Routledge
Taylor & Francis Group

Journal of Research on Educational Effectiveness

ISSN: 1934-5747 (Print) 1934-5739 (Online) Journal homepage: https://www.tandfonline.com/loi/uree20

Experimental Impacts of the "Quality Preschool for Ghana" Interventions on Teacher Professional Well-being, Classroom Quality, and Children's School Readiness

Sharon Wolf, J. Lawrence Aber, Jere R. Behrman & Edward Tsinigo

To cite this article: Sharon Wolf, J. Lawrence Aber, Jere R. Behrman & Edward Tsinigo (2019) Experimental Impacts of the "Quality Preschool for Ghana" Interventions on Teacher Professional Well-being, Classroom Quality, and Children's School Readiness, Journal of Research on Educational Effectiveness, 12:1, 10-37, DOI: 10.1080/19345747.2018.1517199

To link to this article: https://doi.org/10.1080/19345747.2018.1517199

JOURNAL OF RESEARCH ON EDUCATIONAL EFFECTIVENESS
2019, VOL. 12, NO. 1, 10–37
https://doi.org/10.1080/19345747.2018.1517199

INTERVENTION, EVALUATION, AND POLICY STUDIES OPEN ACCESS Check for updates

Experimental Impacts of the "Quality Preschool for Ghana" Interventions on Teacher Professional Well-being, Classroom Quality, and Children's School Readiness

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

ABSTRACT

We assessed the impacts of a teacher professional development program for public and private kindergartens in the Greater Accra Region of Ghana. We examined impacts on teacher professional well-being, classroom quality, and children's readiness during one school year. This cluster-randomized trial included 240 schools (teachers $N = 444$; children $N = 3{,}345$, $M_{age} = 5.2$) randomly assigned to one of three conditions: teacher training (TT), teacher training plus parental-awareness meetings (TTPA), and controls. The programs incorporated workshops and in-classroom coaching for teachers and video-based discussion groups for parents. Moderate impacts were found on some dimensions of professional well-being (reduced burnout in the TT and TTPA conditions, reduced turnover in the TT condition), classroom quality (increased emotional support/behavior management in the TT and TTPA conditions, support for student expression in the TT condition), and small impacts on multiple domains of children's school readiness (in the TT condition). The parental-awareness meetings had counteracting effects on child school readiness outcomes. Implications for policy and practice are discussed for Ghana and for early childhood education in low- and middle-income countries.

KEYWORDS

early childhood education
kindergarten
classroom quality
school readiness
Ghana
teacher training
and coaching

Introduction

Since the start of the 21st century, there has been a dramatic increase globally in emphasizing the period of early childhood development (ECD). Estimates suggest that 250 million children younger than 5 years of age in developing countries are at risk of not reaching their developmental potential, as indicated by stunting and living in

CONTACT Sharon Wolf wolfs@upenn.edu Graduate School of Education, University of Pennsylvania, 3700 Walnut Street, Philadelphia, PA 19104, USA.

[a]Graduate School of Education, University of Pennsylvania, Philadelphia, Pennsylvania, USA
[b]Global TIES for Children, Institute for Human Development and Social Change, New York University, New York, New York, USA
[c]Departments of Economics and Sociology, School of Arts and Sciences, University of Pennsylvania, Philadelphia, Pennsylvania, USA
[d]Innovations for Poverty Action, Accra, Ghana

Supplemental data for this article can be accessed on the publisher's website.

extreme poverty (Lu, Black, & Richter, 2016). Compared to other regions, sub-Saharan Africa has the highest prevalence of children at risk, with 38% of the 143 million children younger than 5 stunted and 50% living in extreme poverty (Lu et al., 2016). Pre-primary education, or early childhood education (ECE), is one approach to support ECD, and estimates suggest that the benefit–cost ratio of increasing ECE enrollments in developing countries are considerable (Engle et al., 2011).

High-quality ECE can improve children's early learning skills, which are crucial for children's transition and adaptation to school (e.g., Blair, 2002; Morris et al., 2014; Yoshikawa et al., 2013). These "school readiness skills" include both early academic skills and behavioral and social skills. As part of a global initiative to improve children's learning, Target 4.2 of the United Nation's Sustainable Development Goal 4 on education aims to expand access to high-quality ECE and improve school readiness outcomes for all children (United Nations, 2015). Yet as ECE enrollments have expanded across low- and middle-income countries (LMICs), concerns about the quality of services have increased (Britto, Yoshikawa, & Boller, 2011; Engle et al., 2011).

ECE in Ghana

Ghana is a lower-middle-income country in West Africa with a gross domestic product per capita of $4,300 and a population of 26.9 million people (Central Intelligence Agency, 2016). It is estimated that one-third (32.6%) of Ghanaian 3- and 4-year-olds do not meet school readiness indicator thresholds, including following directions, working independently, avoiding distraction, getting along with others, and avoiding aggression (McCoy et al., 2016). In 2004, the government adopted the National Early Childhood Care and Development Policy, which highlighted access to quality early education as central to improving ECD and learning as well as to reducing inequalities in learning outcomes. In 2007, 2 years of pre-primary education—called *kindergarten 1* (KG1; the equivalent to pre-K in the United States) and *kindergarten 2* (KG2; the equivalent to kindergarten in the United States), respectively—were added to the universal basic education system that had previously begun in the first grade of primary school.

As a result, Ghana has one of the highest pre-primary enrollment rates on the continent at 75% net enrollment in 2015–2016 (Ghana Ministry of Education, 2016). Yet several reports have concluded that educational quality and learning outcomes are low (e.g., Ghana Ministry of Education, 2014), including in the kindergarten (KG) sector (Ghana Education Service, 2012). The 2012 Government Kindergarten Situational Report concluded that the curriculum established in 2004 was sound but that teachers had not incorporated the new pedagogy into their practice. Thus, the Ghana Education Service (the implementing arm of the Ministry of Education) declared training KG teachers a top education policy priority. A secondary priority was to engage parents in their child's KG education at home and in school.

The private ECE sector in Ghana has grown in recent years. As the country's migration and urbanization continue to grow rapidly (CIA, 2016), the private education sector has been quick to respond to families' demands by providing educational services needed by the local community faster than the public sector has been able to. The 2014–2015 Education Management Information System data indicated that the number of public pre-

primary schools increased by 2.1%—from 13,861 in 2013–2014 to 14,147 by 2014–2015—while the number of private pre-primary schools increased at four times that rate—8.7%, or from 11,983 to 13,031—within the same period (Republic of Ghana, 2015). A large majority of these private schools charge very low fees and cater to low-income families (known as low-fee private schools), and many struggle to cover their costs (Abdul-Hamid, Baum, De Brular, Lusk-Stover, & Tettey, 2017). While public and private KG schools are required to follow the national curriculum, a major concern for policy makers remains educational quality and lack of standards enforcement in both sectors.

Improving KG Quality in Ghana

There have been relatively few systematic randomized studies of interventions to improve ECE quality in developing countries. The studies to date found that classroom quality can be improved by in-service teacher training interventions, but these improvements do not necessarily result in better outcomes for children. Two factors may limit the success of quality improvement efforts in Ghana specifically and low-income countries generally.

First, there is growing concern about a "motivation crisis" among teachers in LMICs (Moon, 2007) and in Ghana specifically (Bennell & Akyeampong, 2007). Teachers in LMICs face many challenges, including increasing workloads due to educational reform, low and unreliable teacher remuneration, lack of professional recognition, challenging working conditions (i.e., large class sizes), lack of accountability, minimal professional development opportunities, and lack of voice (Bennell & Akyeampong, 2007; Guajardo, 2011; Wolf, Aber, Torrente, McCoy, & Rasheed, 2015). Low teacher motivation and attendance, as well as high rates of turnover (Osei, 2006), are serious challenges to improving educational quality and child learning (Bennell & Akyeampong, 2007). Meta-analyses show that successful in-service teacher training, coupled with ongoing monitoring and coaching, is a proven way to improve teachers' instructional practice in both the United States (Kraft, Blazar, & Hogam, 2017) and in LMICs (Ganimian & Murnane, 2016) for primary school teachers. Such ongoing support may be necessary to combat this issue.

The second factor is the role of parents in children's learning. Parental involvement in school may be particularly beneficial for younger children, while lack of parental participation may limit the effectiveness of school-based programs (e.g., Connor, Son, Hindman, & Morrison, 2005). Regular communication between parents and schools allows them to work together toward children's learning and development and has been shown to improve longer-term academic outcomes for preschool and KG children in the United States (Miedel & Reynolds, 1999). While research on parent involvement in school in sub-Saharan Africa is much more limited, a study in South Africa found that many ECE teachers perceived parents to be uninterested or uninvolved in their children's education, and there was little reported communication between home and school (Bridgemohan, van Wyk, & van Staden, 2005).

