can influence.

d. *Evidence for Education Summit organized by IPA Kenya.* Our paper submission was recently accepted to the IPA Kenya Evidence for Education (E4A) conference in Nairobi, Kenya December 5-7th, 2017. Bridget Gyamfi of IPA Ghana presented the QP4G project and findings.

5.5 Scholarly progress

To date, several publications are either published, under revision, or in preparation based on data collected for this study. Below, we list those studies.

1. Wolf, S., Aber, J.L., & Behrman, L. (under review). Experimental Evaluation of the 'Quality Preschool for Ghana' Intervention on Teacher Professional Well-Being, Classroom Quality and Children's School Readiness. *Journal of Research in Educational Effectiveness.*

2. McCoy, D.C., & Wolf, S. (revision under review). Changes in Classroom Quality Predict Ghanaian Preschoolers' Gains in Academic and Social-Emotional Skills. *Developmental Psychology.*

3. Wolf, S., & McCoy, D. C. (2017). Household Socioeconomic Status and Parental Investments: Direct and Indirect Relations with School Readiness in Ghana. *Child Development.* [E-pub ahead of print].

4. Wolf, S., Raza, M., Kim, S., Aber, J.L, Behrman, J., & Seidman, E. (revision under review). Measuring classroom process quality in pre-primary classrooms in Ghana using the TIPPS. *Early Childhood Research Quarterly.*

5. Chan, W. (revision under review). The Relation Between Power in Normal and Binomial Outcomes in Cluster Randomized Trials. *Psychological Methods.*

6. Pesando, L., Wolf, S., Behrman, J., Tsinigo, E., Aber, J. (under review) Public versus Private? Examining Kindergarten School Choice, Parental Resources, and Children's School Readiness in Ghana. *International Journal of Educational Development.*

5.6 Next steps

The next steps involve conducting a longer-term follow-up testing longer lasting effects of the training in primary school and testing the replicability of the program in a rural context.

a. **QP4G Phase II:** IPA has secured funds to launch a longer-term follow-up of the QP4G study. This longer-term follow-up seeks to follow the KG children in their transition to primary school to assess if there are (i) longer-term impacts of high quality ECE (vs. fade-out), (ii) differences in longer-term impacts based on primary school characteristics, and (iii) differences in longer-term impacts by child age and gender and by select parental characteristics.

b. **Rural QP4G**: IPA proposes to (i) adapt the QP4G training to increase the emphasis on teaching academic outcomes (i.e., early literacy and numeracy) in addition to social and behavioral outcomes, and (ii) to meet the needs of teachers and children in rural areas of the country. By extending the model to resource-poor rural communities and with a focus on teacher professional well-being and improved teacher-child interaction quality, the program aims to ultimately improve the KG sector on a national scale to increase children's readiness for primary school across the country.

References

Alderman, H., Orazem, P., & Paterno, E. (2001). School quality, school cost, and the public/private school choices of low-income households in Pakistan. *The Journal of Human Resources*, *36*(2), 304.

Bennell, P., & Akyeampong, K. (2007). *Teacher motivation in Sub-Saharan Africa and South Asia*. London: DfID.

Bidwell, K., & Watine, L. (2014). Exploring early education problems in peri-urban settings in Africa: Final Report. Newhaven, CT: Innovations for Poverty Action.

Chan, W. (revision under review). The relation between power in normal and binomial outcomes in cluster randomized trials. *Psychological Methods.*

Collins, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research*, *12*(3), 229–238.

Conn, K. M. (2017). Identifying effective education interventions in Sub-Saharan Africa: A meta-analysis of impact evaluations. *Review of Educational Research, 87(5),* 863-898.

Educational Management Information System (EMIS). (2004). Report on basic statistics and planning parameters for basic education in Ghana 2003/2004. Accra: Ministry of Education.

Educational Management Information System (EMIS). (2015). Report on basic statistics and planning parameters for basic education in Ghana 2014/2015. Accra: Ministry of Education.

Grisay, A., & Mahlck, L. (1991). The quality of education in developing countries: A review of some research studies and policy documents. Issues and methodologies in educational development: An IIEP Series for Orientation and Training, 3. Paris: International Institute of Educational Planning

Hedges, L. V. (2009). Effect sizes in nested designs. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The Handbook of Research Synthesis and Meta-analysis* (pp. 337-356). New York, NY: Russell Sage Foundation.

Kabay, S., Wolf, S., & Yoshikawa, H. (2017). “So that his mind will open”: Parental perceptions of preschool in urbanizing Ghana. *International Journal of Educational Development, 57*, 44-53.

Maslach, C., Jackson, S. E., & Leiter, M. P. (1996). Maslach burnout inventory manual. 3rd ed. Palo Alto, CA: Consulting Psychologists Press.

McCoy, D.C., & Wolf, S. (under review). Changes in classroom quality predict Ghanaian preschoolers' gains in academic and social-emotional skills. *Developmental Psychology.*

Ministry of Education (MOE). (2016) Education sector performance report 2016. Accra: Ministry of Education.

Morris, P., Mattera, S. K., Castells, N., Bangser, M., Bierman, K., & Raver, C. (2014). *Impact findings from the Head Start CARES demonstration: National evaluation of three approaches to improving preschoolers' social and emotional competence* (OPRE Report 2014-44). New York: MDRC.

Osei, G.M. (2006). Teachers in Ghana: Issues of training, remuneration and effectiveness. *International Journal of Educational Development*, *26*, 38–51.

Özler, B., Fernald, L. C. H., Kariger, P., McConnell, C., Neuman, M., & Fraga, E. (2016). *Combining preschool teacher training with parenting education: A cluster-randomized controlled trial* (Policy Research Working Paper 7817). Washington, DC: World Bank, 2016.

Pesando, L., Wolf, S., Behrman, J., Tsinigo, E., Aber, J. (under review) Public versus private? Examining kindergarten school choice, parental resources, and children's school readiness in Ghana. *Demography*

Pisani, L., Borisova, I., & Dowd, A. J. (2015). International development and early learning assessment technical working paper. Save the Children. Washington, DC. Available at: http://resourcecentre.savethechildren.se/library/international-development-and-early-learning-assessment-technical-paper

Raver, C. C., Jones, S. M., Li-Grining, C. P., Metzger, M., Champion, K. M., & Sardin, L. (2008). Improving preschool classroom processes: Preliminary findings from a randomized trial implemented in Head Start settings. *Early Childhood Research Quarterly*, *23*(1), 10–26.

Seidman, E., Raza, M., & Kim, S. (2017). *Assessment of pedagogical practices and processes in low & middle income country secondary school classrooms: Findings from Uganda.* Manuscript submitted for publication.

Seidman, E., Raza, M., Kim, S., & McCoy, J. M. (2013). *Teacher instructional practices and processes system (V.5) – TIPPS: Manual and scoring system.* New York: New York University.

Torrente, C., Aber, J. L., Witteveen, D., Gupta, T., Johnston, B., Shivshanker, A., Annan, J., & Bundervoet, T. (2012). Baseline report: Teacher survey results. New York: New York University.

Wolf, S., & McCoy, D. C. (2017). Household Socioeconomic Status and Parental Investments: Direct and Indirect Relations with School Readiness in Ghana. Child Development. [E-pub ahead of print].

Wolf, S., & McCoy, D. C. (2017). Household socioeconomic status and parental investments: direct and indirect relations with school readiness in Ghana. *Child Development.* [E-pub ahead of print].

Wolf, S., Aber, J. L., Torrente, C., McCoy, M., & Rasheed, D. (2015). Cumulative risk and teacher well-being in the Democratic Republic of the Congo. *Comparative Education Review, 59(4)*, 717-742.

