Received: 5 June 2018 | Revised: 20 April 2019 | Accepted: 30 April 2019

DOI: 10.1111/desc.12878

SPECIAL ISSUE ARTICLE

Developmental Science WILEY

Longitudinal causal impacts of preschool teacher training on Ghanaian children's school readiness: Evidence for persistence and fade-out

Sharon Wolf[1] | J. Lawrence Aber[2] | Jere R. Behrman[1] | Morgan Peele[1]

[1]University of Pennsylvania, Philadelphia, Pennsylvania

[2]New York University, New York, New York

Correspondence
Sharon Wolf, 3700 Walnut Street, Philadelphia, PA 19104.
Email: wolfs@upenn.edu

Funding information
UBS Optimus Foundation, Grant/Award Number: UBS9307; World Bank Group, Grant/Award Number: 7182260 and WBK5595

Abstract

Preschool programs have expanded rapidly in low- and middle-income countries, but there are widespread concerns about whether they are of sufficient quality to promote children's learning and development. We conducted a large school-randomized control trial ('Quality Preschool for Ghana' – QP4G) of a one-year teacher training and coaching program, with and without parental-awareness meetings, designed to improve preschool quality and child development. We followed 3,435 children in 240 schools in the Greater Accra Region of Ghana, a country with universal pre-primary education. A previous study reported positive impacts of teacher training (but not teacher training plus parental-awareness meetings) at the end of the implementation year on some dimensions of classroom quality, teacher well-being, and children's school readiness (Wolf et al., [2019] *Journal of Research on Educational Effectiveness*, 12, 10–37). The present study analyzed a new round of data collected 1 year after the end of implementation to assess (a) the extent of persistence in impacts on child development and (b) whether such impacts vary by select child, household, and school characteristics. We found impacts of the teacher training intervention on children's overall school readiness were sustained (d = 0.13), but were only marginally statistically significant. When broken down by domain, impacts on social–emotional skills specifically persisted. Persistent negative effects of teacher training plus parental-awareness meetings varied by the literacy status of the male parent such that negative impacts were concentrated in children in households with non-literate male heads.

KEYWORDS

Ghana, persistence, preschool, quality, randomized control trial, sub-Saharan Africa, teacher training

1 | INTRODUCTION

Investments in early childhood education (ECE) by both parents and governments in low- and middle-income countries (LMICs) have increased over the past decade (Yoshikawa, Wuermli, Raikes, Kim, & Kabay, 2018). Global goals for expansion of pre-primary education services are reflected in the United Nations' Sustainable Development Goal 4 on Quality Education, which specifies that all children have access to 1 year of pre-primary education so that they arrive at primary school ready to learn (Indicator 4.2.1; United Nations, 2015). While access has expanded, concerns exist about service quality, including whether ECE services are structured to

promote young children's learning and development and to satisfy parents' desires for and perceptions of quality education (Britto, Yoshikawa, & Boller, 2011). Thus, whether the expansion of ECE in LMICs is improving child development and school readiness is unclear.

We address two critical questions about ECE quality improvements. *First, do efforts to improve existing ECE quality that have positive short-run effects (i.e. over one school year) also have sustained gains for children*? The longer-term persistence or fade-out of gains from ECE has received some attention in the United States (Bailey, Duncan, Odgers, & Yu, 2017; Jenkins et al., 2018), and growing attention in LMICs (Berlinksi, Galiani, & Gertler, 2009). Given limited resources for investing in education, ECE programs that have lasting impacts are more cost-effective than programs for which impacts fade after exposure to the program ends. *Second, do sustained ECE impacts differ by select child and household characteristics and in private* versus *public schools?* This question is important because governments may be particularly interested in improving educational outcomes for marginalized groups, including girls and children from poorer socioeconomic backgrounds, and because the prevalence of private preschools is increasing rapidly (Bidwell, Parry, & Watine, 2014).

One key reason for examining sustained effects of quality ECE comes from a few prominent studies showing strikingly positive impacts in adulthood and large benefit–cost ratios of small-scale interventions, including the Perry Preschool (e.g. Schweinhart et al., 2005; Heckman, Moon, Pinto, Savelyev, & Yavitz, 2010) and Abecedarian studies (e.g. Ramey & Campbell, 1979; Campbell et al., 2012). The larger experimental literature analyzing longer-term impacts of ECE quality improvements in high-income countries has found mixed results for persistent impacts 1 year after interventions (e.g. Ansari, 2018; Barnett et al., 2018; Jenkins et al., 2018). The Head Start Impact Study in the Unites States was the first evaluation of a large-scale public program. Head Start produced modest positive effects on immediate cognitive and social–emotional outcomes (Puma et al., 2010; Shager et al., 2013), but impacts on longer-term outcomes have been mixed (Puma et al., 2010, 2012). Similar results – with positive gains for children at the end of treatment but fade-out in early elementary school – are consistent with earlier work (Barnett, 1995; Deming, 2009), though a follow-up study of Tulsa's universal pre-K program found sustained impacts on academic outcomes in middle school (Gormley, Phillips, & Anderson, 2018). Importantly, a recent study re-analyzed data from the Head Start Impact Study and found significant variation in impacts based on child, center, and neighborhood characteristics (Morris et al., 2018).