In Ghana, levels of at-home cognitive stimulation are relatively low, with only 33.1% of children having been read to in the three days immediately prior to data collection (versus an average of 54.1% in all developing countries; McCoy et al., 2016). Ghanaian parents have been shown to value early education and demand academically focused, rigorous instruction from teachers (Bidwell & Watine, 2014; Kabay, Wolf, & Yoshikawa,

2017), and school involvement has been shown to partially mediate the positive associations between socioeconomic status and Ghanaian children's school readiness skills (Wolf & McCoy, 2017). In contexts such as Ghana, with low adult-literacy rates and educational levels, many parents' may not feel efficacious in promoting their child's learning. Aligning parents' and schools' expectations for ECE and supporting parents to engage in their child's education may be critical for sustainably changing teacher practice and children's development.

The Quality Preschool for Ghana Interventions and Theory of Change

The Quality Preschool for Ghana (QP4G) project aimed to build capacity and support for implementation of the 2004 KG curriculum and to enhance the quality of KG education in Ghana. The goal of QP4G was to develop and rigorously evaluate a scalable model of transformational teacher training to provide high-quality ECE services to children and to test the 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 Ghanaian children's school readiness skills. *Preschool* in this study refers to the two years of pre-primary education in Ghana called *kindergarten*.

The *teacher-training program* included training workshops (five days in September, two days in January, and one day in May) and in-classroom coaching (six visits over the course of the school year) administered by trained district government ECE coordinators. The training for teachers was led by professional teacher trainers at the National Nursery Teacher Training Center in Accra, a teacher-training facility affiliated with the Ministry of Education that provides ECE certification courses for teachers. The content focused on integrating play- and activity-based, child-centered teaching practices into teaching instructional content and covered five areas: (1) how children learn—developing a child-friendly environment, (2) classroom management, (3) incorporating child-centered and activity-based approaches to teaching language and literacy, (4) incorporating child-centered and activity-based approaches to teaching math, and (5) assessment and planning. The first half of each training day consisted of lectures and discussions, and the second half focused on practicing the techniques learned and creating teaching and learning materials to implement activities in the classroom.

District government ECE coordinators attended the teacher training as well as two days of training on their roles as monitors and coaches plus a one-day refresher training halfway through the school year. They were trained on the same content as the teachers as well as their role as coaches to guide teachers and answer their questions on the implementation of the curriculum. The coaching visits focused on practical ways that teachers could integrate the lessons from the training in their teaching, including positive classroom management (e.g., how to use consistent rewards and routines), assessment and planning, and integrating play-based activities into literacy and math lessons. At each visit, teachers were observed for a minimum of one hour, followed by debriefing sessions where teachers reflected on their practice and were provided with feedback on what they did well, as well as areas for improvement.

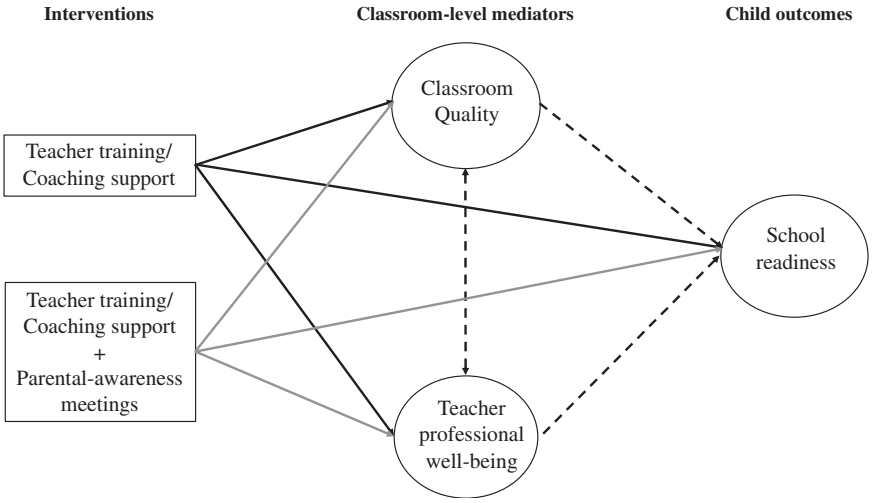

Figure 1. Quality preschool for Ghana theory of change. *Note.* Solid lines represent causal relations. Dashed lines represent noncausal relations. Only solid lines are examined in this study.

The *parental-awareness meetings* consisted of three meetings administered through school parent–teacher associations (PTAs) over the course of the school year. It was offered to all parents with KG children in the school and administered by the same trained district government ECE coordinators. Each meeting consisted of a video (content was developed for the intervention) followed by a discussion led by the district coordinator. The video themes were (1) the importance of play-based learning, (2) parents' role in child learning, and (3) encouraging parent–teacher and parent–school communication. The aim was to increase parental involvement with their child's education at home and in school and increase parent–teacher communication.

Finally, treatment schools were randomly assigned to receive reinforcement messages from the trainings (weekly text messages for teachers and picture-based paper flyers for parents; examples shown in Appendix B available in the online supplemental data) to test the impact of using affordable, scalable reinforcements.

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 received both the teacher training and parental-awareness meetings, it was anticipated that this combined package of programs would change parent–teacher communication and increase parental school involvement. This package of programs was considered a separate condition to teacher training only (see Figure 1).

Targeting teacher–child interactions. Teacher–child interactions have emerged as a key aspect of classroom quality in predicting children's outcomes (Phillips, Mekos, Scarr, McCartney, & Abbott–Shim, 2000; Pianta & Hamre, 2009) and were the primary target of the interventions. The "Teaching Through Interactions" framework focuses on the domains of classroom process quality that contribute to student learning and development, with processes related to instructional support, emotional support, and behavioral management as critical elements of quality (Pianta & Hamre, 2009). Similar

systematic research examining domains of process quality in preschool classrooms in sub-Saharan Africa does not yet exist.

Targeting teacher professional well-being. Aspects of teachers' professional well-being are related to process quality (La Paro et al., 2009; Pianta et al., 2005) and directly to student outcomes (Tsouloupas, Carson, Matthews, Grawitch, & Barber, 2010). Professional development interventions can improve teacher job satisfaction, self-efficacy, well-being, and performance (e.g., Brown, Jones, LaRusso, & Aber, 2010; Jennings, Frank, Snowberg, Coccia, & Greenberg, 2013; Piper & Zuilkowski, 2015). In Ghana, where teachers are more likely to be ill-equipped for the challenges of teaching, there have been concerns about a large-scale lack of motivation (Bennell & Akyeampong, 2007) and high rates of teacher turnover (Osei, 2006). Through providing teachers with support—in terms of increased behavioral management techniques and instructional skills as well as in-class coaching—the QP4G program aimed to improve teachers' professional well-being as a key part of improving children's learning experience.

Targeting school readiness skills. The development of nonacademic skills, such as social-emotional skills and executive functions, can facilitate children's engagement with the learning process and academic learning (Raver, 2002). When children can regulate their emotions and behaviors, show high levels of prosocial behavior, and sustain their attention on learning-related tasks, they benefit more from instruction in the classroom (Blair, 2002; Raver, 2002). Many teacher professional development programs in the United States have focused on the classroom emotional climate, teachers' instructional support, and behavior management techniques as pathways to improve children's nonacademic and academic skills development (e.g., Brown et al., 2010; Morris et al., 2014). Recent experimental research has shown that similar classroom processes facilitate the development of social-emotional skills in preschool children in Chile (Yoshikawa et al., 2015) and in KG children in Ecuador (Araujo, Carneiro, Cruz-Aguayo, & Schady, 2016). The QP4G study addressed similar processes in Ghanaian KG classrooms.

The Current Study

The current study evaluated the effectiveness of the QP4G teacher training and coaching program with KG classrooms in both private and public schools to improve (1) teacher well-being, (2) the quality of teacher practices and interactions with children, and (3) children's school readiness skills over one school year. Additional goals were to test the value of combining a scalable (low-cost) parental-awareness intervention with teacher in-service training and to examine several important sources of potential heterogeneity of impact, primarily impacts in public versus private schools. We examined two primary research questions (RQ1 and RQ2) and two secondary research questions (RQ3 and RQ4):

RQ 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?

RQ 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 (1) a control group and (2) the QP4G teacher training only?

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

RQ 4: Do impacts vary by key child characteristics (gender, grade level, baseline school-readiness skills) and school sector (public vs. private)?