Wolf, S., Aber, J.L., & Behrman, L. (under review). Experimental evaluation of the 'quality preschool for ghana' intervention on teacher professional well-being, classroom quality and children's school readiness. *Journal of Research in Educational Effectiveness.*

Wolf, S., Raza, M., Kim, S., Aber, J.L, Behrman, J., & Seidman, E. (revision under review). Measuring classroom process quality in pre-primary classrooms in Ghana using the TIPPS. *Early Childhood Research Quarterly.*

Wolf, S., Torrente, C., Frisoli, P., Weisenhorn, N., Shivshaker, A., Annan, J., & Aber, J. L. (2015). Preliminary impacts of a professional development intervention on teacher wellbeing in the Democratic Republic of the Congo. *Teaching and Teacher Education, 51,* 24–36.

Yoshikawa, H., Leyva, D., Snow, C. E., Treviño, E., Barata, M., Weiland, C., ... & Arbour, M. C. (2015). Experimental impacts of a teacher professional development program in Chile on preschool classroom quality and child outcomes. *Developmental Psychology*, *51*(3), 309-322.

Appendices

Appendix A

Appendix A Table 1. Sample Attrition Across All Waves

	Children			Teachers		
	Baseline	Midline	Endline	Baseline	Midline	Endline
	N (% of baseline sample)					
Control	1,180	931 (78.9%)	812 (68.8%)	139	97 (69.8%)	89 (64.0%)
TT	1,167	1,025 (87.8%)	926 (79.4%)+	155	128 (82.6%)*	110 (71.0%)
TTPA	1,088	1,019 (94.7%)	919 (84.4%)+	150	122 (81.3%)+	110 (73.3%)
Total	3,435	2,975	2,657	444	347	309

+ $p < .07$.

Appendix A Table 2. Impacts on Teacher Professional Wellbeing and Classroom Quality at Midline and Endline

	Midline				Endline			
	b	se	p-value	*effect size (d_{wt})*	b	se	p-value	*effect size (d_{wt})*
Motivation								
Teacher	0.105	(0.061)	0.085	0.352^{+}	-0.014	(0.065)	0.835	-0.052
Teacher + Parent	0.036	(0.061)	0.549	0.121	0.020	(0.065)	0.761	0.076
Burnout								
Teacher	-0.329	(0.121)	0.007	**-0.396****	-0.170	(0.120)	0.155	-0.187
Teacher + Parent	-0.453	(0.121)	0.000	**-0.546*****	-0.264	(0.120)	0.027	**-0.290***
Job Satisfaction								
Teacher	0.088	(0.085)	0.302	0.221	0.009	(0.085)	0.914	0.023
Teacher + Parent	-0.051	(0.085)	0.547	-0.129	-0.064	(0.085)	0.454	-0.159
Fidelity Checklist								
Teacher	1.253	0.209	0.000	**0.535*****	0.476	0.189	0.012	**0.214***
Teacher + Parent	1.412	0.209	0.000	**0.603*****	0.487	0.189	0.010	**0.219****
Facilitating Deeper Learning								
Teacher	-0.062	0.096	0.520	-0.154	-0.111	0.085	0.187	-0.256
Teacher + Parent	-0.079	0.096	0.412	-0.197	-0.142	0.084	0.093	-0.327^{+}
Supporting Student Expression								
Teacher	0.215	0.096	0.025	**0.479***	-0.168	0.081	0.038	**-0.398***
Teacher + Parent	0.095	0.090	0.325	0.212	-0.304	0.081	0.000	**-0.720*****
Emotional Support & Behavior Management								
Teacher	0.162	0.052	0.002	**0.617*****	0.069	0.045	0.117	0.582
Teacher + Parent	0.167	0.052	0.001	**0.636*****	0.038	0.045	0.352	0.324
Sample size = 444								

Notes. Estimates are computed using observed scores, in two level models: teachers nested in schools. Effect sizes calculated accounting for the 2-level model structure (Hedges, 2009).

+ p <.10, *p < .05, **p < .01, ***p < .001

TT = Teacher training condition; TTPA = Teacher training plus parent awareness training condition.

Models include the following control variables: private (vs. public) sector status of the school, six district dummies, a dummy variable for if the school was assigned to receive teacher text messages, a dummy for if the school was assigned to receive parent flyers, a series of five dummy variables accounting for within-sample mobility, teacher gender, age, level of education, years of teaching experience. Models for teacher professional well-being outcomes also include the baseline score for each respective outcome.

All impact estimates computed from 20 multiply imputed datasets.

[a] Odd ratio presented rather than effect size.

Appendix A Table 3. Impacts on Children's School Readiness at Midline and Endline

	Midline				Endline			
	B	Se	p-value	*effect size (d_{wt})*	b	se	p-value	*effect size (d_{wt})*
Composite IDELA score								
Teacher	0.023	0.011	0.033	**0.149***	**0.013**	**0.008**	**0.093**	**0.109+**
Teacher + Parent	0.000	0.010	0.999	0.000	-0.007	0.008	0.330	-0.062
Early numeracy								
Teacher	0.020	0.009	0.038	**0.106***	0.007	0.009	0.437	0.043
Teacher + Parent	-0.006	0.009	0.503	-0.034	-0.009	0.009	0.289	-0.058
Early literacy								
Teacher	0.027	0.014	0.045	**0.136***	0.016	0.012	0.184	0.091
Teacher + Parent	0.004	0.013	0.783	0.018	-0.007	0.012	0.575	-0.038
Social-emotional								
Teacher	0.032	0.014	0.017	**0.166***	**0.023**	**0.011**	**0.031**	**0.138***
Teacher + Parent	0.021	0.013	0.125	0.105	0.001	0.011	0.982	0.006
Executive function								
Teacher	0.010	0.013	0.448	0.054	**0.017**	**0.010**	**0.099**	**0.106+**
Teacher + Parent	-0.013	0.013	0.314	-0.071	-0.010	0.011	0.350	-0.062
Sample size = 3,435								

Notes. Estimates are computed using observed scores, in three level models: children nested in classrooms nested in schools. Effect sizes calculated accounting for the 3-level model structure (Hedges, 2009).

Sample includes children present at baseline and follow-up.

* $p < .05$.

TT = Teacher training condition; TTPA = teacher training plus parent awareness training condition.

Models include the following control variables: private (vs. public) sector status of the school, six district dummies, a dummy variable for if the school was assigned to receive teacher text messages, a dummy for if the school was assigned to receive parent flyers, a series of five dummy variables accounting for within-sample mobility, child gender, age, KG level (1, 2, or 3 if KG1 and KG2 were combined in one classroom, as a categorical variable), and baseline score for each respective outcome.

All impact estimates computed from 100 multiply imputed datasets.

Appendix A Table 4. Moderation of Treatment Impacts on Outcomes, By Public and Private Sector Status at Midline

	B	(SE)	p-value
Teacher professional well-being			
Motivation			
TT	0.109	(0.083)	0.193
TTPA	-0.001	(0.081)	0.990
Private school	-0.108	(0.073)	0.145
TT*Private	-0.005	(0.100)	0.957
TTPA*Private	0.072	(0.101)	0.475
Burnout			
TT	-0.081	(0.164)	0.621
TTPA	-0.196	(0.159)	0.216
Private school	0.347	(0.144)	0.016*
TT*Private	-0.439	(0.197)	0.026*
TTPA*Private	-0.483	(0.199)	0.015*
Job satisfaction			
TT	0.171	(0.116)	0.139
TTPA	-0.010	(0.112)	0.930
Private school	-0.009	(0.102)	0.927
TT*Private	-0.144	(0.139)	0.301
TTPA*Private	-0.067	(0.140)	0.632
Teacher turnover			
TT	-0.023	(0.084)	0.786
TTPA	0.038	(0.081)	0.644
Private school	0.349	(0.074)	0.000
TT*Private	-0.186	(0.101)	0.065
TTPA*Private	-0.255	(0.102)	0.012
Classroom processes			
Facilitating deeper learning			
TT	-0.002	(0.131)	0.988
TTPA	0.015	(0.127)	0.909
Private school	0.298	(0.116)	0.010
TT*Private	-0.108	(0.158)	0.492
TTPA*Private	-0.188	(0.159)	0.236
Supporting student expression			
TT	0.340	(0.129)	0.008
TTPA	0.064	(0.125)	0.610
TT*Private	-0.173	(0.155)	0.264
TTPA*Private	-0.016	(0.156)	0.919
Emotional support & behavior management			
TT	0.140	(0.073)	0.056
TTPA	0.151	(0.071)	0.032
Private school	0.030	(0.064)	0.639
TT*Private	0.052	(0.088)	0.553
TTPA*Private	0.049	(0.089)	0.578
Child school-readiness composite			
TT	0.008	(0.014)	0.536
TTPA	-0.007	(0.013)	0.611

Private school	0.020	(0.013)	0.123
TT*Private	0.025	(0.018)	0.153
TTPA*Private	0.014	(0.017)	0.436

TT = Teacher training condition; TTPA = teacher training plus parent awareness training condition.