The potential for ECE to support children's development in a sustained way may be larger in parts of the world where children face extreme levels of risk. Compared to other regions, sub-Saharan Africa (SSA) has the largest number of young children experiencing malnutrition and poverty (Black et al., 2017), as well as the largest number and proportion of 3- and 4-year-olds (29.4 million; 44%) failing to meet cognitive and social–emotional milestones (McCoy et al., 2016). Given extremely low learning levels for primary school students across SSA countries (Sandefeur, 2016), ECE may be one approach to boosting children's school readiness and ultimately improving longer-term learning outcomes.

Research Highlights

- One of the first studies in sub-Saharan Africa to test persistence of impacts of a preschool teacher training program on child development 1 year after program ended.
- Small impacts on children's overall school readiness persisted (effect size = 0.13).
- The counteracting negative effects of adding a parental-awareness intervention on children's school readiness also persisted but varied by literacy status of the male household head.
- Findings have implications for future research on the persistence and fade-out of preschool interventions on children's development in sub-Saharan Africa.

The Republic of Ghana is a unique context to study these issues at the interface of developmental science, intervention science, and educational policy and practice (Wuermli, Tubbs, Peterson, & Aber, 2015). The government has been investing in 2 years of universal pre-primary education since 2007, and Ghana has among the highest net enrollment rates in Africa (UNESCO, 2015). At the same time, levels of at-home stimulation provided to Ghanaian children are low. Compared to 13 other countries categorized as medium on the Human Development Index (a composite indicator of life expectancy, education, and income per capita), Ghana ranked 12th in terms of the levels of cognitive stimulation parents reported engaging in with their child (Bornstein & Putnick, 2012).

Wolf, Aber, Behrman, and Tsinigo (2019) undertook a previous study to examine short-term impacts during a single school year of a preschool quality improvement – the Quality Preschool for Ghana (QP4G) program. This cluster-randomized trial included 240 schools (teachers N = 444; children N = 3,345, M_{age} = 5.2) randomly assigned to: 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. At the end of treatment, moderate impacts were found on some dimensions of teacher professional well-being, classroom, and multiple domains of children's school readiness. The parental-awareness meetings had counteracting effects on child school readiness. In this study, we analyze longer-run impacts on children's school readiness, 1 year after completion of the QP4G program.

1.1 | Pre-primary education and child development in LMICs

Low-income countries (e.g. Haiti, Nepal, and Uganda) are defined as those with a gross national income (GNI) per capita of $995 or less in 2017 and lower middle-income economies (e.g. Ghana, India,

and Vietnam) between $996 and $3,895 (World Bank, 2018). Over the past 15 years, there has been a rapid expansion of ECE services around the world (UNESCO, 2015). However, enrollment and quality of services have been generally lower in LMICs (Neuman, Josephson, & Chua, 2015). Furthermore, ECE enrollment rates vary widely across regions, with an average of 18% of 3- and 4-year-olds in sub-Saharan Africa (SSA) compared to 62% in Latin America and the Caribbean (McCoy et al., 2018).

Research on ECE quality is less established. The literature on quality improvement interventions is nearly exclusively focused on the United States and other high-income countries. However, there are a small and increasing number of studies in LMICs, mostly focused on short-run impacts, typically within a year of program initiation. Yoshikawa et al. (2015) found positive impacts of a teacher-workshop and coaching intervention on classroom quality that did not translate into gains for Chilean child development outcomes over two school years of implementation. (Note that Chile is now classified as a high-income country.) Araujo, Carneiro, Cruz-Aguayo, and Schady (2016) randomly assigned Ecuadorian children to kindergarten teachers across 204 schools. A one-standard-deviation increment in classroom quality predicted growth in children's language, math, and executive function of 0.11, 0.11, and 0.07 standard deviations over one school year, respectively. Finally, a recent experimental study in Malawi found that a 1-year teacher training program delivered alone in child-care centers led to improved classroom quality but no gains in child developmental at 18 and 36 months after baseline (Özler et al., 2018). When delivered with a parenting intervention, there were improvements of 0.2–0.3 *SD* in children's language skills and prosocial behaviors at the 18-month follow-up only (Özler et al., 2018).

1.2 | Pre-primary effects for different subgroups

Examining whether certain types of children, or children in different types of schools, benefit more from exposure to high-quality ECE is also a needed area of investigation. In this study, we consider impact heterogeneity for three sets of subgroup characteristics based on previous literature and the context: school (i.e. private vs. public), household, and child characteristics. Given the scarce evidence to date on impact heterogeneity of ECE for various subgroups in SSA, we consider such tests at this stage as 'hypothesis generating' classified as civil rather than hypothesis testing.

1.2.1 | *Public* versus *private schools*

As demand for ECE has expanded, in particular in peri-urban communities surrounding large cities, the private sector has grown significantly and helped fill large gaps in communities where public preschools have been slower to open (Bidwell et al., 2014; Lewis, 2013). In Ghana, the public and private sectors differ in terms of structural characteristics. For example, there are no requirements for teacher credentials in the private sector, while teachers in the public sector are required to have Diplomas in Basic Education obtained from approved colleges of education (Asare & Nti, 2014). Public-sector teachers are classified as civil servants and thus receive guaranteed remuneration levels and job security, while private-sector teachers are generally paid less with no job security (Osei, 2006).

Importantly, a large majority of private schools in Ghana charge low fees and cater to low-income families (known as low-fee private schools), and many struggle to cover their costs (Baum, Abdul-Hamid, & Wesley, 2018). Yet, parents often perceive them to be of better quality than public schools (Dixon & Tooley, 2012; Zuilkowski, Piper, Ong'ele, & Kiminza, 2018). In our sample, children who attend private schools come from families who are better educated and have more wealth (), with 72% of children in public schools living in households that were 'at risk' for poverty, compared to only 28% of those in private schools. Given the differences in the teacher workforce and children and families across public- and private-sector schools, examining whether there are differences in effects of professional development programs across the two sectors is critical.