Method

The implementation and first-year evaluation of the QP4G interventions occurred between September 2015 and June 2016. The research design was a cluster randomized trial, where schools were randomly assigned to one of three treatment arms: (1) teacher training (TT; 82 schools), (2) teacher training plus parental-awareness training (TTPA; 79 schools), and (3) control group (79 schools). The trial was registered in the American Economic Associations' registry for randomized controlled trials (RCT ID: AEARCTR-0000704). The school year in Ghana begins in September and ends in July. All data presented in this study were collected in September to October 2015 (baseline) and May to June 2016 (follow-up).

Randomization

First, randomization was conducted at the school level and stratified by district and public/private sector to one of the three treatment arms. Second, half of the TT treatment schools were randomly assigned to receive weekly text messages for teachers ($N = 40$ schools), and half of the TTPA treatment schools were randomly assigned to receive teacher text messages and picture-based paper flyers for parents ($N = 40$ schools). The research design is shown in Figure 2.

Sample Size Estimation

A sample size of 160 schools (for two-way comparisons) with seven children per class and two teachers/classrooms per school was assumed. With 80% power at the 5% significance level, and assuming an intraclass correlation (ICC) of 0.15 for child outcomes and 0.10 for teacher outcomes, this was sufficient to detect effect sizes of 0.17 standard deviations (minimum detectable effect size) for child outcomes and 0.33 for teacher outcomes.

Sampling Procedures

Six of the sixteen districts in the Greater Accra region were selected. These districts were rated as the most disadvantaged districts on the 2014 UNICEF District League Table (a social accountability index that ranks regions and districts based on development and delivery of key basic services, including education, health, sanitation, and governance; UNICEF, 2015) and were within a two-hour driving distance from Accra (for teachers to be able to attend the training, which was located at a training center in

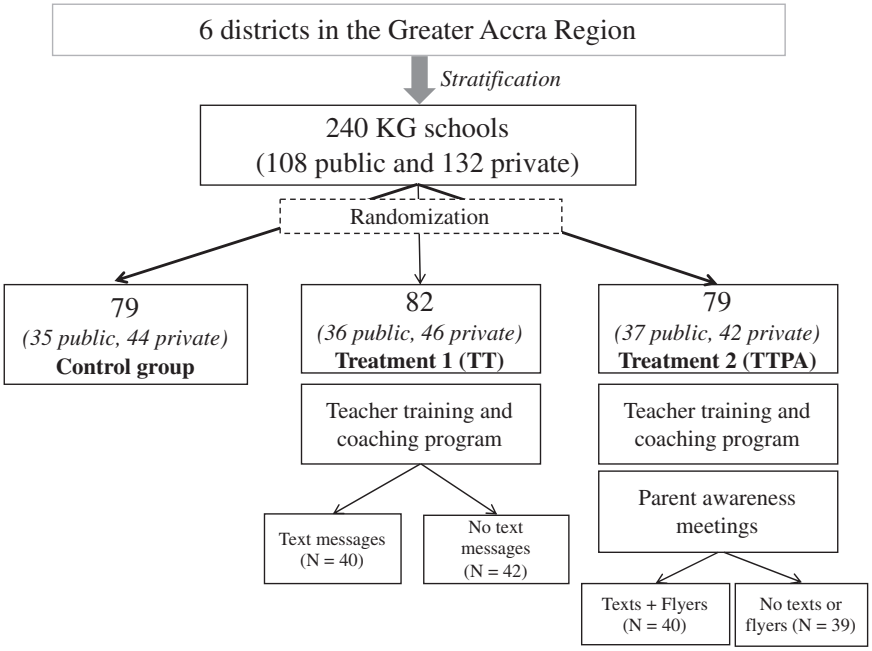

Figure 2. Research design. *Note.* The six districts in the Greater Accra Region include: Ga South, Adenta, Ledzokuku-Krowor, Ga Central, La Nkwantanang-Madina, and Ga West. School randomization is stratified by district and public- versus private-sector status.

Accra). On average, teachers traveled 1.5 hours each way for the training. The six districts were Ga South, Adenta, Ledzokuku-Krowor, Ga Central, La Nkwantanang-Madina, and Ga West.

School sample. All schools in the six districts were identified using the GES Educational Management Information System (EMIS) database, which lists all registered schools in the country. Schools were then randomly sampled stratified by district and within district by public and private schools. Eligible schools had to be registered with the government and have at least one KG class. A school listing was then conducted to confirm the presence of each school and to obtain information on each school's head teacher and proprietor. Because there were fewer than 120 public schools across the six districts, every public school was sampled. Private schools (490 total) were sampled within districts in proportion to the total number of private schools in each district relative to total for 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 sampled schools refused or was not eligible to participate in the study (because, for example, it does not have a KG program or no longer existed despite being listed in the EMIS data set). Then, 11 schools were replaced from the original 240 (6 had inaccurate location and contact information and could not be found, 3 refused to participate, and 2 did not have a KG program). The final baseline sample consisted of 240 schools, finalized prior to randomization. In the follow-up assessment, three schools dropped out of the study (two control schools, one TT school), and two (control) schools closed down.

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

Child sample. Class rosters for all KG classrooms were collected. An average of 15 children (8 from the KG1 teacher's classroom and 7 from the KG2 teacher's classroom) were randomly selected from each class roster to participate in direct assessments. If a school had fewer than 15 KG children enrolled across both classrooms, all children were selected. Assessors also randomly selected up to 10 additional children on the initial visit (a "reserve" list). If a selected child from the first 15 was not in school that day, assessors returned up to two times to assess the child. If the child was still not present on the third visit, a child from the reserve list was used to replace that child. At baseline, the total sample of children was 3,435 children, with an average of 14.3 children per school (range = 4–15). For schools with only one KG classroom, 15 children were randomly sampled from the classroom. The analysis does not include children who enrolled in school after baseline.

Measures

All measures were collected at baseline (beginning of the school year) and follow-up (end of the school year). All data within a school—teacher surveys, classroom observations, and child outcomes—were collected during the same visit at each round.

Teacher professional well-being. Teachers answered a survey in English (the language of Ghana's education system). 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 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.

Motivation. Teacher's motivation was measured using five items adapted from Bennell and Akyeampong (2007) as reported in Wolf, Aber et al. (2015). Items were answered on the following scale: 1 = *false*, 2 = *mostly false*, 3 = *sometimes*, 4 = *mostly true*, and 5 = *true*. 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" ($M = 4.6$, $SD = .59$, $\alpha = .77$).

Burnout. Teacher burnout was measured using 11 items from the Maslach Burnout Inventory (Maslach, Jackson, & Leiter, 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" ($M = 2.03$, $SD = .90$, $\alpha = .75$).

Job satisfaction. Teacher's job satisfaction was measured using six items adapted from Bennell and Akyeampong (2007) as reported in Wolf, Aber et al. (2015). Items were answered on the following scale: 1 = *true*, 2 = *somewhat true,* 3 = *somewhat false,* and 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 ($M = 3.09$, $SD = .69$, $\alpha = .73$).

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-fifth of teachers ($N = 97$) had left their position by follow-up.

Classroom outcomes. All teachers were videotaped teaching a lesson in their classrooms for 30 to 60 minutes at each wave. Videos were coded with two instruments: an implementation fidelity checklist and a tool to assess the quality of teacher–child interactions.

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" ($M = 3.51$, $SD = 2.22$).

Teacher–child interaction quality. All videos were coded using the Teacher Instructional Practices and Processes System (TIPPS; Seidman, Raza, & Kim, 2017; Seidman, Raza, Kim, & McCoy, 2013). 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 LMICs. 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 KG settings). More information about the assessment tool can be obtained by referring to Seidman and colleagues (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* (three items: connects lesson to teaching objectives; provides specific, high-quality feedback; and uses scaffolding; $\alpha = .42$), *emotional support and behavior management* (seven items: positive climate; negative climate (reverse-scored); sensitivity and responsiveness; tone of voice; positive behavior management; provides consistent routines; and student engagement in class activities; $\alpha = .83$), and *supporting student expression* (four items: considers student ideas and interests;

encourages students to reason and problem solve; connects lesson to students' daily lives; and models complex language; $\alpha = .63$). See Wolf, Raza and colleagues (2018) 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 prespecified 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 a set of TIPPS calibration criteria within three attempts to be certified as a TIPPS observer. The calibration criteria look at not only agreement but also the degree of deviation from master codes—both 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 cutoffs). Collectively, the criteria enhance the likelihood of achieving acceptable levels of interrater 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 interrater 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 variance in scores broke down into differences in individual raters and shared variance across raters. On average across items, 71.1% of the variance was shared across raters.