Models include the following control variables: private (vs. public) sector status of the school, six district dummies, a dummy variable for if the school was assigned to receive teacher text messages, a dummy for if the school was assigned to receive parent flyers, a series of five dummy variables accounting for within-sample mobility, child gender, age, KG level (1, 2, or 3 if KG1 and KG2 were combined in one classroom, as a categorical variable), and baseline score for each respective outcome.

All impact estimates computed from multiply imputed datasets.

Impact estimates with no covariates included in models

Appendix A Table 5. Individual- and School-level Impact Estimates at Midline with No Covariates and Cluster-adjusted Standard Errors

	Individual-level			School-level		
	b	SE	p-value	b	SE	p-value
Professional wellbeing						
Motivation						
Teacher	0.129	0.060	0.034	0.151	0.062	0.015
Teacher + Parent	0.004	0.068	0.956	0.014	0.062	0.819
Burnout						
Teacher	-0.340	0.131	0.010	-0.390	0.129	0.003
Teacher + Parent	-0.219	0.131	0.095	-0.250	0.129	0.053
Job satisfaction						
Teacher	0.066	0.107	0.536	0.042	0.101	0.676
Teacher + Parent	0.066	0.022	0.835	-0.015	0.101	0.879
Classroom quality						
Fidelity Checklist						
Teacher	1.522	0.218	0.000	1.632	0.239	0.000
Teacher + Parent	1.577	0.220	0.000	1.426	0.252	0.000
Facilitating Deeper Learning						
Teacher	0.055	0.092	0.552	0.072	0.098	0.461
Teacher + Parent	-0.021	0.089	0.818	-0.031	0.098	0.755
Supporting Student Expression						
Teacher	0.250	0.093	0.008	0.290	0.096	0.003
Teacher + Parent	0.164	0.093	0.079	0.144	0.096	0.134
Emotional Support & Behavior Management						
Teacher	0.201	0.053	0.000	0.191	0.053	0.000
Teacher + Parent	0.175	0.055	0.002	0.161	0.054	0.003
Child outcomes						
School readiness						
Teacher	0.024	0.012	0.051	0.029	0.013	0.024
Teacher + Parent	0.016	0.012	0.187	0.021	0.013	0.104
Numeracy						
Teacher	0.021	0.014	0.128	0.029	0.015	0.046
Teacher + Parent	0.012	0.014	0.376	0.021	0.015	0.160
Literacy						
Teacher	0.023	0.017	0.183	0.025	0.019	0.177
Teacher + Parent	0.009	0.018	0.609	0.015	0.019	0.439
Social-emotional						
Teacher	0.035	0.013	0.009	0.040	0.014	0.003
Teacher + Parent	0.031	0.014	0.021	0.037	0.014	0.008
Executive Function						
Teacher	0.018	0.013	0.164	0.023	0.013	0.093
Teacher + Parent	0.013	0.013	0.310	0.013	0.014	0.337

Notes. TT = Teacher training condition; TTPA = teacher training plus parent awareness training condition.
No control variables included in models.
Analyses use unimputed data, with teachers and children who were present in the sample at baseline and each subsequent wave. School-level outcomes calculated by aggregating the mean score for either teacher or child outcomes in each school.

Appendix A Table 6. Individual- and School-level Impact Estimates at Midline with No Covariates and Cluster-adjusted Standard Errors

	Individual-level			School-level		
	b	SE	p-value	b	SE	p-value
Professional wellbeing						
Motivation						
Teacher	-0.015	0.063	0.810	-0.046	0.072	0.522
Teacher + Parent	-0.098	0.068	0.150	-0.081	0.072	0.266
Burnout						
Teacher	-0.06	0.134	0.663	-0.056	0.146	0.700
Teacher + Parent	-0.052	0.136	0.702	-0.008	0.147	0.959
Job satisfaction						
Teacher	-0.068	0.108	0.529	-0.063	0.102	0.538
Teacher + Parent	-0.105	0.111	0.351	-0.042	0.102	0.679
Classroom quality						
Fidelity Checklist						
Teacher	1.140	0.243	0.000	1.209	0.243	0.000
Teacher + Parent	0.940	0.246	0.000	0.974	0.245	0.000
Facilitating Deeper Learning						
Teacher	0.030	0.118	0.802	0.124	0.118	0.293
Teacher + Parent	-0.083	0.114	0.467	-0.060	0.119	0.613
Supporting Student Expression						
Teacher	0.010	0.093	0.916	0.045	0.094	0.636
Teacher + Parent	-0.152	0.091	0.095	-0.112	0.095	0.240
Emotional Support & Behavior Management						
Teacher	0.108	0.053	0.044	0.117	0.058	0.046
Teacher + Parent	0.101	0.057	0.080	0.101	0.059	0.087
Child outcomes						
School readiness						
Teacher	0.013	0.011	0.224	0.017	0.011	0.133
Teacher + Parent	0.000	0.011	0.981	0.005	0.011	0.684
Numeracy						
Teacher	0.006	0.012	0.592	0.015	0.013	0.231
Teacher + Parent	-0.004	0.013	0.760	0.005	0.013	0.681
Literacy						
Teacher	0.018	0.016	0.274	0.021	0.017	0.205
Teacher + Parent	0.003	0.017	0.841	0.008		0.610
Social-emotional						
Teacher	0.021	0.011	0.050	0.022	0.012	0.061
Teacher + Parent	0.007	0.011	0.567	0.009	0.012	0.440
Executive Function						
Teacher	0.007	0.012	0.525	0.009	0.012	0.433
Teacher + Parent	-0.005	0.012	0.681	-0.004	0.012	0.703

Notes. TT = Teacher training condition; TTPA = teacher training plus parent awareness training condition.
No control variables included in models.
Analyses use unimputed data, with teachers and children who were present in the sample at baseline and each subsequent wave. School-level outcomes calculated by aggregating the mean score for either teacher or child outcomes in each school.

Appendix A Table 7. Impact Estimates on Teacher Professional Wellbeing and Classroom Quality at Midline and Endline with District and Private Sector Status Indicators Only

	Midline				Endline			
	b	SE	p-value	effect size (dwt)	b	SE	p-value	effect size (dwt)
Professional wellbeing								
Motivation								
Teacher	0.131	0.051	0.009	**0.439**	0.015	0.054	0.788	0.057
Teacher + Parent	0.058	0.051	0.259	0.194	-0.001	0.054	0.979	-0.004
Burnout								
Teacher	-0.341	0.107	0.001	**-0.411**	-0.152	0.11	0.167	-0.167
Teacher + Parent	-0.264	0.108	0.014	**-0.318**	-0.187	0.111	0.092	-0.205
Job satisfaction								
Teacher	0.055	0.107	0.496	0.139	0.004	0.076	0.962	0.010
Teacher + Parent	-0.010	0.022	0.907	-0.025	-0.013	0.077	0.870	-0.032
Classroom quality								
Fidelity Checklist								
Teacher	1.334	0.218	0.000	**0.570**	0.703	0.159	0.000	**0.316**
Teacher + Parent	1.437	0.220	0.000	**0.614**	0.578	0.160	0.000	**0.260**
Facilitating Deeper Learning								
Teacher	-0.038	0.092	0.639	-0.095	-0.097	0.072	0.175	-0.223
Teacher + Parent	-0.120	0.089	0.140	-0.300	-0.165	0.072	0.022	**-0.380**
Supporting Student Expression								
Teacher	0.163	0.093	0.036	**0.364**	-0.087	0.068	0.200	-0.206
Teacher + Parent	0.040	0.093	0.608	0.089	-0.254	0.069	0.000	**-0.603**
Emotional Support & Behavior Management								
Teacher	0.166	0.053	0.000	**0.632**	0.091	0.039	0.019	0.768
Teacher + Parent	0.140	0.055	0.002	**0.533**	0.074	0.039	0.059	0.624
Sample size = 444								

Notes. These results include covariates for stratification variables (district dummies and private/public status) only.