1.2.2 | Household characteristics

A compensatory hypothesis posits that the most disadvantaged individuals have the most 'room to grow', and therefore stand to benefit most from environmental supports. Disadvantage is often defined by household socioeconomic status (SES), which includes both income and parental education (Bradley & Corwyn, 2002). The compensatory hypothesis has been supported in the early childhood literature in the United States (McCartney, Dearing, Taylor, & Bub, 2007; Weiland & Yoshikawa, 2013) and is the basis for many publicly funded programs that aim to 'level the playing field' for low-income children (e.g. Head Start).

Alternatively, it is possible that within a disadvantaged population, more advantaged children will benefit from improved access to quality education. For example, a study in Bangladesh found that a stipend program aimed to reduce achievement gaps between relatively low- and high-SES families showed improved achievement only for higher-SES females (Behrman, 2015). Given these divergent findings, we examine moderation by SES (measured as household wealth and literacy status of the household head) without a priori hypotheses about the direction of moderating effects.

1.2.3 | Child characteristics

Similar to the compensatory hypothesis, some evidence suggests that the benefits of quality ECE may be greatest for children with low initial skills (Zaslow et al., 2016). Hamre and Pianta (2005) found that increments in classroom emotional support were only beneficial for children facing initially high 'functional risk', as defined by low attention and social skills, and/or high behavioral problems. But McCoy and Wolf (2018) found the opposite in Ghana. Specifically,

children with higher school readiness skills at school entry had larger gains in learning in high-quality classrooms than those with lower skills.

Gender is another important characteristic related to children's schooling outcomes. Ghanaian girls have historically experienced lower educational outcomes than boys (UNESCO, 2014), and gender parity in school enrollment in Ghana declines with school progression (UNGEI, 2012). Few studies have examined gender disparities in educational outcomes in ECE settings in Ghana. Finally, differences in developmental processes may depend on grade level (a proxy for both child age and schooling experience). Children's foundational academic, social, and executive function skills develop rapidly during the preschool period, reflecting both increasing environmental demands for these skills (e.g. in the context of preschool classrooms), and children's increasing neurodevelopmental capacity for higher-order thinking (Shonkoff & Phillips, 2000; Zelazo & Carlson, 2012). In this study, we examine whether baseline skills, gender, and grade level moderate impacts of children's school readiness outcomes in a relatively under-studied context.

1.3 | The Quality Preschool for Ghana (QP4G) Program

In 2004, the Government of Ghana adopted the National Early Childhood Care and Development Policy, which highlighted access to quality kindergarten education as central to improving child development and 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 primary grade one. Ghanaian children are exposed to pre-primary education at a much higher rate than in most peer countries, with 75% of 4- and 5-year-olds enrolled in pre-primary education (Ministry of Education, 2016). But several reports have concluded that classroom quality and learning outcomes are low among (e.g. Ghana Ministry of Education, 2014), including in kindergarten (Ghana Education Service, 2012). A 2012 Government Kindergarten Situational Report concluded that the curriculum established in 2004 was sound, but that teachers had not incorporated it into their practice (Ibid.).

The Quality Preschool for Ghana (QP4G) project aimed to train teachers on the KG curriculum (Republic of Ghana, 2004) and to enhance the quality of KG education. The goal was to develop and rigorously evaluate a scalable model of transformational teacher training to provide high-quality ECE to children, and to test the benefits of engaging parents via an awareness campaign designed to align parental expectations with the KG curriculum and pedagogy. 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 2 years of pre-primary education in Ghana called *kindergarten*.

1.3.1 | Targeting teachers

Ghanaian teachers, like teachers in many LMICs, face many challenges including increasing workloads due to educational reform, low and unreliable teacher remuneration, lack of professional recognition, large class sizes, and minimal professional development opportunities (Bennell & Akyeampong, 2007; Schwartz et al., 2019; 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). Teachers with high burnout and poor well-being have lower quality classrooms (Klusmann, Kunter, Trautwein, Lüdtke, & Baumert, 2008) and students with lower motivation and poorer achievement (Hindman & Bustamente, 2019; Jennings, 2016; McLean & Connor, 2015). Meta-analyses show that professional development programs that include ongoing mentoring and coaching can successfully change teachers' pedagogical practices in both high-income countries (Kraft, Blazar, & Hogan, 2018) and LMICs (Ganimian & Murnane, 2016), and can improve teacher well-being (Jennings, 2016).

The *teacher training and coaching program* (TT) of QP4G included training workshops (5 days in September, 2 days in January, and 1 day in May) and in-classroom coaching (six visits over 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 (NNTTC) 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: (a) how children learn – developing child-friendly environments, (b) classroom management, (c) incorporating child-centered and activity-based approaches to teaching language and literacy, (d) incorporating child-centered and activity-based approaches to teaching math, and (e) assessment and planning. The first half of each training day consisted of lectures and discussions. The second half focused on practicing the techniques covered in the first half and creating teaching and learning materials to implement activities in the classroom.

District-government ECE coordinators attended the teacher training and 2 days of additional training in their roles as monitors and coaches, plus a 1-day refresher training halfway through the school year. They were thus able to answer questions from teachers on the curriculum implementation. The coaching visits focused on practical ways teachers could integrate 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 1 hr, followed by debriefing sessions in which teachers reflected on their practices and were provided with feedback on what they did well, as well as areas for improvement.