Child school readiness outcomes. School readiness was directly assessed in four domains: early literacy, early numeracy, social-emotional skills, and executive function. The instrument used was the International Development and Early Learning Assessment (IDELA), developed by Save the Children (Pisani, Borisova, & Dowd, 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 consisted of 38 items grouped into six subtasks, and it 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" ($\alpha = .74$).

Early numeracy. The domain of early numeracy consisted of 39 items grouped into eight subtasks and 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 for shape identification showed the child a picture with six shapes and asked the child to identify the circle ($\alpha = .72$).

Social-emotional skills. Social-emotional skills consisted of 14 items grouped into five subtasks covering 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 that 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 ($\alpha = .69$).

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 ($\alpha = .83$).

School readiness composite. For the primary impact analysis, scores on the four domains were combined with equal weight 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 .67 to .97, with an average kappa value of .86.

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, six district dummies, a dummy variable for whether the school was randomly assigned to receive teacher text messages, a dummy for whether the school was randomly assigned to receive text messages and 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.

Analytic Strategy

Baseline equivalency. We first conducted a baseline equivalency analysis to ensure that the randomization yielded treatment and control groups that were statistically equivalent. We tested whether the mean values for a set of school, teacher, and child characteristics differed by treatment group (see Table 1). Of note, there were some differences between child outcomes at baseline, with children in the TTPA condition performing slightly higher in early literacy and social-emotional skills than children in the TT and control conditions. Second, we conducted an omnibus F test with each set of characteristics in a multivariate analysis of variance (MANOVA) with treatment status as the grouping variable to assess whether overall the set of predictors was statistically differentiated across treatment groups. Overall, there were no meaningful differences across the three treatment arms for baseline school characteristics (omnibus $F\ (2) = 0.97$, $p = .520$), teacher characteristics (omnibus $F\ (2) = 1.06$, $p = .380$), or child characteristics (omnibus $F\ (2) = 0.99$, $p = .429$). Thus, we interpret the few differences between the intervention

Table 1. Means and mean differences in teacher and child characteristics at baseline, by treatment condition.

	Control	TT	TTPA	*F* statistic	*p* value
School characteristics	Mean or %				
Private school status	55.7%	56.1%	53.2%	0.08	.923
No. of years school has been established	23	23	19	0.95	.389
School has written rules/regulations for staff	38.5%	48.8%	35.9%	1.52	.222
Total number of KG children in school	54	63	60	0.64	.529
Total number of KG teachers on the payroll	2.0	2.3	2.2	0.98	.376
Main language of instruction in KG1					
English only	10.5%	13.5%	7.5%	0.68	.509
Mother tongue only	4.5%	1.4%	1.5%	0.90	.407
Mixture of English and mother tongue	85.1%	85.1%	91.0%	0.70	.496
Sample size (total =240)	79	82	79		
Teacher characteristics					
Female	97.9%	97.4%	97.3%	0.05	.953
Age	35.3	35.7	35.2	0.07	.933
Years as a teacher	6.55	6.16	6.64	0.22	.801
Years as a teacher in current school	3.37	3.47	3.21	0.17	.842
At least secondary high school (%)	97.1%	93.5%	91.3%	2.18	.114
Has any post-secondary training	60.0%	62.3%	58.7%	0.22	.804
Has training in ECD	65.7%	72.1%	64.0%	1.25	.288
Motivation	4.66	4.75	4.64	1.83	.162
Burnout	2.05	2.03	2.12	0.78	.459
Job satisfaction	3.05	3.12	3.08	0.285	.755
Sample size (total =444)	143	153	148		
Child characteristics					
Female	50.0%	48.5%	49.0%	0.27	.764
Age	5.25	5.17	5.25	1.02	.361
KG1 (vs. KG2)	53.5%	52.1%	52.6%	0.24	.789
School readiness composite (% correct)	50.9	51.8	52.2	1.66	.190
Early literacy	43.6	44.6	46.0	3.80	.023
Early numeracy	38.8	39	40.0	1.54	.214
Social-emotional	36.3	37.2	38.4	3.61	.027
Executive function	46.4	45.9	46.3	0.21	.814
Sample size (total =3435)	1,088	1,180	1,167		

Note. TT = Teacher training condition; TTPA = Teacher training plus parental awareness condition; KG = Kindergarten; ECD = Early childhood development.

groups and the control group at baseline as occurring by chance. In addition, differences were examined within treatment conditions between schools assigned to receive text and flyer reinforcements on all baseline scores for dependent variables. One difference was detected, with teachers in the TT group who were assigned to receive text messages having slightly lower job satisfaction at baseline compared to teachers in the TT group, which was not assigned to receive text messages ($p < .05$).

Differential attrition analysis. Of the teachers working in the schools at baseline, 97 (21.8%) were no longer in the schools at follow-up (with the exception of four of these teachers who left the sample due to their school dropping out of the study). Similarly, 460 children (13.4%) transferred or left their school between baseline and follow-up (with the exception of 31 who left the sample due to their school dropping out of the study; see Appendix Table 1 of the online supplemental data). We conducted multilevel logistic regression analyses, with an indicator if the teacher or child left the study sample, to assess whether there was differential attrition of teachers or of children by treatment status (internal validity) or by characteristics (external validity). For the *teacher sample*, the question of treatment status is considered in the impact analysis because we found significant differences in teacher turnover rates (30.7% in control, 16.9% in TT,

18.7% in TTPA; $\chi^2(2) = 9.56$, $p = .008$). To assess external generalizability of the sample of teachers who stayed, we assessed baseline motivation, burnout, and job satisfaction; age; gender; education level; years of teaching experience; and private-sector status. Of these predictors, only one—baseline job satisfaction—significantly predicted teacher attrition, such that teachers with higher levels of baseline job satisfaction were less likely to leave the study sample ($b = -0.49$, $SE = 0.25$, $p < .05$). For the *child sample*, treatment status did not significantly predict whether children left the study sample, indicating that our experimental design was not compromised (14.4% in control, 13.1% in TT, 12.7% in TTPA; $\chi^2(2) = 1.59$, $p = .453$). To assess external validity of the sample of children who stayed, we assessed baseline levels of school readiness, child gender, child age, and private-sector status. We found that baseline school readiness scores predicted a lower likelihood of leaving the study sample ($b = -0.92$, $SE = 0.36$, $p < .05$) and age predicted a higher likelihood of leaving the study sample ($b = 0.15$, $SE = 0.05$, $p < .01$).

Missing data imputation. We used multiple imputation (with Stata's "ice" command) to address missing data on all missing variables, using three rounds of data collection (baseline and follow-up, as well as a second round of follow-up data). While the data are not missing completely at random, 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, when including a large set of covariates in estimating multiple chains of models, including those that predict differential attrition, assumptions of MAR have been shown to be robust. Our imputation approach met the standards of the *What Works Clearinghouse Version 4.0 Standards Handbook* (IES, 2017).

We conducted the imputation in two steps to account for the nested structure of the data. First, using a rich set of teacher demographic and background variables, outcome scores for professional well-being and classroom quality across all waves, and treatment status indicators, we imputed 20 teacher-level data sets. All impact estimates on teacher and classroom-level data were computed on these 20 data sets (using Stata's "mi estimate" command). Ten teachers refused to have their classrooms videotaped at follow-up, and these data were imputed in addition to the data for teachers who left the sample. Second, we randomly selected 10 of these teacher data sets. We merged each individual data set with child outcome data and basic child demographic characteristics from all waves of data. For each of the 10 teacher data sets, we imputed 10 child data sets, resulting in 100 child-level data sets. Impact estimates on child outcomes were computed on these 100 data sets.

Impact analysis. To account for the nested, non-independent nature of the data (i.e., students nested within classrooms and classrooms nested within schools), we employed three-level (for child outcomes) and two-level (for teacher and classroom outcomes) multilevel modeling in Stata (Version 14.0). First, we estimated unconditional models to estimate the ICCs, or the proportion of variance in each of the teacher/classroom and student outcomes attributable to students, teachers/classrooms, and schools. Second, impact analyses were conducted with a select set of covariates. We nested children and teachers in the baseline schools from which they were sampled. The multiply-imputed data sets were used in all analyses with Stata's "mi estimate" command, which uses Rubin's combining rules to compute pooled coefficients and standard errors across data sets (Rubin, 1987).

Separate models were fitted to estimate main intervention impacts on (1) teacher professional well-being (i.e., motivation, burnout, and job satisfaction), (2) classroom quality factors (i.e., fidelity checklist, facilitating deeper learning, emotional support and behavior management, and supporting student expression), and (3) 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 skills, and executive function) to assess whether 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 X_{ijk} is the vector of child covariates (gender, age, and baseline score).