TT = Teacher training condition; TTPA = teacher training plus parent awareness training condition.

All impact estimates computed from 20 multiply imputed datasets.

Appendix A Table 8. Impact Estimates on Child School Readiness Outcomes at Midline and Endline with District and Private Sector Status Indicators Only

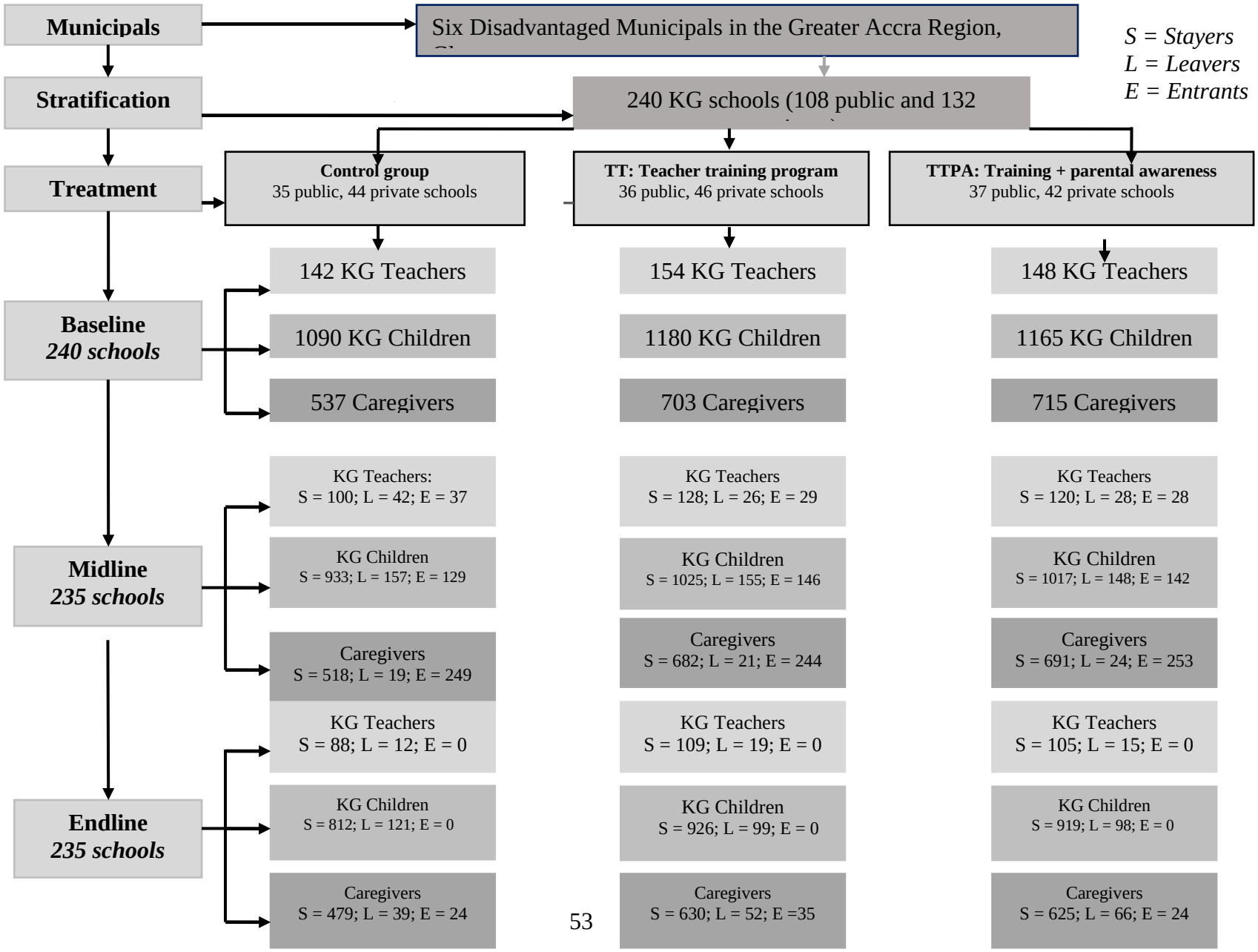
	Midline				Endline			
	b	se	p-value	*effect size (d_{wt})*	b	se	p-value	*effect size (d_{wt})*
Composite IDELA score								
Teacher	0.024	0.012	0.053	**0.155**	0.012	0.009	0.177	0.103
Teacher + Parent	0.013	0.012	0.286	0.084	0.004	0.009	0.627	0.034
Early numeracy								
Teacher	0.023	0.014	0.107	0.124	0.010	0.012	0.377	0.064
Teacher + Parent	0.011	0.014	0.435	0.059	0.001	0.011	0.907	0.006
Early literacy								
Teacher	0.026	0.016	0.110	0.131	0.018	0.014	0.194	0.099
Teacher + Parent	0.016	0.016	0.328	0.078	0.008	0.013	0.537	0.045
Social-emotional								
Teacher	0.037	0.013	0.004	**0.191**	0.027	0.010	0.009	**0.159**
Teacher + Parent	0.036	0.013	0.005	**0.185**	0.016	0.010	0.120	0.095
Executive function								
Teacher	0.008	0.012	0.519	0.043	0.007	0.010	0.472	0.044
Teacher + Parent	-0.011	0.012	0.379	-0.059	-0.001	0.010	0.948	-0.006
Sample size = 3,435								

Notes. These results include covariates for stratification variables (district dummies and private/public status) only.

TT = Teacher training condition; TTPA = teacher training plus parent awareness training condition.

All impact estimates computed from 100 multiply imputed datasets.

Appendix A Figure 1: Sample flow chart



Appendix B

Improving the implementation of developmentally appropriate pedagogy and child primary school readiness in Ghana: A resource cost analysis for the Quality Preschool for Ghana (QP4G) Program

by

Edward Tsinigo[a], Jere R. Behrman[b], and Sharon Wolf[c]

25 October 2017

[a]Senior Research Associate, Innovations for Poverty Action, Dzorwulu, Accra, Ghana.

[b]W.R. Kenan Jr. Professor of Economics and Sociology, University of Pennsylvania, Philadelphia, PA, USA.

[c]Assistant Professor, Graduate School of Education, University of Pennsylvania, Philadelphia, PA, USA.

Summary

- We estimate total resource costs (including the time value of participants and trainers) and direct budgetary expenditures for the Quality Preschool for Ghana (QP4G) in-service teacher training and coaching program.
- The total resource costs of implementing the QP4G program per school, teacher, and child were $841.66, $402.10, and $16.08, respectively.
- When considering direct budget expenditures only, the total costs of implementing the QP4G program per school, teacher, and child were $512.35, $244,77, and $9.79, respectively.
- The QP4G program would be more cost-effective in increasing school readiness and the implementation of KG-specific pedagogy than any other program with a cost-effectiveness ratio (CER) of a one standard-deviation improvement of more than $107.23 per child and $744.63 per teacher, respectively, per year.