1.3.2 | Targeting parents

Parental involvement in school may be 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 kindergarten children in the United States (Miedel & Reynolds, 2000). Levels of at-home cognitive stimulation in Ghana are relatively low, with only one-third of preschool-aged children having been read to in the three days immediately prior to data collection (vs. an average of 54.1% in all developing countries) (McCoy et al., 2016). At the same time, Ghanaian parents have been shown to value early education and demand academically focused, rigorous instruction from teachers (Bidwell et al., 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, 2019). Aligning parents' and schools' expectations for ECE and supporting parents to engage in their children's education may be critical for sustainably changing teacher practice and children's development.

As part of the QP4G program, three *parental-awareness meetings* were held through school parent–teacher associations (PTAs) over the school year. They were offered to all parents with kindergarten children in the school and administered by the same trained district-government ECE coordinators. Each meeting consisted of viewing videos developed for the intervention followed by discussions led by district coordinators. The video themes were (a) the importance of play-based learning, (b) parents' roles in children's learning, and (c) encouraging parent–teacher and parent–school communication. The meetings aimed to increase parent involvement with their children's education at home and in school.

1.3.3 | Theory of change

The *theory of change* was that teachers would use fewer rote repetition, memorization, and teacher-led strategies in favor of more activity-based, child-centered teaching strategies, and use more proactive positive behavior management practices. This would improve classroom quality (measured through teacher–child interactions) and teacher professional well-being (measured through teachers' motivation, burnout, and job satisfaction), and ultimately children's school readiness. For schools that received both TTPA, it was anticipated that this combined package would improve parents' understanding of play-based learning, strengthen parent–teacher communication, and increase parental involvement and engagement with teachers, and consequently have larger impacts on child outcomes.

1.3.4 | Findings of QP4G impacts over the program year

We undertook an initial study to test impacts on (a) teacher professional well-being, (b) classroom quality, and (c) children's school readiness and learning skills over one school year (Wolf et al., 2019). We summarize key 1-year impact results as background for the present study.

Teacher professional well-being and classroom quality

There were no impacts on motivation or job satisfaction, but there were reductions in teacher burnout in TT (d_{wt}[1] = −0.40) and TTPA (d_{wt} = −0.59). Additionally, TT impacted teacher turnover, reducing odds that teachers left KG classrooms by the third term by 58% (OR = 0.42, $p < 0.05$). QP4G also improved some elements of classroom quality, increasing the number of developmentally appropriate activities used in classrooms in both treatment conditions (d_{wt} = 0.56 in TT and 0.61 in TTPA). There were no impacts of either treatment condition on teachers' facilitation of deeper learning, but both treatment conditions increased levels of emotional support and positive behavior management observed in the classroom (d_{wt} = 0.65 in TT and 0.66 in TTPA), and TT (but not TTPA) increased levels of supporting student expression in classrooms (d_{wt} = 0.52).

Children's school readiness

We first assessed impacts on a composite score of children's school readiness, 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. TT increased children's school readiness (d_{wt} = 0.16). Impacts were statistically significant for three of the four domains: early numeracy (d_{wt} = 0.11), early literacy (d_{wt} = 0.11), and social–emotional skills (d_{wt} = 0.18). There were no TTPA impacts on children's school readiness. Compared to TT, children in TTPA had statistically lower scores on overall school readiness, and on domains of literacy and numeracy.

Contrary to our prediction, adding three parental-awareness meetings, administered through school PTAs by local government district coordinators, counteracted some of the positive impacts, specifically on children's school readiness outcomes. It is possible that parents disagreed with the messages from the training and favored traditional, teacher-directed, academically rigorous approaches (e.g. Bidwell et al., 2014). Also, it is possible that implementation challenges led parents to feel frustrated by the meetings and to push back on school change efforts. Furthermore, the study took place in peri-urban and semi-rural communities in the fastest growing and most diverse region in the country. Research in human development indicates that urbanization is a powerful force in shaping changing expectations for children's learning (Greenfield, 2009), and research with parents in this region of Ghana suggests that parents view preschool as a way to prepare children for academic learning and socialization (Kabay et al., 2017). The messages relayed in the QP4G program may have been interpreted by parents as threatening their goals for their children's academic preparation and socialization.

1.4 | The current study

We collected 1-year follow-up longitudinal data for the QP4G children. The current study goes beyond the initial QP4G study summarized above by addressing two primary questions: (1) Did the

short-term (project year) impacts on children's school readiness persist 1 year after children were exposed to QP4G? And (2) were there differential sustained impacts by children's gender, grade levels, baseline skills, household wealth, male household head literacy level, and attendance at private versus public schools?

2 | METHODS

The implementation and first-year evaluation of the QP4G intervention occurred between September 2015 and June 2016. Schools were randomly assigned to one of three treatment arms noted above: (a) TT, 82 schools; (b) TTPA, 79 schools; and (c) control group, 79 schools. The trial was registered in the American Economic Association 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 the initial study were collected in September–October 2015 (baseline) and May–June 2016 (follow-up 1). The data for the present study were collected in May–June 2017 (follow-up 2).