Level 2 (classroom-level) model:

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

Where 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 γ_{00k} is the school-level random intercept; Z_k is the vector of school-level covariates (district dummies, private or public status, and four dummy variables for different school mobility scenarios [i.e., three treatment schools combined their separate KG classrooms into one KG1 and one KG2 classroom; one school split into two schools between baseline and follow-up; nine schools started with a combined KG1 and KG2 class and split into two separate classrooms midyear; and 12 teachers switched to teach a different KG class within the school midyear]); TT_k is an indicator for schools assigned to the teacher training condition; and $TTPA_k$ an indicator for schools assigned to teacher training plus parental awareness.

Third, as a secondary exploratory analysis, we examined whether intervention impacts varied by child characteristics (gender, child baseline scores, and grade level [KG1 and KG2]) and by school sector (private and public). Impact variation 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.

Experimental analyses testing more than one outcome must acknowledge the issue of multiple comparisons. We analyze three contrasts for each outcome: TT vs. control, TTPA vs. control, and TT vs. TTPA. While this increases our chance of Type 1 error, because this is the first trial to test teacher training and parental-awareness programs focused on promoting play-based, child-centered ECE in West Africa and the first time measures of teacher and classroom quality have been used in the context, we do this to avoid Type 2 error. A Bonferroni correction can address this issue when all outcomes are independent from one another, but this requirement is not satisfied in our particular

case (i.e., if the treatment moves outcome A, it probably moves outcome B too). Outcomes are assessed in three theoretically informed categories—teacher professional well-being, classroom quality, and child outcomes—with nine total. We use $p < .05$ as a threshold for statistical significance, and thus treatment impacts on at least one outcome for each treatment condition would need to be significant to be considered better than chance. In consideration of the issue of multiple comparisons, we are discerning in interpreting results based on the pattern of impacts across each of the three sets of outcomes and do not interpret our secondary exploratory questions as part of the primary analysis.

Results

The descriptive statistics for all outcome variables and their intercorrelations are presented in Table 2, and ICCs for all outcomes are shown in Table 3. The majority of variance in teacher professional well-being and classroom quality (74%–88%) was accounted for by differences across teachers rather than across schools. The majority of variance in child outcomes was accounted for by differences across children (58%–83%) and secondarily across classrooms (14%–42%). A very small portion of the variance (0%–9%) was accounted for across schools, indicating that classrooms are more

Table 2. Descriptive statistics and bivariate correlations of outcome variables at follow-up.

	Mean	*SD*	Range	1	2	3	4	5	6
1 Child school readiness composite	0.565	0.178	0–1						
Teacher professional well-being									
2 Teacher motivation	4.71	0.44	1–5	0.035					
3 Teacher burnout	2.01	0.90	1–6	**−0.040**	**−0.174**				
4 Teacher job satisfaction	3.08	0.68	1–4	−0.017	**0.130**	**−0.284**			
Observed classroom quality									
5 Facilitating deeper learning	2.39	0.65	1–4	**0.042**	−0.076	−0.051	0.031		
6 Emotional support and behavior management	3.07	0.37	1–4	**0.104**	−0.052	−0.080	0.063	**0.304**	
7 Supporting student expression	1.75	0.67	1–4	−0.035	−0.052	0.014	0.000	**0.360**	**0.157**

Notes. Bold numbers indicate that correlation is statistically significant at $p < .05$. Correlations with school readiness use child-level data ($N = 3435$); correlations among teacher variables include teacher-level data ($N = 444$).

Table 3. Variance partition for teacher/classroom and child outcomes at follow-up.

	Proportion of variance		
	Child	Teacher/classroom	School
Teacher professional well-being			
Motivation	–	0.874	0.126
Burnout	–	0.883	0.117
Job satisfaction	–	0.731	0.269
Classroom quality			
Facilitating deeper learning	–	0.778	0.222
Emotional support and behavior management	–	0.867	0.133
Supporting student expression	–	0.852	0.148
Child outcomes			
School readiness composite	0.581	0.419	0.000
Early numeracy	0.622	0.378	0.000
Early literacy	0.543	0.371	0.086
Social-emotional	0.826	0.141	0.034
Executive function	0.765	0.199	0.037

important than schools in explaining variance across child outcomes. This may largely be the case because the classrooms within each school were of different grade levels. The results for the two primary research questions (RQ1 and RQ2) are presented first, followed by results for the two secondary research questions (RQ3 and RQ4).

Teacher Training Attendance

We first present the average number of days teachers attended the training and received coaching visits as an overview on implementation. Of the teachers selected to participate in the program at baseline, teachers in the TT arm 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 2-day refresher training, and 0.7 days ($SD = 0.5$) of the final 1-day refresher. In the TTPA arm, 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 ($SD = 0.9$) of the first 2-day refresher training, and 0.8 days ($SD = 0.4$) of the final 1-day refresher. Of note, a few teachers in the control group ($n = 3$) 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.

Primary Research Questions (RQ1 and RQ2)

We first present the results addressing the first two research questions, focused on main impacts of the TT condition relative to the control group, and impacts of the TTPA condition relative to the control group and the TT condition.

Impacts on Teacher Professional Well-being. Table 4 displays 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 (d_{wt}[1] $= -.40$, $p < .01$) and the TTPA condition ($d_{wt} = -0.59$, $p < .001$). In addition, the TT condition impacted teacher turnover, reducing the probability that a teacher would leave the KG classroom by the third term by 29.6% ($OR = 0.42$, $p < .05$). There were marginally statistically significant impacts of the TTPA condition on turnover ($OR = 0.53$, $p < .10$).

[1] d_{WT} represents a standardized mean difference between treatment and control schools. 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 school, classroom/teacher, 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 data sets using Rubin's combining rules. This same approach was utilized to estimate d_{WT} for this and other main effects presently reported.

Table 4. Impacts on teacher professional well-being and classroom quality.

	b	(*SE*)	*p* value	d_{wt} or *OR*
Teacher professional well-being				
Motivation				
TT	0.103	(0.068)	.132	0.345
TTPA	0.033	(0.072)	.648	0.111
Burnout				
TT	−0.330	(0.155)	.035*	−0.398
TTPA	−0.492	(0.159)	.002**	−0.593
Job satisfaction				
TT	0.130	(0.096)	.175	0.328
TTPA	0.043	(0.100)	.663	0.108
Teacher turnover[a]				
TT	−0.868	(0.414)	.036*	0.420[b]
TTPA	−0.629	(0.351)	.073+	0.533[b]
Classroom outcomes				
Fidelity checklist (no. of activities)				
TT	1.310	(0.246)	.000***	0.560
TTPA	1.434	(0.266)	.000***	0.613
Facilitating deeper learning				
TT	−0.045	(0.106)	.672	−0.113
TTPA	−0.063	(0.116)	.588	−0.158
Emotional support and behavior management				
TT	0.170	(0.065)	.010**	0.647
TTPA	0.172	(0.066)	.010**	0.655
Supporting student expression				
TT	0.235	(0.110)	.033*	0.524
TTPA	0.042	(0.116)	.719	0.094

Note. TT = teacher training condition; TTPA = teacher training plus parent-awareness meetings condition.
Estimates are computed using observed scores, in two-level models: teachers nested in schools. Effect sizes are calculated accounting for the two-level model structure (Hedges, 2009).
Sample size for TT vs. control =296 teachers nested in 161 schools. Sample size for TTPA vs. control =291 teachers nested in 158 schools. All impact estimates computed from 20 multiply imputed data sets. Models include the following control variables: private- (vs. public-) sector status of the school, six district dummies, a dummy variable for whether the school was assigned to receive teacher text messages, a dummy for whether the school was assigned to receive parent flyers, and a series of five dummy variables accounting for within-sample mobility, teacher gender, age, level of education, and years of teaching experience. Models for teacher professional well-being outcomes also include the baseline score for each respective outcome.
$+p < .10$.
$*p < .05$.
$**p < .01$.
$***p < .001$.
[a]Model estimated with multinomial logistic regression.
[b]Odds ratio presented.

To test differences between the TT and TTPA conditions, all models were re-run with TT as the reference group. There were no statistically significant differences between TT and TTPA arms, indicating that despite some differences in the magnitudes of the impacts estimates across the two treatment conditions, impacts were not statistically different.

Impacts on Classroom Outcomes. Table 4 also 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 ($d_{wt} = 0.56$ in TT and 0.61 in TTPA, $p < .001$). In the control group, 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.