1.1 Introduction

There is considerable literature on assessing the benefits of interventions at various levels of education. The present analysis builds on evidence from a cluster randomized trial, the Quality Preschool for Ghana (QP4G) program[6], which had significantly positive effects on some important components of child development, classroom quality, and teacher satisfaction. But to

[6] We acknowledge the funding provided by UBS Optimus Foundation and The Strategic Impact Evaluation Fund (SIEF) for the QP4G program. The analysis and the conclusions drawn are solely the responsibility of the authors.

understand fully whether an intervention is justified from a public policy perspective or should be expanded, knowledge of the impacts alone is not sufficient. Equally important are the costs of the intervention. Two interventions, for example, may have the same effects on learning, but one may cost twice as much as the other – ceteris paribus, clearly, the latter is preferred. Or, for another example, one intervention may have 50% greater effects than another but cost twice as much, and therefore ceteris paribus be less cost-effective. All too often, there is emphasis purely on the impacts, without attention to the costs – thus leading to little guidance about whether an intervention is desirable. Therefore, in this report, we estimate costs for the QP4G program in order to provide insight for judging the merits of this intervention.

The relevant costs are the resource costs incurred to undertake the intervention. The resource costs are the opportunity costs of resources used to undertake an intervention in terms of what is foregone by this use. Note that the resource costs, in general, are <u>not</u> the same as current expenditures by providers, though often studies in the literature assume that they are. This is for several reasons:

a. Usually, there are user costs, e.g., the time that users have to devote to benefit from the intervention. If these are not included in the estimation of the costs, then the true costs are likely to be underestimated and the estimates are likely to be biased towards interventions that impose greater burdens on users if they also reduce supplier costs. If the users are relatively poor members of society, shifting costs to users is likely to be regressive.

b. Expenditures by suppliers for some important programs, e.g., conditional cash transfers or unconditional transfers, include not only resource costs but also transfers to program participants. Transferring purchasing power from one individual to another typically incurs some resource costs (e.g., the costs of the time of persons engaged in making such transfers), but these usually are small in comparison with the total expenditures. While transfers are a real budgetary item, they themselves are not resource costs. If Kwame gives Mary $100 (whether directly or through the public sector), for example, the only resource cost is the use of resources to make the transfer, not the amount of the transfer itself.

c. If governments tax people to obtain revenues to finance an intervention, then the taxes may distort incentives for people's behaviors so that society uses resources less efficiently. These distortion costs often are thought to be on the order of magnitude of 20-25% of public expenditures [1].

d. Some of the additional resource costs of an intervention may be incurred in the future. If a preschool program increases eventual schooling of the children who are affected, for example, then there may be additional resource costs years in the future to provide the teachers, materials, etc. for that additional schooling.

To empirically measure resource costs is challenging, in part because of these considerations. But that is not a good reason to ignore resource costs since, again, they are critical in coming to a judgement about the desirability of an intervention. Therefore, we proceed to make the best estimates that we can for the QP4G program. While resource costs, as noted, are not the same as the provider expenditures, the provider expenditures typically include some major components of the resource costs, and that is the case for the QP4G program, in part because this is not a transfer program so there is no confounding of transfers in the expenditures. For comparison with the resource costs, we also estimate under alternative assumptions the impacts on supplier budgetary expenditures below.

1.2 Description of the QP4G Program

Most kindergarten teachers in private schools and a number in public schools[7] in Ghana are untrained [2]. Many have only a primary education, leaving most teachers without the required training or content knowledge to carry out their tasks [2]. Relatedly, the 2012 Government Kindergarten Situational Report concludes that the 2004 curriculum established is sound, but that teacher behavior has not yet adapted to reflect the new pedagogy. Rectifying this is a top priority for the Ghana Education Service (GES). In order to improve instructional quality, New York University and the University of Pennsylvania in partnership with Innovations for Poverty Action (IPA), the Ghana Education Service, and the National Nursery Teacher Training Center (NNTTC), developed and implemented the QP4G program. The teacher in-service training component of the QP4G program (the component under examination in this report) was an experimental eight-day in-service kindergarten teacher training program with monitoring and coaching visits. The QP4G program began with a five-day course at the start of the first term, followed by a two-day refresher training four months later at the start of the second term, and a one-day refresher four months after that at the start of the third term. The QP4G program was implemented from September 2015 to July 2016.

[7] The 2015 Education Management Information System (EMIS) data showed that in the 2014/2015 school year, for instance, 38% and 95% of teachers in public and private kindergartens, respectively, did not have formal training directed towards teaching kindergarten.

The QP4G program focused on helping teachers provide age-appropriate and play-based instructional techniques and an encouraging, positive classroom environment according to the national curriculum. The National Nursery Teacher Training Center (NNTTC) in Accra implemented the QP4G program within the existing governmental structure of the GES in six municipal education office jurisdictions in the Greater Accra Region. The municipals were La Nkwantanang-Madina; Ga Central; Ledzorkuku-Krowor; Adentan; Ga East and Ga South. The District Support Focal Persons (DSFPs) conducted the monitoring and coaching visits after the training sessions. The DSFPs are staff in the district education offices who were mandated to conduct monitoring and coaching visits to the participating schools. They occupied such positions as Deputy District Directors of Education in charge of supervision, Basic Schools Coordinators, Early Childhood Education Coordinators, and Circuit Supervisors. Each participating school received at least two visits per term, lasting one hour each. Using monitoring checklists, the DSFPs provided feedback, a copy of which was left with the school, municipal GES, and IPA. A total of 161 schools, with a student population of 8425 children, participated in the training and coaching activities of the QP4G program. A total number of 337 KG teachers participated in the teacher training program. The QP4G program served an average kindergarten class size of 25 children, with a range from four to 70 children, aged from three to eight years.

A two-year impact evaluation of the program across 240 public and private schools in the six municipalities was conducted to assess its impact on the quality of teachers' instructional practice and children's primary school readiness. The results show that the program (a) increased the number of developmentally appropriate, kindergarten-specific pedagogical activities used while teaching and improved the quality of teacher-child interactions; and (b) led to a small positive improvement in children's primary school readiness [3, 4]. Given the potential to scale-up this program, one important question is whether the QP4G program is cost-effective. Thus, we investigate how much resources need to be expended to change instructional practice within the kindergarten classroom to improve classroom and child outcomes in a sustained and cost-effective manner. The cost and cost-effectiveness analysis of the QP4G program seeks to estimate the cost of replicating on a larger scale the implementation of the QP4G program to achieve results similar to those observed in the impact evaluation.

1.3 Cost Components of the QP4G Program

The cost analysis of the QP4G program was done from both a societal and a budgetary perspective. Data on the program costs were obtained from (a) IPA's record of transaction

details on the direct [monetary] expenses incurred in implementing the QP4G program[8] and (b) a [telephone] survey on participants' time spent and salary. Data on teachers' salaries were obtained from direct surveys of teachers. The program costs were determined based on the details and the resources used in implementing the QP4G program. This ensures that the opportunity costs of all the resources expended in implementing the QP4G program are accounted for. The total cost of implementing the QP4G program consists of the training costs, coaching costs, and administrative costs (Appendix Table 1).

a. ***Training costs*:** Training costs included the trainer costs, trainee costs, cost of supplies and materials as well as the time cost of the trainees. Trainer costs[9] were the costs of the honorarium for the NNTTC trainers, transportation for the trainers, venue rental for the training program, teaching and learning materials, as well as meals for the participants. Trainee costs were the transportation costs (directly reimbursed to trainees) and time costs of the participants who attended the QP4G program. This time is a real resource cost because trainees were diverted from other activities for eight days, even if there were no impact on supplier expenditures. Costs of supplies and materials included expenses associated with stationaries and training materials. Stationaries used for the training program included pens/pencils, sticky notes, exercise books, flip charts, and markers. Each training participant received a participants' reference guide, training manual, visual aids, posters, and a certificate of participation.

b. ***Coaching costs*:** Coaching costs were costs incurred by the coaches (who were district ECE coordinators) who participated in the monitoring and coaching activities. Coaching costs were comprised of the transportation (directly reimbursed) and time costs of coaches as well as the cost of the supplies and materials used for the monitoring and coaching activities. The supplies and materials included stationaries and monitoring forms.

c. ***Administrative costs*:** Administrative costs were costs incurred by the program managers or administrators in implementing the QP4G program. The administrative costs were facility rental for training coaches, travels, supplies and materials as well as the salaries of the IPA staff who managed the implementation of the QP4G program.