Randomization was stratified by district and sector to TT, TTPA, or control. The six most disadvantaged of the 16 districts in the Greater Accra region were selected. The selected districts were rated as the most disadvantaged districts in the region according to 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) that were within a 2-hr drive from Accra (for teachers to be able to attend the training in Accra). The Greater Accra Region is also the most developed part of Ghana, has the smallest proportion of socioeconomically disadvantaged citizens of all the regions, and is rife with ethnic diversity (Owusu & Agyei-Mensah, 2011). According to the 2012 Population and Housing Census, the Akans are the predominant ethnic groups in Ghana (47.5%), followed by the Mole–Dagbani (16.6%), the Ewe (13.9%), the Ga-Dangme (7.4%), and other groups (14.6%). In the Greater Accra Region, approximately 39.7% are Akan, 27.4% are Ga-Dangme, 20.1% are Ewe, 5.2% are Mole–Dagbani, and 7.6% are other ethnic groups (Ghana Statistical Service, 2013). Importantly, there is tremendous variation in socioeconomic status across districts and neighborhoods. For example, of the six districts included in this study, the average ranking on 'disadvantage' ranged from 93 to 187 (average of 139) out of 216 districts in the country (UNICEF, 2015). The districts were Ga South, Adenta, Ledzokuku-Krowor, Ga Central, La Nkwantanang-Madina, and Ga West. Appendix Table 4 provides descriptive statistics about the schools, teachers, and children and their households at baseline by treatment condition.

2.1 | Power analysis

The sample included 160 schools (for two-way comparisons) with seven children per class and two teachers–classrooms per school. With 80% power at the 5% significance level, and assuming an ICC of 0.15 for child outcomes, this was sufficient for minimum detectable effect sizes (MDES) of 0.17 for child outcomes. For subgroup effects, assuming the same number of schools and teachers–classrooms but half of the children per classroom, the MDES was 0.21. For subgroups at the school level (i.e. public vs. private), the MDES was 0.27.

2.2 | Sampling and data collection procedures

2.2.1 | School sample

All schools in the six districts were identified using the Ghana Education Service Educational Management Information System (GES-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 all districts.

All KG teachers in the schools were invited to participate in the training. The majority of schools had two KG teachers, though 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/ classrooms.

2.2.2 | Child sample

Class rosters for KG classrooms were collected. An average of 15 children (eight from KG1 and seven from KG2) were randomly selected from each school roster to participate in direct assessments. If a school had fewer than 15 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 replaced that child. For schools with only one KG classroom, 15 children were randomly sampled from the classroom. At baseline, the total sample of children was 3,435 children, with an average of 14.3 children per school (range = 4–15).

2.2.3 | Caregiver sample

Phone numbers for caregivers of children sampled were collected from school administrative records. But a number of these phone numbers were no longer correct when we tried to call the caregivers, so only

2,149 caregivers were reached, 12 of whom declined to participate (response rate of 99.1%), resulting in 2,137 caregivers for the original 3,435 children (64%). Children whose caregivers were reached were different from those not reached: more likely to be enrolled in private schools (56.0% vs. 49.2%, $t = 3.89$, $p < 0.01$) and had slightly higher literacy scores (45.5% vs. 43.7%, $t = 2.34$, $p < 0.05$). The groups were not statistically different on other child outcomes or locales of the schools.

2.2.4 | Data collection procedures

Following verbal assent, children's school readiness skills were assessed directly in their schools. When data collectors arrived at schools, they worked with head teachers to designate a few quiet spaces on the school grounds that were out of sight of other children where assessments could be conducted (e.g. an empty classroom). Primary caregivers of these children (41.6% mothers, 44.6% fathers, and 13.8% other) were then contacted via telephone to participate in a survey in which data on household SES and parental investments were collected. A pre-interview was conducted to determine whether the caregiver was the child's primary caregiver, defined as '… the person who takes primary responsibility for the child's education and who could best talk about the child and his/her experiences in school and at home. It may be the child's parent, a family member, guardian, or another individual'.

2.2.5 | Assessment development and adaptation

Extensive work was done to ensure that all measures were contextually appropriate. For the *caregiver surveys*, items were selected from existing scales and pilot-tested. First, we conducted five cognitive interviews with caregivers to assess whether they understood each question, both consistently across constructs and in the way the item was intended (Collins, 2003). Next, the survey was piloted with 20 caregivers, and distributions for all items were assessed. Suitable items were then selected for use.

The child assessment tool was translated into three local languages (Twi, Ewe, and Ga). Surveys were translated and then back-translated by different persons to check for accuracy. Any discrepancies were discussed and addressed. 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. While the tool used was designed to be implemented widely across culturally diverse settings, the current protocol was piloted with 20 children to confirm its appropriateness for the study context. Only minor adaptations were required (e.g. removing the word 'please' from the start of several questions and simplifying instructions by removing excessive words).

The majority of caregivers (63%) reported speaking more than one language at home. Of all of the languages reported spoken at home, 46.1% reported speaking English as one of those languages (a second language for nearly all of those); 34.1% reported any use of Twi/Fante (18.6% exclusively); 5.5% reported any use of Ewe (4.2% exclusively); 7.0% reported any use of Ga (6.3% exclusively); 1.9% reported any use of Dangme (0.2% exclusively); 2.6% reported any use of Hausa (1.2% exclusively), and 4.4% reported any use of another language (4.2% exclusively). Assessors spent several minutes chatting and playing games with children to make them comfortable before beginning the assessment. As schools in this sample reported using a mixture of English and local language for instruction, part of this initial introduction was intended to help the assessor to gauge children's linguistic preferences. Assessors then administered the assessment in the language he/she deemed most appropriate for the child, including Twi/Fante only (39.0%), Ewe only (1.3%), Ga only (5.0%), English only (37.9%), and mixed English and local language (16.9%).