Table 5. Impacts on children's school readiness outcomes.

	b	(*SE*)	*p* value	d_{wt}
School readiness composite				
TT	0.025	(0.010)	.010**	0.163
TTPA	0.004	(0.010)	.691	0.026
Post-hoc estimates by domain				
Early numeracy				
TT	0.020	(0.009)	.031*	0.107
TTPA	−0.005	(0.010)	.609	−0.027
Early literacy				
TT	0.022	(0.011)	.045*	0.110
TTPA	−0.006	(0.013)	.670	−0.030
Social-emotional				
TT	0.035	(0.013)	.010**	0.180
TTPA	0.014	(0.013)	.286	0.128
Executive function				
TT	0.020	(0.014)	.148	0.109
TTPA	0.007	(0.014)	.589	0.038

Note. TT = teacher training condition; TTPA = teacher training plus parent-awareness training condition.
Estimates are computed using observed scores, in three-level models: children nested in classrooms nested in schools. Effect sizes calculated accounting for the three-level model structure (Hedges, 2009).
Sample size for TT vs. control =2,268 children nested in 296 teachers nested in 161 schools. Sample size for TTPA vs. control =2,255 children nested in 291 teachers nested in 158 schools. All impact estimates computed from 100 multiply imputed data sets. Models include the following control variables: private- (vs. public-) sector status of the school, six district dummies, a dummy variable for whether the school was assigned to receive teacher text messages, a dummy for whether 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.
*$p < .05$. +$p < .10$.

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 ($d_{wt} = 0.65$ in the TT condition and 0.66 in the TTPA condition, $p < .001$). Finally, the TT condition increased levels of supporting student expression in classrooms ($d_{wt} = 0.52$, $p < .01$), but there were no statistically significant impacts in 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. Thus, even though some of the coefficient estimates from the TT arm were significantly different from zero, they were not significantly different from the insignificant estimate for the TTPA arm given the degree of precision of both estimates. For example, for *supporting student expression*, the coefficient estimate for the TT arm is statistically different from zero ($b = 0.235$, $SE = 0.110$, $p < .05$) but not statistically different from the TTPA arm impact estimate ($b = 0.042$, $SE = 0.116$, $p > .10$) given the standard errors of the two point estimates.

Impacts on Child School Readiness Outcomes. Table 5 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 whether the findings were driven by any particular developmental domains. The TT program increased children's school readiness ($d_{wt} = .16$, $p < .01$). When broken down by domain, impacts were statistically significant for three of the four domains, including early numeracy ($d_{wt} = .11$, $p < .05$), early literacy ($d_{wt} = .11$, $p < .05$), and social-emotional skills ($d_{wt} = .18$, $p < .01$). There were no impacts on children's school readiness in the TTPA condition.

Compared to the TT condition directly, children in the TTPA condition had marginally statistically lower scores on overall school readiness ($b = -.021$, $SE = 0.011$, $p < .06$). When considered by domain, TTPA had lower scores on early numeracy ($b = -.023$, $SE = 0.011$, $p < .05$) and marginally statistically lower scores on early literacy ($b = -.027$, $SE = .015$, $p < .07$). Social-emotional skills were lower but not statistically different ($b = -.021$, $SE = .013$, $p = .102$).

Secondary Research Questions

Impacts of Reinforcements (RQ3). Our third research question was to assess whether added reinforcements to teachers via text message or the flyers to parents on school readiness strengthened the impact of each program. These indicators were included in the main models and are presented in Appendix Table 3 in the online supplement. We found no added impacts on any outcome of the text message reinforcements to teachers or flyers to parents.

Moderation by Child Characteristics and Public- and Private-Sector Schools (RQ4). Our fourth research question was concerned with impact variation. We assessed whether impacts on school readiness were moderated by three child characteristics: gender, baseline school readiness, and grade level (KG1 and KG2). We found no significant interactions between treatment status for any child characteristics (results shown in Appendix Table 4 in the online supplement).

We then assessed whether program impacts on teacher, classroom, and child outcomes were moderated by school status (i.e., public vs. private sector; results shown in Appendix Table 5 in the online supplement). Of the eight primary outcomes assessed, we found two statistically significant interactions between treatment status and public- vs. 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 ($b = -0.42$, $SE = .20$, $p < .05$ and $b = -0.45$, $SE = .20$, $p < .05$, respectively). Results indicate that impacts on reduced burnout were larger in private schools.

Second, the interaction term from the logistic regression model predicting teacher turnover between private-school status and the TT condition was statistically significant ($b = -1.04$, $SE = .14$, $p < .001$), as well as between private-school status and the TTPA ($b = -1.13$, $SE = .14$, $p < .001$). Figure 3 illustrates the nature of these differences, showing predicted probability of teacher turnover by treatment condition in private- and public-sector schools separately. The treatment reduced the predicted probability of

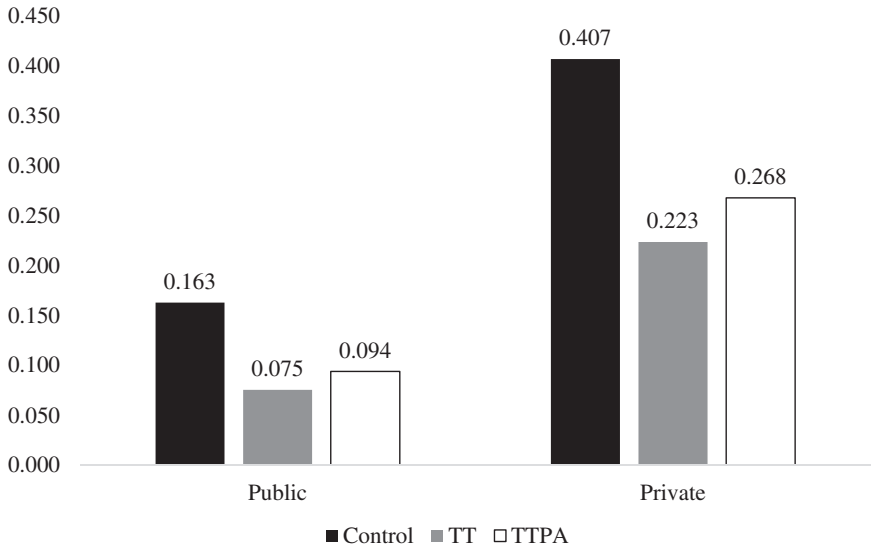

Figure 3. Predicted probability of teacher turnover by treatment status for public and private schools.

teacher turnover in private schools to a larger degree, from 40.7% to 22.3% (TT condition) and to 26.8% (TTPA condition).

Discussion

This article presented results from an impact evaluation of in-service teacher training, with and without parental-awareness meetings, in preschools on teachers, classrooms, and children in Ghana at the end of one school year of intervention. Moving beyond the question of whether access to early childhood education improves child outcomes (i.e., "Education for All"), we considered the question of how to improve educational quality and learning outcomes (i.e., "Learning for All") in a country where ECE enrollment rates are among the highest in sub-Saharan Africa (UNESCO, 2015). Initiatives to improve quality are particularly needed given the increased investment by governments of LMICs and by international donors in the preschool period of early childhood development (e.g., Britto et al., 2011) and the growing commitment to Sustainable Development Goal 4 (United Nations, 2015). To our knowledge, this is one of the first studies in sub-Saharan Africa to assess impacts of a pre-primary 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.

Teacher Training and Coaching

We found moderate impacts of the teacher training and coaching on some dimensions of teacher professional well-being (reduced burnout, $d = -0.40$ and -0.59, and 58% reduction in the odds of midyear job turnover) and on improved classroom quality ($d = 0.52$–0.66). The impacts on improved professional well-being are noteworthy, in particular in a context where teaching professionalization is lacking and turnover rates

are high (Bennell & Akyeampong, 2007; Guajardo, 2011; Osei, 2006) and particularly in the private sector, where teachers have fewer educational qualifications, are younger, and receive much lower remuneration (Wolf, Tsinigo, Behrman & Aber, 2016). We found small impacts on improved children's school readiness ($d = 0.16$).

Two domains of classroom quality were impacted—supporting student expression and emotional support and behavior management—but not the third: facilitating deeper learning. Of importance, the facilitating deeper learning factor had low internal reliability, and thus we cannot rule out the possibility that the lack of impacts on this domain was due to poor reliability of the construct. Research in Chile found similar effects of in-service teacher training on observed levels of classroom emotional support but not instructional support (Yoshikawa et al., 2015), concluding that a focus on behavior management, along with teachers' perceptions that they were receiving support, may have led to increased warm and respectful interactions and positive emotions and expectations in the classroom. Of note, these did not translate into improved child outcomes. Similarly, Özler et al. (2018) found that an intensive 5-week teacher training in Malawi childcare centers improved classroom quality but not child outcomes. Thus, it is notable that the less-intensive training evaluated in this study improved both classroom quality and children's outcomes, albeit using a different set of measures for both.