[8] This included the partners' costs at GES and NNTTC (other than staff time), given that IPA paid for them.

[9] The trainer costs also included the cost of developing the curriculum (i.e., training manual) for the QP4G program. The cost of developing the curriculum constituted about 8% of the QP4G program cost. This cost would not be incurred on a regular basis in an ongoing program and was therefore not included in this cost analysis.

1.3.1 Time costs

The time costs were the dollar value or cost of the time spent by the program participants. The participants contributed their time to the QP4G program that would have been engaged in productive work in their respective roles had they not participated in the program. The inclusion of these economic costs accounts for the opportunity cost of their time. The realized income method was used to value the time spent by the program participants – only their monthly "take-home" pay or salaries were used in computing the hourly wage. To compute the cost of the participants' time, we first measured the time spent on the program. The program participants spent time commuting to and from the training venue as well as participating in the training program - the training time. The training time per participant was the sum of the time spent traveling to/from the training program and during the training, multiplied by the number of training days. The coaching time was the time spent by the coaches traveling to/from schools and conducting monitoring and coaching activities in the participating schools. The coaching time per participant was the sum of the time spent traveling to and from schools for monitoring and coaching, multiplied by the average number of monitoring visits conducted within the academic year. The total time spent by the coaches was the sum of the training and coaching time. On average, the teachers and coaches spent nearly 77 hours and 85 hours, respectively, in participating in the QP4G program (Table 6~~Table 1~~). Note that these estimates are conservative, as we assume that all teachers and coaches participated in the program activities to the fullest extent.

Table 6~~1~~. Time spent by the program participants

	Trainer	Teacher	Coach
Teacher training			
Average time spent traveling to and from training (min)	129	184	166
Time spent during training (min)	390	390	390
Number of days of training	8	8	3
Total training time spent (min)	4152	4592	1668
Monitoring and coaching			
Average time spent traveling to and from school (min)	-	-	158
Average time spent in school (min)	-	-	86
Average number of monitoring sessions per term	-	-	14
Total coaching time spent (min)	-	-	3416
Total time spent (minutes)	4152	4592	5084

	Trainer	Teacher	Coach
Total time spent (hours)	69.2	76.5	84.7

Data on the teachers' monthly salaries were obtained via administered surveys with teachers. For the trainers and the coaches, salary data were obtained through a short telephone survey after program implementation. Participants' salaries were measured as a categorical variable, due to the reluctance of some participants to provide actual salary figures. We converted the categorical responses of the monthly salary of the participants to Ghana cedis amounts by assigning the median of each category and then converting the salary into US dollars as with the other costs. In estimating the hourly wage of program participants, we used their net salary per month. The hourly wage per participant was the ratio of the monthly salary to the product of the hours worked a day and the number of days worked a month. This worked out to be $0.96, $1.91, and $1.88 per hour for the teachers, trainers, and coaches, respectively. Using the average time spent on the QP4G program, including time in training and in transit, (Table 2), the costs of each teacher and trainer's time in participating in the QP4G program were $73.57 and $132.40, respectively.

Table 7~~2~~. Monthly Salary and Cost of Participants' Time on the QP4G Program

Item	Trainer	Teacher	Coach
Average monthly salary	$336.74	$169.18	$330.05
Hourly wage	$1.91	$0.96	$1.88
Time cost per-participant	$132.40	$73.57	$158.90
Number of participants	9	337	47
Total cost of participants' time	$1,191.60	$24,792.35	$7,468.27

1.4 Resource Cost Analysis of the QP4G Program

Table 3 shows the resource costs and costs per participant[10] of implementing the QP4G program. The resource costs occurred within nine months of one academic year and largely at the start of the school year; as such, no adjustments were made to account for inflation or discounting within the school year. The total costs of the QP4G program were divided by the number of schools, children, and teachers to obtain the costs per participant. The QP4G program costs totaled $135,506.91 per academic year for an estimated 8425 kindergarten

[10] The average class size served was 25 students (range = 4 – 70). Thus, we calculate costs per student as 1/25 the cost per teacher. We also assume 2 teachers / classrooms per school in calculating school-level costs.

children across 161 schools. Of this, approximately 67% was due to the training program. The trainee cost constituted the largest share of the training cost with a value of $43,336.63. A substantial part was due to transportation costs for the teachers who participated in the QP4G program. The transportation cost totaled $18,544.28 with an average of $55.03 per teacher. Additionally, each teacher allocated $73.57 of his/her time to the training. Trainer costs amounted to $38,893.38 and comprised of NNTTC's professional fees of honorarium, transportation, and meals for the training participants.

Nearly 15% of the program costs went to the monitoring and coaching visits. A noticeably large portion of the coaching cost of $9,789.05 was used to cover the transportation costs of the coaches. These costs represent a significant influx of resources into the municipal education offices in boosting the DSFPs' core monitoring and supervisory mandates. Notably, while these visits are part of the work of district ECE coordinators, lack of resources often prohibits the number of visits conducted. The impact estimates showed that teachers who received the training program implemented an average of 4.6 developmentally appropriate practices, compared to the average of 3.1 activities implemented by the control group classrooms [3]. Notably, the program increased the number of activities teachers used in the classroom ($p < .001$, $d_{wt} = 1.68$) [3], indicating that teachers apparently demonstrated greater commitment to transferring the knowledge, skills, and attitudes acquired to the classrooms. Part of the cost of the monitoring and coaching visits was $2,796.93 for supplies and materials such as stationaries and monitoring forms. Each municipal directorate was given a proportionate amount of administrative fees based on the number of participating schools in the municipality for supplies and materials.

Across all activities, administrative costs constituted 19% of the total cost of implementing the QP4G program. The salaries of the full-term IPA administrative staff who managed the implementation of the QP4G program drove the largest portion of the administrative costs (i.e., $20,758.34). In addition, administrative travels cost $3,541.41 for the administrative staff who worked on the QP4G program. Altogether, salaries and travel constituted most of the administrative cost of implementing the QP4G program.

In addition to the program costs, as noted above, there may be distortion costs due to distortions created to raise the funds needed for the program. We do not have a measure of distortion costs, but they often are assumed to be on the order of magnitude of 25% of the public revenues that need to be arranged to support a program [1]. Though we do not incorporate such distortion costs into our estimates, readers should be aware that they might

Table 8~~3~~. Resource Costs and Cost Distribution for the QP4G Program

Ingredient	Component	Resource Costs[11]	% of Total	Overall %	*School*	*Teacher*	*Child*
Trainer fees		$38,893.38	27.02%		$241.57 (10.23)	$115.41 (4.89)	$4.62 (0.20)
Transport costs for trainees	Training	$18,544.28	13.69%	66.55%	$115.18 (9.98)	$55.03 (4.77)	$2.20 (0.19)
Training supplies and materials		$7,954.95	5.53%		$49.41 (10.36)	$23.61 (4.95)	$0.94 (0.20)
Trainees' time		$24,792.35	17.23%		$153.99 (0.03)	$73.57 (0.01)	$2.94 (0.00)
Transport costs for coaches		$9,789.05	6.80%		$60.80 (10.12)	$29.05 (4.83)	$1.16 (0.19)
Coaching supplies and materials	Coaching	$2,796.93	1.94%	14.80%	$17.37 (1.80)	$8.30 (0.86)	$0.33 (0.03)
Coaches' time		$7,468.27	5.19%		$46.39 (0.32)	$22.16 (0.15)	$0.89 (0.01)
Facilities - venue for training		$324.81	0.23%		$2.02 (2.78)	$0.96 (1.33)	$0.04 (0.05)
Administrative travels	Administration	$3,541.41	2.46%	18.65%	$22.00 (10.38)	$10.51 (4.96)	$0.42 (0.20)
Adm. supplies and materials		$643.14	0.45%		$3.99 (2.30)	$1.91 (1.10)	$0.08 (0.04)
Adm. Personnel		$20,758.34	14.42%		$128.93 (10.02)	$61.60 (4.79)	$2.46 (0.19)
Total Resource Cost		**$135,506.91**	**100.00%**	**100.00%**	**$841.66 (68.30)**	**$402.10 (32.63)**	**$16.08 (1.31)**
Total Resource Cost with 25% increase in transportation costs		$143,475.59			$891.15 (75.92)	$425.74 (36.27)	$17.03 (1.46)
Total Resource Cost with 50% increase in transportation costs		$151,444.27			$940.65 (83.54)	$449.39 (39.91)	$17.98 (1.61)

Notes: Figures in parentheses are standard deviations.