3 | MEASURES

Table 1 displays the bivariate correlations among all study variables.

3.1 | Child development outcomes

Child outcomes were directly assessed in four domains: early literacy, early numeracy, social–emotional skills, and executive function using the International Development and Early Learning Assessment (IDELA) (Pisani, Borisova, & Dowd, 2018). The initial set of items was inspired by and conceptually adapted from existing assessments of children's school readiness. See Pisani et al. (2018) for details on the development of the IDELA items. Recent studies have assessed the psychometric properties and factor structure of the IDELA (Wolf et al., 2017), as well as the measurement invariance of the factors across five countries (Halpin et al., 2019).

Early literacy consisted of 38 items grouped into six constructs: 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 (e.g. 'Here is my friend mouse. Mouse starts with/m/. What other word starts with/m/? Cow, doll, milk') ($\alpha = 0.74$ and 0.88 at baseline and follow-up 2, respectively).

Early numeracy consisted of 39 items grouped into eight constructs: number knowledge, basic addition and subtraction, one-to-one correspondence, shape identification, sorting abilities based on color and shape, size and length differentiation, and completion of a simple puzzle. An example item assessing shape identification showed the child a picture with six shapes and asked the child to identify the circle ($\alpha = 0.72$ at baseline and follow-up 2).

Social–emotional skills consisted of 14 items grouped into five constructs: self-awareness, emotion identification, perspective taking and empathy, friendship, and conflict and problem solving. An example item of conflict solving involved asking children to imagine

TABLE 1 Bivariate correlations among key study variables

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	School readiness composite	1.00																		
2	TT	**0.05**	1.00																	
3	TTPA	**−0.02**	**−0.51**	1.00																
4	Male household head not literate	**−0.05**	−0.01	0.03	1.00															
5	Male household head literate	**0.06**	0.03	−0.02	**−0.65**	1.00														
6	No male head in household	**−0.02**	−0.03	0.00	**−0.19**	**−0.62**	1.00													
7	Household wealth	**0.22**	**0.07**	−0.04	**−0.35**	**0.41**	**−0.17**	1.00												
8	KG1 class at baseline	**−0.34**	−0.01	0.03	0.01	−0.03	0.02	0.00	1.00											
9	KG2 class at baseline	**0.33**	0.00	0.04	0.00	0.01	−0.01	−0.01	**−0.81**	1.00										
10	Mixed KG1/ KG2 class at baseline	0.01	0.02	**−0.11**	−0.02	0.03	−0.02	0.03	**−0.32**	**−0.30**	1.00									
11	Baseline school readiness score	**0.70**	0.00	0.03	−0.03	0.03	0.00	**0.13**	**−0.42**	**0.43**	−0.02	1.00								
12	Private school	**0.21**	0.03	0.00	**−0.16**	**0.22**	**−0.12**	**0.49**	**−0.10**	**−0.04**	**0.22**	**0.11**	1.00							
13	Child is male	**−0.04**	0.01	0.00	0.04	0.01	**−0.06**	−0.02	0.02	−0.01	−0.01	−0.02	−0.02	1.00						
14	Child age at baseline	**0.23**	−0.02	0.03	**0.13**	**−0.15**	**0.06**	**−0.25**	**−0.32**	**0.34**	−0.02	**0.34**	**−0.26**	0.00	1.00					
15	Ga East district	**−0.04**	0.01	−0.03	**−0.05**	**0.06**	−0.02	**0.05**	−0.03	**0.06**	**−0.04**	**−0.04**	−0.01	0.02	**−0.05**	1.00				
16	Ga South district	**−0.04**	**−0.04**	**0.04**	**0.09**	**−0.11**	**0.05**	**−0.22**	0.01	−0.02	0.01	**−0.09**	**−0.19**	0.02	**0.10**	**−0.25**	1.00			
17	Adenta district	**0.09**	**0.05**	−0.01	−0.02	**0.06**	**−0.05**	**0.09**	0.02	**0.04**	**−0.09**	**0.07**	**0.11**	**−0.07**	**−0.04**	**−0.17**	**−0.23**	1.00		
18	Ledzokuku district	**−0.04**	0.01	0.01	−0.01	0.01	0.00	0.03	0.00	**−0.03**	**0.05**	**−0.04**	**0.06**	0.01	−0.03	−0.23	**−0.31**	**−0.21**	1.00	
19	Ga Central district	**0.13**	0.01	−0.02	−0.02	0.01	0.01	**0.11**	−0.03	**−0.06**	**0.14**	**0.14**	**0.16**	0.00	0.02	**−0.17**	**−0.22**	**−0.15**	**−0.21**	1.00
20	Madina district	**−0.07**	**−0.05**	−0.01	0.00	−0.01	0.01	−0.02	**0.03**	0.02	**−0.08**	0.00	**−0.10**	0.00	−0.02	**−0.15**	**−0.20**	**−0.14**	**−0.19**	**−0.13**

Note: Bold correlations indicate statistical significance at $p < 0.05$.

they are playing with a toy and another child wants to play with the same toy and asking the children what they would do to resolve that conflict. 'Correct' answers in the Ghanaian context as agreed upon by the assessors during training included talking to the child, taking turns, sharing, and getting another toy (α = 0.69 and 0.67 at baseline and follow-up 2, respectively).