The QP4G training included instruction at the start of the school year (with two shorter refresher trainings) and in-classroom coaching and mentoring over the course of the school year, all implemented by local professionals (teacher trainers and district government education coordinators). The in-service training and coaching helped teachers incorporate play-based and child-centered methods into literacy and numeracy lessons, as well as improved behavior- and classroom-management skills. The trainings did not focus on academic content (e.g., the importance of teaching phonological awareness as an early literacy skill), but rather centered on pedagogical approaches to teaching such content (e.g., incorporating activities and games into instruction), as well as incorporating positive behavior management practices to foster school readiness behavioral skills. Nonetheless, post hoc analyses indicated that both academic (literacy and numeracy) and social-emotional skills improved. This pattern of results converges with findings from studies outside of Ghana that improving the processes of supportive teaching can improve children's school readiness (e.g., Morris et al., 2014) and provides evidence of similar processes in a Ghanaian context. The findings are also consistent with research in the United States that finds that improvements in children's social-emotional skills are related to improvements in academic outcomes (Durlak et al., 2011). Future data collection will provide evidence on whether these impacts are sustained into the next academic year.

The medium-sized effects we observed for teacher and classroom measures are similar to those found in ECE teacher professional development interventions in high-income countries (e.g., Morris et al., 2014; Raver et al., 2008). The small effect sizes observed for child outcomes align with related ECE interventions in the United States and with other educational interventions in LMICs that have been found to improve child learning outcomes (McEwan, 2015). While classroom process quality is considered a major "driver" of child learning (e.g., Pianta et al., 2005), the question naturally arises that with moderate impacts on classroom outcomes, why were impacts on child outcomes smaller? A recent meta-analysis summarized the research on ECE teacher professional

development interventions and found a similar pattern of results, with an average effect size on classroom process quality of $d = 0.68$ and child outcomes of $d = 0.14$ (Egert, Fukkink, & Eckhardt, 2018). The reasons for this discrepancy are not known, although given the large role of family and home characteristics in promoting child learning (e.g., Connor et al., 2005), the ability of a classroom-based intervention to improve child learning may be limited. In addition, the large majority of research on classroom process quality has been conducted in high-income countries. Conceptualization and operationalization of process quality in sub-Saharan African ECE contexts is much less developed. Thus, it is plausible that our measures do not account for all classroom quality characteristics necessary to promote child learning outcomes in Ghana.

Future initiatives should focus on how to translate the large improvements in teaching quality into 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 with which the field must continue to grapple. Identifying the classroom processes that most strongly predict learning in this context should be a priority so that low-intensity interventions can be targeted strategically. Furthermore, improving the home and community drivers of learning may also be of high priority.

The impact of the teacher training on school readiness skills was equally effective for boys and girls and KG1 and KG2 children and for children who were more or less school-ready at baseline. The equally positive effects for boys and girls and for the relatively more school-ready and less school-ready children also suggest that the program did not increase inequalities among the targeted, relatively disadvantaged, population, in contrast to some recent results for schooling programs in Bangladesh, another developing country (Behrman, 2015). But also, equally positive effects for boys and girls and for children of different levels of school readiness mean that the program did not reduce inequalities among the targeted, relatively poor, population. Attention to who benefits from educational programs, as well as how to ensure the most marginalized benefit, is a critical area for research to continue to consider.

Our findings also suggest that there are significant gains from teacher training in both private and public schools and larger reductions in teacher burnout and turnover in the private sector. In this sample, teachers in the private sector had lower educational qualifications and less teaching experience and were significantly younger compared with public-sector teachers. The training may have improved self-efficacy among private school teachers who had relatively fewer skills and training at baseline than public school teachers. These analyses are not part of the primary questions in this study, and as a result we consider them "hypothesis generating" rather than hypothesis testing. Future research should consider differential needs of professional development for teachers in public and private schools.

Parental-Awareness Meetings

Contrary to our prediction, we found that adding three parental-awareness meetings, administered through school PTAs by local government district coordinators, did not

improve the effectiveness of the teacher training. Rather, these meetings counteracted some of the positive impacts of the teacher training, specifically on children's school-readiness outcomes. Of importance, these meetings focused on increasing parental engagement and awareness of developmentally appropriate education in early childhood as opposed to parenting skills. 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 and socialization (Kabay et al., 2017). Perhaps parents did not agree with the messages of the teacher training program to promote child-centered and play-based learning 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 meetings, some parents complained more to teachers about their child's behavior and academic problems. For example, when asked about parents reactions to the meetings, one teacher said: "When the topics were discussed, some of them did not agree. They were like 'I gave birth to my child so why shouldn't I beat him if the child is misbehaving?'" A second teacher said: "Like getting the child some learning materials so that as they are playing they can be learning at the same time. And sometimes they have to stop using the cane but a parent voiced out and said that her kids are stubborn so without the cane. … So I told her that with some parents they said the kids are always happy whenever they are around so they should avoid the cane and they will be fine." These experiences indicate that in some schools, parents pushed back on teachers when they did not agree with the messages promoted in the meetings and perhaps advocated for teachers to continue to use the old methods.

Alternatively, perhaps the counteracting effect was due to the medium of the parental-awareness meetings. The trainings 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 trainings were not implemented with fidelity and that parents' experiences varied widely based on the district education coordinator who was implementing the program. Poor implementation may have led to parents to feel frustrated with the school, since many schools held the meetings during the day and parents had to leave work to attend. Thus, our conclusion 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 reaches parents.

Notably, a recent study in Malawi found that a more intensive, 12-module, group-based parenting-support program focused on parenting skills administered through childcare centers by teachers and their mentors combined with intensive teacher training was effective in improving early childhood developmental outcomes (Özler et al., 2018) This suggests that more intensive parenting programs administered through schools by local (not district) personnel can be effective. However, it is possible that such programs need to have frequent meetings for parents to internalize the messages. In addition, this program, like many parenting programs that have been studied to date, focused on younger children (18 months old), while the children in the QP4G intervention were 4 to 6 years old. Parents may be less receptive to messages of cognitive

stimulation and play for older children, and perhaps the implications of these practices for cognitive and academic benefits need to be emphasized in order to be appealing to parents.

Finally, we tested the added value of providing teachers and parents with reinforcements of the messages of the trainings via biweekly text messages for teachers and picture-based paper flyers delivered to parents three times in the second and third terms of the academic year. We found no impacts of these additional "nudge-like" reinforcements of the training messages.

Limitations and Conclusions

These findings must be interpreted within the limitations of the study. First, the study's sample is limited in its generalizability to the six peri-urban districts in the Greater Accra Region in Ghana included in the sample. Second, due to time and resource constraints, we collected very few data on the implementation of the parental-awareness training and parents' engagement in and perceptions of this training. Thus, we are left speculating about the unexpected findings of the parental-awareness meetings due to what we were not able to observe. Similarly, teachers were observed for only one class at both time points, limiting our understanding of how the intervention changed teaching practices to two snapshots of the classroom. Furthermore we were unable to reliably measure a key dimension of classroom quality: facilitating deeper learning.

Our findings suggest that brief and relatively affordable in-service teacher training, built into existing governmental systems, can improve key dimensions of classroom quality and early childhood development in both private and public schools in Ghana. In QP4G, teachers not only incorporated the specific activities taught in the training program but also improved two dimensions of teacher–child interactions. Although our findings are promising, they are also somewhat disappointing in that they only translated into small effect sizes on children's school readiness that are, however, comparable to results of previous studies. This points to the potential need to combine home-based interventions with school-based interventions, although it raises caution on how home-based interventions are implemented. There remain many important questions about how to translate large program effects on teachers and classrooms into large effects on young children's learning and development that will have lasting impacts on lifelong learning.

Acknowledgments

This article 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, and Shams Muzaffar, and our talented data collection supervisors and enumerators. We also thank the Ghana Education Service for their support of this project. And we would like to thank the dedicated teachers, children, and families who participated in the study. We thank the UBS Optimus Foundation and the World Bank Strategic Impact Evaluation Fund (SIEF) for 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 second author for research activities that directly contributed to this article.

Funding

World Bank Strategic Impact Evaluation Fund, Grant ID#WBK5595 and UBS Optimus Foundation, Grant ID#UBS9307

ARTICLE HISTORY

Received 20 October 2017
Revised 1 June 2018
Accepted 7 August 2018

References

Abdul-Hamid, H., Baum, D. R., De Brular, L. L., Lusk-Stover, O., & Tettey, L. O. (2017). *SABER engaging the private sector in education pilot country report: Ghana 2015. Systems Approach for Better Education Results (SABER) country report: Engaging the Private Sector*. Washington, D.C.: World Bank Group.