[11] Based on Dhaliwal et al.'s (2012) recommendations, the program costs for the base year 2015 were adjusted to the year of the analysis (i.e., 2017) using a 10% discount rate.

increase the real resource costs by this order of magnitude. Also, as we note above, there may be future costs after the QP4G program year if, for example, the program results in more schooling for the students exposed, which means more teachers and other resources to support their additional schooling due to the program will be required. Again, we do not incorporate such costs into our estimates, but readers should be aware that our estimates may be downward biased because we do not include these costs.

1.5 Cost-Effectiveness Ratios (CER) of the QP4G Program

How much does it cost per unit gain to improve the implementation of developmentally appropriate pedagogy and children's school readiness outcomes in kindergarten classrooms in Ghana through QP4G? As seen in Table 3, the total costs of implementing the QP4G program per school, teacher, and child were $841.66, $402.10, and $16.08, respectively. Comparison of our resource cost estimates with particular benefits[12] gives measures of the cost-effectiveness of the QP4G program to compare with alternative means of obtaining the same benefits. For this reason, the CER estimates are important. Based on our estimates of the QP4G program (Table 4), the real annual resource costs of increasing school readiness for a child by one standard deviation was $107.23 and of increasing the implementation of developmentally appropriate pedagogy per a teacher by one standard deviation was $744.63 . The cost-effectiveness ratios for improving the implementation of developmentally appropriate pedagogy were more favorable, compared to improving child school readiness, because of the observed substantially higher program impact. Therefore, QP4G would be more cost-effective in increasing school readiness than any other programs with CERs more than $107.23 per child per year and more effective in increasing the implementation of developmentally appropriate pedagogy than any other programs with CERs more than $744.63 per teacher per year. As this represents the first ever cost-effectiveness analysis of an in-service kindergarten teacher training program in Ghana, we are unable to determine whether these CERs are high or low unless more programs are evaluated. Notwithstanding, our estimates provide benchmark estimates about the cost structure of a cost-effective in-service teacher training program with monitoring and coaching visits.

[12] We used the effect sizes of 0.54 and 0.15 standard deviation changes in the implementation of developmentally appropriate pedagogy and children's school readiness, respectively, as the base for the CERs. These effect sizes were observed at the end of the school year (i.e., 2016). A comparison also was made using impact results observed a year later (i.e., 2017).

Table 4. Cost-Effectiveness Ratios of the QP4G Program

Outcomes	Effect Size[a]	Total Cost	Cost Distributions Per			Cost-Effectiveness Ratios	
			School	*Teacher*	*Child*	*School*	*Teacher*
Using Impact Results at the End of the School Year (i.e., 2016)							
Actual Resource Costs							
Implementation of developmentally appropriate pedagogy	0.54	$135,506.91	$841.66	$402.10	$16.08	$1,558.63	$744.63
Child school readiness	0.15	$135,506.91	$841.66	$402.10	$16.08	$5,611.05	$2,680.65
Resource costs with 25% increase in transportation costs							
Implementation of developmentally appropriate pedagogy	0.54	$143,475.59	$891.15	$425.74	$17.03	$1,650.28	$788.41
Child school readiness	0.15	$143,475.59	$891.15	$425.74	$17.03	$5,941.02	$2,838.29
Resource costs with 50% increase in transportation costs							
Implementation of developmentally appropriate pedagogy	0.54	$151,444.27	$940.65	$449.39	$17.98	$1,741.94	$832.20
Child school readiness	0.15	$151,444.27	$940.65	$449.39	$17.98	$6,270.98	$2,995.93
Using Impact Results a Year Later (i.e., 2017)[b]							
Actual Resource Costs							
Implementation of developmentally appropriate pedagogy	0.21	$108,580.97	$674.42	$322.20	$12.89	$3,211.50	$1,534.28
Child school readiness	0.11	$108,580.97	$674.42	$322.20	$12.89	$6,131.05	$2,929.08
Resource costs with 25% increase in transportation costs							
Implementation of developmentally appropriate pedagogy	0.21	$116,027.36	$720.67	$344.29	$13.77	$3,431.75	$1,639.50
Child school readiness	0.11	$116,027.36	$720.67	$344.29	$13.77	$6,551.52	$3,129.95
Resource costs with 50% increase in transportation costs							
Implementation of developmentally appropriate pedagogy	0.21	$123,473.74	$766.92	$366.39	$14.66	$3,651.99	$1,744.72
Child school readiness	0.11	$123,473.74	$766.92	$366.39	$14.66	$6,971.98	$3,330.83

[a] The QP4G program impacts are given in terms of standard deviations. The number of SDs is a measure of how much the quality of teacher-child interaction or school readiness changed in the treatment group relative to the control group. All effect sizes are statistically significant.

[b] Following the recommendation of Dhaliwal et al. (2012), we inflate the present value of the resource costs in the base year to the year of analysis.

1.5.1 Sensitivity of the CERs to increase in transportation costs

An important concern for policy makers and donors is the cost of scaling-up the QP4G program to other regions of the country. The QP4G program was implemented within the Greater Accra Region with the inclusion of poor municipalities with close proximity to the training center – the NNTTC. Scaling-up the program in remote areas within or outside the region would mean that the costs of providing the program to such areas would increase. We, therefore, conducted sensitivity analyses to assess the robustness of the CERs using a range of plausible assumptions about the program costs while keeping constant the number of training days (and the costs for those) and participants. We investigated the sensitivity of the CERs to 25% and 50% increases in transportation costs. Our estimates from Table 4 demonstrated that a 25% and 50% increase in the transportation cost would increase the CER from $744.63 to $788.41 and $832.20 , respectively, per standard deviation improvement in the implementation of developmentally appropriate pedagogy. In addition, the cost per child for a standard deviation improvement in school readiness would increase from $107.23 to $113.53 and $119.84 for a respective 25% and 50% increase in transportation costs of the QP4G program. Using the impact results a year after (i.e., 2017) yields higher CERs because of the smaller estimated effect sizes of 0.21 and 0.11 standard deviation improvements in the implementation of developmentally appropriate pedagogy and child primary school readiness, respectively.

1.6 Government's Budgetary Expenditures (BE) and BE Ratios

Table 5 indicates the budgetary expenditures under two assumptions: (a) the time of trainees is not included because they would have been paid in any case and (b) the time of trainees, coaches, and administrative personnel is not included because they would have been paid in any case. With direct expenditures excluding time cost for teachers, the QP4G program would imply governmental expenditures of $110,714.55. Furthermore, the government would expend $82,487.95 on the QP4G program with direct expenditures excluding time costs for teachers, coaches, and administrative personnel. We note that the budgetary expenditures in the two scenarios about impacts of the program on payments to personnel are substantially lower than the resource cost estimates. However, we emphasize again that these do NOT represent resource costs if there are any positive opportunity costs to the time of personnel. With direct expenditures excluding time costs for teachers, coaches, and administrative personnel and using the impact results at the end of the school year (i.e., 2016) [Appendix Table 2], the government would spend $453.28 per a teacher for a one standard deviation change in increasing the implementation of developmentally appropriate pedagogy on the QP4G program.

The budgetary expenditure ratio would be $65.27 per child for a one standard deviation improvement in child school readiness.