Executive function was assessed with ten items grouped into working memory (i.e. forward-digit span) and impulse control (i.e. head–toes task, adapted from McClelland et al. (2014) as described in Pisani, Borisova, and Dowd (2015)). 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 assessors touched their heads, and vice versa, in a series of five items (α = 0.83 and 0.79 at baseline and follow-up 2, respectively).

A *school readiness composite* was created by summing with equal weight scores on the four domains measured via direct assessment.

3.1.1 | Reliability

Interrater reliabilities on the child outcome measures were 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 ranged from 0.82 to 0.90, with an average of 0.87.

3.2 | Household characteristics

All household characteristics were collected directly from caregivers via telephone surveys in October 2015.

Household wealth was measured using the Simple Poverty Scorecard for Ghana (Schreiner & Woller, 2010). This tool was developed using the national 2005–6 Ghana Living Standards Survey, administered by the Ghana Statistical Service, to construct a scorecard that estimates the likelihood that a household falls below the poverty line using ten indicators: number of household members; number of school-aged children enrolled in school; highest grade completed by female head or spouse; primary job of male head; construction material used for house roof; main lighting source for dwelling space; main source of household drinking water; presence of working stove; possession of working iron; and possession of working radio, radio cassette, record player, or 3-in-1 radio system. Scores on each response were assigned numerical values with total scores ranging from 1 to 100, where higher scores indicate greater household wealth.

Male household head literacy was measured with the question used by the Ghana Statistical Service (Schreiner & Woller, 2010) asked of the caregiver: 'Can the male head/spouse read a phrase/sentence in English?' Responses were categorized as 0 = 'No' (15.5%), 1 = 'Yes' (64.5%), and 2 = 'No male head/spouse' (15.0%).

3.3 | Covariates

We included covariates to improve the precision of our impact estimates: private-sector status of the school, six district dummies, and five dummy variables accounting for within-sample mobility between baseline and follow-up 1 (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). In addition, we included child gender (1 = male), 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.[2]

3.4 | Analytic plan

3.4.1 | Baseline equivalence and attrition analysis

Baseline equivalency analyses confirmed that randomization was successful in yielding statistically equivalent treatment and control groups. We next undertook differential attrition analysis for follow-up 2. Six hundred and seventy-nine children from baseline (19.8%) were not in the follow-up 2 sample. Treatment status did not significantly predict whether children left the study sample at either follow-up, indicating that our experimental design was not compromised. To assess external validity of the sample of children that stayed, we assessed baseline levels of school readiness, child gender, child age, and private-sector status. Higher school readiness predicted a lower likelihood ($b = -0.70$, $SE = 0.06$, $p < 0.05$), and child age predicted a higher likelihood of leaving the sample ($b = 0.14$, $SE = 0.04$, $p < 0.001$) by follow-up 2.

3.4.2 | Missing data imputation

We used multiple imputation (with Stata's 'ice' command) to handle missing data on all missing variables, following the What Works Clearinghouse Version 4.0 Standards Handbook (Institute for Education Science (IES) 2017), including dependent variables and treatment status, using three rounds of data collection (baseline and both follow-ups). While the data were not missing completely at random (MCAR), if variables that strongly predict attrition are incorporated into the missing data strategy, the plausibility of the missing-at-random (MAR) assumption increases (Young & Johnson, 2015). In other words, 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. Impact estimates were computed using 100 imputed datasets.[3]

3.4.3 | Impact analysis

To account for the nested, non-independent nature of the data (i.e. students nested within classrooms that are nested within schools), we employed three-level models to estimate impacts on child

outcomes. Impact analyses were conducted with a select set of covariates. We nested children in the baseline teachers and schools from which they were sampled. The multiply imputed datasets 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 datasets (Rubin, 1987). Notably, 56.8% of the variation was explained at the child level, 43.2% at the classroom level, and 0% at the school level. Nonetheless, given the school-randomized study design, we modeled all three levels in our estimation models.

The primary outcome was school readiness (i.e. composite IDELA score) at follow-up 2. As a post hoc test, we estimated impacts on each of the four domains of school readiness (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:

$$Y_{ijk} = \boldsymbol{B}_{0jk} + \boldsymbol{B}'_{1jk}\boldsymbol{X}_{ijk} + e_{ijk}$$

where $\boldsymbol{X}_{ijk}$ is the vector of child covariates (gender, age, and baseline score).

Level 2 (Classroom-level) Model:

$$\boldsymbol{B}_{0jk} = \gamma 00k + u0jk$$

where $\boldsymbol{B}_{0jk}$ is the classroom-level random intercept.

Level 3 (School-level) Model:

$$_{00k} = \pi_{000} + \pi_{001}\mathbf{TT}_k + \pi_{002}\mathbf{TTPA}_k + \pi'_{003}\boldsymbol{Z}_k + v_{00k}$$

where γ_{00k} is the school-level random intercept; $\boldsymbol{Z}_k$ is the vector of school-level covariates (district dummies, private or public, and four dummy variables for different school-mobility scenarios); and $\mathbf{TT}_k$ and $\mathbf{TTPA}_k$ are indicators for schools assigned to the two treatment arms.

Next, we examined whether intervention-sustained impacts were moderated by child characteristics (gender, child baseline scores, grade level [KG1 and KG2]), by household characteristics (household wealth, male household head presence, and literacy), and by private versus public school. Moderation of impacts was tested by adding cross-level interaction terms between each treatment condition (school-level) and child/household variable (child-level). Moderation by sector was calculated with an interaction term between school-sector status (1 = private, 0 = public) and treatment status.