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., & Schady, N. (2016). Teacher quality and learning outcomes in kindergarten. *Quarterly Journal of Economics*, *131*(3), 1415–1453.

Behrman, J. A. (2015). Do targeted stipend programs reduce gender and socioeconomic inequalities in schooling attainment? Insights from rural Bangladesh. *Demography*, *52*(6), 1917–1927.

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.

Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist*, *57*(2), 111–127.

Bridgemohan, R., van Wyk, N., & van Staden, C. (2005). Home-school communication in the early childhood development phase. *Education*, *126*(1), 60.

Britto, P. R., Yoshikawa, H., & Boller, K. (2011). Quality of early childhood development programs in global contexts: Rationale for investment, conceptual framework and implications for equity. *Social Policy Report*, *25*(2), 3–23.

Brown, J. L., Jones, S. M., LaRusso, M. D., & Aber, J. L. (2010). Improving classroom quality: Teacher influences and experimental impacts of the 4rs program. *Journal of Educational Psychology*, *102*(1), 153–167.

Central Intelligence Agency. (2016). Ghana. *The World Factbook*. Retrieved from http://www.cia-worldfactbook.us/africa/ghana.html

Collins, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research : an International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *12*(3), 229–238.

Connor, C. M., Son, S., Hindman, A. H., & Morrison, F. J. (2005). Teacher qualifications, classroom practices, family characteristics, and preschool experience: Complex effects on first graders' vocabulary and early reading outcomes. *Journal of School Psychology*, *43*(4), 343–375.

Egert, F., Fukkink, R. G., & Eckhardt, A. G. (2018). Impact of in-service professional development programs for early childhood teachers on quality ratings and child outcomes: A meta-analysis. *Review of Educational Research*, *88*(3), 401–433.

Engle, P. L., Fernald, L. C. H., Alderman, H., Behrman, J., O'Gara, C., Yousafzai, A., … Iltus, S. (2011). Strategies for reducing inequalities and improving developmental outcomes for young children in low-income and middle-income countries. *The Lancet*, *378*(9799), 1339.

Ganimian, A. J., & Murnane, R. J. (2016). Improving education in developing countries: Lessons from rigorous impact evaluations. *Review of Educational Research*, *86*(3), 719–755.

Ghana Education Service. (2012). *Programme to Scale-Up Quality Kindergarten Education in Ghana*. Ghana Ministry of Education. Accra: Ghana. Retrieved from https://issuu.com/sabre-tom/docs/10_12_12_final_version_of_narrative_op

Ghana Ministry of Education. (2014). *Ghana 2013 National Education Assessment: Technical Report*. Accra, Ghana.

Ghana Ministry of Education. (2016). *Education Sector Performance Report, 2016*. Ghana.

Guajardo, J. (2011). *Teacher Motivation: Theoretical Framework, Situation Analysis of Save the Children Country Offices and Recommended Strategies*. Washington, DC: Save the Children.

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.

Institute for Education Science (IES). (2017). *WWC Version 4.0 Standards Handbook*. IES: Washington, D.C. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v4.pdf

Jennings, P. A., Frank, J. L., Snowberg, K. E., Coccia, M. A., & Greenberg, M. T. (2013). Improving classroom learning environments by Cultivating Awareness and Resilience in Education (CARE): Results of a randomized controlled trial. *School Psychology Quarterly*, *28*(4), 374–390.

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.

Kraft, M. A., Blazar, D., & Hogan, D. (2017). *The Effect of Teacher Coaching on Instruction and Achievement: A Meta-Analysis of the Causal Evidence*. Brown University Working Paper.

Lu, C., Black, M. M., & Richter, L. M. (2016). Risk of poor development in young children in low-income and middle-income countries: An estimation and analysis at the global, regional, and country level. *The Lancet Global Health*, *4*(12), e916–e922.

La Paro, K. M. L., Hamre, B. K., Locasale-Crouch, J., Pianta, R. C., Bryant, D., Early, D., et al. (2009). Quality in kindergarten classrooms: Observational evidence forthe need to increase children's learning opportunities in early education classrooms. *Early Education and Development*, *20*, 657–692.

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

McCoy, D. C., Peet, E. D., Ezzati, M., Danaei, G., Black, M. M., Sudfeld, C. R., … Fink, G. (2016). Early childhood developmental status in low-and middle-income countries: National, regional, and global prevalence estimates using predictive modeling. *PLoS Medicine*, *13*(6), e1002034.

McEwan, P. J. (2015). Improving learning in primary schools of developing countries: A meta-analysis of randomized experiments. *Review of Educational Research*, *85*(3), 353–394.

Miedel, W. T., & Reynolds, A. J. (1999). Parent involvement in early intervention for disadvantaged children: Does it matter? *Journal of School Psychology*, *37*(4), 379–402.

Moon, B. (2007). School-based teacher development in sub-Saharan Africa: Building a new research agenda. *Curriculum Journal*, *18*(3), 355–371.

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*(1), 38–51.

Özler, B., Fernald, L. C. H., Kariger, P., McConnell, C., Neuman, M., & Fraga, E. (2018). Combining preschool teacher training with parenting education: A cluster-randomized controlled trial. *Journal of Development Economics*, *133*, 448–467.

Phillips, D., Mekos, D., Scarr, S., McCartney, K., & Abbott–Shim, M. (2000). Within and beyond the classroom door: Assessing quality in child care centers. *Early Childhood Research Quarterly*, *15*(4), 475–496.

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, *38*(2), 109–119.

Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science*, *9*(3), 144.

Piper, B., & Zuilkowski, S. S. (2015). Teacher coaching in Kenya: Examining instructional support in public and nonformal schools. *Teaching and Teacher Education*, *47*, 173–183.

Pisani, L., Borisova, I., & Dowd, A. J. (2015). *International development and early learning assessment technical working paper*. Washington, DC: Save the Children.

Raver, C. C. (2002). Emotions matter: Making the case for the role of young children's emotional development for early school readiness. *Social Policy Report*, *16*(3), 1–18.

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.

Republic of Ghana. (2015). Education management information system: Basic statistics and planning parameters for basic education in Ghana 2014/2015. Ministry of Education.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: Wiley.

Seidman, E., Kim, S., Raza, M., Ishihara, M., & Halpin, P. (2018). Assessment of pedagogical practices and processes in low and middle income countries: Findings from secondary school classrooms in Uganda. *Teaching and Teacher Education*, *71*, 283–296.

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: NYU.

Tsouloupas, C. N., Carson, R. L., Matthews, R., Grawitch, M. J., & Barber, L. K. (2010). Exploring the association between teachers' perceived student misbehaviour and emotional exhaustion: The importance of teacher efficacy beliefs and emotion regulation. *Educational Psychology*, *30*(2), 173–189. Retrieved 2016 from http://data.worldbank.org/indicator/SE.ENR.PRSC.FM.ZS

UNESCO. (2015). *Education for all global monitoring report regional overview: Sub-Saharan Africa*. Paris: UNESCO.

United Nations. (2015). *Transforming our world: The 2030 agenda for sustainable development*. New York: Author.

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., & McCoy, D. C. (2017). Household socioeconomic status and parental investments: Direct and indirect relations with school readiness in Ghana. *Child Development*. Advance online publication. doi:10.1111/cdev.12899

Wolf, S., Raza, M., Kim, S., Aber, J. L., Behrman, J., & Seidman, E. (2018). Measuring process quality in pre-primary classrooms in Ghana using the Teacher Instructional Practices and Processes System (TIPPS). *Early Childhood Research Quarterly*, *45*, 18–30.

Wolf, S., Torrente, C., Frisoli, P., Weisenhorn, N., Shivshanker, 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*, *52*, 24–36.

Wolf, S., Tsinigo, E., Behrman, J. R., & Aber, J. L. Testing and scaling-up supply- and demand side interventions to improve kindergarten educational quality in Ghana: Baseline report. *World Bank Strategic Impact Evaluation Fund*. Washington, D.C.

Yoshikawa, H., Leyva, D., Snow, C. E., Treviño, E., Barata, M. C., 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.

Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M. R., Espinosa, L. M., Gormley, W. T., … Zaslow, M. J. (2013). *Investing in our future: The evidence base on preschool education*. New York: Foundation for Child Development.

Young, R., & Johnson, D. R. (2015). Handling missing values in longitudinal panel data with multiple imputation. *Journal of Marriage and Family*, *77*(1), 277–294.