However, with one-year later impact results (i.e., 2017), the budgetary expenditures per school, child, and teacher as well as the respective budgetary expenditure ratios would increase considerably for improved implementation of the developmentally appropriate pedagogy and child school readiness. For instance, the budgetary expenditure ratios for direct expenditures excluding time costs for teachers, coaches and administrative personnel would increase from $453.28 to $861.60 per a teacher for improved implementation of the developmentally appropriate pedagogy. Also, the budgetary expenditure ratios would rise slightly from $65.27 to $65.80 per child for a one standard deviation improvement in school readiness. We acknowledge that it is possible that these scenarios are of possible interest, but it is important to realize that to interpret BERs as CERs requires the assumption that personnel would not have done anything of value for the time they devoted to the QP4G program had there not been such a program.

Table 5. Budgetary Expenditure for the QP4G Program

	Costs[a]	***School***	***Teacher***	***Child***
Direct expenditures excluding time cost for teachers	$110,714.55	*$687.67*	*$328.53*	*$13.14*
		(68.28)	*(32.62)*	*(1.31)*
Direct expenditures excluding time costs for teachers, coaches and administrative personnel	$82,487.95	$512.35	$244.77	$9.79
		(47.71)	(31.44)	(1.27)
Total Resource Cost[b]	$135,506.91	$841.66	$402.10	$16.08
		(68.30)	(32.63)	(1.31)

[a] Based on Dhaliwal et al.'s (2012) recommendations, the program costs for the base year 2015 were adjusted to the year of the analysis (i.e., 2017) using a 10% discount rate.

[b] Total resource costs are the same estimates presented in Table 3. They are included here as a basis of comparison.

1.7 Conclusions

The costs and cost-effectiveness analysis presented in this report reveals critical information needed to consider scaling-up the QP4G program. The program cost about $135,506.91 to implement the eight-day in-service teacher training program with monitoring and coaching visits in the 161 schools across six municipalities in the Greater Accra Region. About 67% of these costs were attributed to the training itself. The average total costs per school, teacher, and child

were approximately $842, $402, and $16. When considering budgetary expenditures only (excluding resource and time costs of participants), expenditures per school, teacher, and child were $512.35, $244,77, and $9.79, respectively.

We then consider the cost-effectiveness ratios (CERs), with estimated effects sizes of 0.54 and 0.15 by the end of the school year in 2016 and 0.21 and 0.11 a year later for improved implementation of developmentally appropriate pedagogy and school readiness among children, respectively. To compare the QP4G program in terms of cost-effectiveness with other programs, the CERs in Table 4 need to be compared with the CERs for the other programs. The QP4G program is more (less) cost-effective than other programs that have CERs greater (less) than those in these tables. Overall, the program would be more cost-effective in increasing school readiness and the implementation of KG-specific pedagogy than any other program with a cost-effectiveness ratio (CER) of a one standard-deviation improvement of more than $107.23 per child and $744.63 per teacher, respectively, per year. The CERs for the QP4G program at first glance may appear more favorable to the implementation of developmentally appropriate pedagogy in schools than the improvement of school readiness because of the higher effect sizes reported for the former. But that inference needs to be heavily qualified because there were not two independent programs for developmentally appropriate pedagogy in schools and for improvement of child primary school readiness. Instead there was a package that affected both of these outcomes. The cost and CER estimates for the QP4G program are sensitive to the underlying assumptions and would increase as reflected in Table 3 and Table 4 were there to be increases in travel and transportation costs in any scale-up process. Finally, comparisons between the CER (Table 4) and BER (Appendix Table 2) estimates illustrate how under the assumptions of direct expenditures excluding time costs for teachers or broader direct expenditures with no teacher, coaches, and administrative personnel payments, governmental expenditures would be less than resource costs, and by how much they therefore would be downward biased estimates of resource costs.

References

1. Harberger AC. New frontiers in project evaluation? A comment on devarajan, squire, and suthiwart-narueput. The World Bank Research Observer. 1997; 12(1):73-9.
2. UBS Optimus Foundation. (2014). Exploring early education programs in peri-urban settings in Africa: Summary findings from Accra, Ghana. Optimus Impact: Research Paper UBS Optimus Foundation. January 2014, Issue 3.
3. Wolf, S., Tsinigo, E., Behrman, J., and Aber, JL. (2017). Testing and scaling-up supply- and demand-side interventions to improve kindergarten educational quality in Ghana. SIEF Midline Report, 2017. Available at http://www.worldbank.org/en/programs/sief-trust-fund/brief/testing-and-scaling-up-supply-and-demand-side-interventions-to-improve-kindergarten-educational-quality-in-ghana
4. Wolf, S., Tsinigo, E., Behrman, J., Aber, JL. and Bonarget, A. (2017). Testing and scaling-up supply- and demand-side interventions to improve kindergarten educational quality in Ghana. SIEF Final Report, 2017.
5. Dhaliwal, I., Duflo, E., Glennerster, R., and Tulloch, C. (2012). Comparative cost-effectiveness analysis to inform policy in developing countries: A general framework with applications for education. J-PAL working paper. Retrieved September 28, 2017. Available at https://www.povertyactionlab.org/sites/default/files/publications/CEA%20in%20Education%202013.01.29_0.pdf

Appendices

Appendix Table 1. Resource Costs Ingredients

Resource Components	Ingredients
Training cost - Trainer costs	a. Honorarium for the NNTTC trainers b. Transportation for the trainers c. Venue rental for the training program d. Printing of teaching and learning materials e. Meals for the participants.
Training cost - Trainee cost	a. Transportation costs b. Cost of time for travelling and participating in the training program c. Time costs of the participants who attended the QP4G program.
Training cost - Costs of supplies and materials	a. Stationaries: Pens/pencils, sticky notes, exercise books, flip charts, and markers. b. Materials: Printing participants' reference guide, training manual, visual aids, posters, and a certificate of participation
Coaching costs	a. Transportation costs during training and monitoring visits b. Cost of time in participating in training and monitoring activities c. The supplies and materials included stationaries and monitoring forms.
Administrative costs	a. Facility rental for training coaches b. Transportation costs c. Supplies and materials as well as the salaries of the QP4G program managers.

Appendix Table 2. Budgetary Expenditure Ratios of the QP4G Program

Outcomes	*Effect Size*[a]	*Total Cost*	***Cost Distributions Per***			***Budgetary Expenditure Ratios Per***	
			School	*Teacher*	*Child*	*School*	*Teacher*
Using Impact Results at the End of the School Year (i.e., 2016)							
No additional budgetary expenditures on trainees (teachers)							
Implementation of developmentally appropriate pedagogy	0.54	$110,714.55	$687.67	$337.00	$13.14	$1,273.46	$624.07
Child school readiness	0.15	$110,714.55	$687.67	$337.00	$13.14	$4,584.45	$2,246.67
No additional budgetary expenditures on teachers, coaches, and administrative personnel							
Implementation of developmentally appropriate pedagogy	0.54	$82,487.95	$512.35	$244.77	$9.79	$948.79	$453.28
Child school readiness	0.15	$82,487.95	$512.35	$244.77	$9.79	$3,415.65	$1,631.81
Using Impact Results a Year Later (i.e., 2017)[b]							
No additional budgetary expenditures on trainees (teachers)							
Implementation of developmentally appropriate pedagogy	0.21	$83,788.61	$520.43	$248.63	$9.95	$2,478.22	$1,183.96
Child school readiness	0.11	$83,788.61	$520.43	$248.63	$9.95	$4,731.15	$2,260.28
No additional budgetary expenditures on teachers, coaches, and administrative personnel							
Implementation of developmentally appropriate pedagogy	0.21	$60,975.64	$378.73	$180.94	$7.24	$1,803.48	$861.60
Child school readiness	0.11	$60,975.64	$378.73	$180.94	$7.24	$3,443.01	$1,644.88

[a] The QP4G program impacts are given in terms of standard deviations. The number of SDs is a measure of how much the quality of teacher-child interaction or school readiness changed in the treatment group relative to the control group. All effect sizes are statistically significant.

[b] Following the recommendation of Dhaliwal et al. (2012), we inflate the present value of the budgetary expenditures in the base year to the year of analysis.