4 | RESULTS

4.1 | Persistent impacts on school readiness

Table 2 presents impact estimates of QP4G on children's school readiness 1 year after the end of the one school-year program, with each row representing the average effect for each treatment condition relative to the control group. We first assessed impacts on our primary outcome of interest: the composite score of children's school readiness. Then, we conducted post-hoc analyses to assess impacts on each domain of school readiness individually to examine whether findings were driven by any particular domain. There were marginally statistically significant persistent impacts of TT on children's overall school readiness (b = 0.015, SE = 0.008, p = 0.065, d_{wt} = 0.13). By domain, there were statistically significant persistent impacts on children's social–emotional development (d_{wt} = 0.13, $p < 0.05$), with marginally statistically significant impacts on executive function (d_{wt} = 0.11, $p < 0.10$). Similar to the first follow-up impact results (Wolf et al., 2019), TTPA did not impact children's school readiness.

Compared to the TT children, the TTPA children had statistically significantly lower scores on overall school readiness (b = −0.022, SE = 0.009, $p < 0.05$). Post hoc analyses showed that the TTPA children had lower scores than the TT children across three domains: marginally statistically significant lower scores on literacy (b = −0.025, SE = 0.014, $p < 0.08$) and social–emotional development (b = −0.022, SE = 0.013, $p < 0.08$), and significant differences in executive function (b = −0.029, SE = 0.012, $p < 0.05$).

4.2 | Impact variation by child characteristics

We next assessed whether impacts on school readiness were moderated by three child characteristics: gender, baseline school readiness, and grade level (KG1, KG2) (Table 3). We found no significant interactions between treatment status and these child characteristics at follow-up 2 (see Panel A, Table 3).

4.3 | Impact variation by household characteristics

We assessed whether impacts on school readiness were moderated by key household characteristics: household wealth, male household head present, and, if present, male household head literacy status (Panel B, Table 3). We found no statistically significant interactions between treatment status and household wealth.

We did find significant interaction effects by baseline male household head's presence and literacy status for TTPA. Specifically, there was a significant interaction between TTPA and children living in households with a male head compared to no male head (b = −0.032, SE = 0.015, $p < 0.05$), as well as whether the male head was literate (b = 0.023, SE = 0.011, $p < 0.05$). Figure 1 provides a graphical depiction of these interaction effects, showing that TTPA had counteracting impacts on children's school readiness for children in households with a non-literate male head. In this case, TTPA had small *negative* impacts on children's school readiness compared to control children living with a non-literate male head.

4.4 | Impact variation by school characteristics

We assessed whether impacts on school readiness were moderated by whether children attended a public- versus private-sector school

(Panel C, Table 3). We found no statistically significant interactions between treatment status and school sector.

5 | DISCUSSION

This study posed and answered two critical questions about influences on early childhood development that have not been examined before in SSA. First, were the short-term impacts of a preschool teacher training program, delivered with and without accompanying parental-awareness meetings, on children's school readiness sustained 1 year after the end of the year in which programs were implemented? To the best of our knowledge, this study constitutes the first test of persistence or fade-out of a preschool teacher professional development program in SSA. During the intervention year, the program increased activity-based learning and positive behavior management classroom practices during the intervention year, and improved teacher burnout, thus improving some of the hypothesized mediating pathways through which the program would improve child development.

Second, were there different patterns of persistence or fade-out of impacts for children who vary in personal (e.g. gender, grade, baseline school readiness), household (e.g. household wealth, literacy status of male household head), and school (public vs. private) characteristics? We investigated whether these characteristics moderated the persistence or fade-out of treated impacts on children's school readiness, as well as whether persistent effects differed in public versus private schools. To address these questions, we followed a large sample of students in the two grades of pre-primary school (referred to in Ghana as KG1 and KG2) from 240 public and private preschools in the Greater Accra region of Ghana after they transitioned to KG2 or Primary 1 classrooms.

The issue of persistence or fade-out of preschool interventions on child outcomes is an important topic in both research and policy circles in the United States (Bailey et al., 2017) and other OECD countries. From a research perspective, such studies offer opportunities to examine how features of children's development may be malleable to intentional efforts to improve their environments. From a policy perspective, governments and communities wish to make investments in interventions that lead to lasting improvements in children's development and so lead to higher returns on their investments.

We first examined impacts on the composite measure of children's school readiness (to control for experiment-wise error) and then by the four subdomains of school readiness (early literacy, early numeracy, social–emotional skills, and executive function skills) for a finer-grained post hoc analysis of impact persistence or fade-out. Our major findings are in Table 2. There were marginally statistically significant sustained impacts of the teacher training intervention (TT) on children's composite school readiness a year after the completion of QP4G. We can interpret these findings in the context of the results of a meta-analysis of cognitive impacts of 67 ECE programs in the United States (Bailey et al., 2017), which report a decline in impacts from an average ES of 0.25 immediately after treatment to ESs 0.09–0.10 1–2 years after treatment. In this study in Ghana, we observed that the initial small impacts from year 1 persisted 1 year later.

TABLE 2 Treatment impacts on child outcomes 1 year after program implementation

	b (SE)	*p*-value	Effect size (d_{wt})
Composite school readiness			
TT versus control	0.015 (0.008)	0.065	0.126^{+}
TTPA versus Control	−0.007 (0.007)	0.345	−0.059
Early numeracy			
TT versus control	0.005 (0.009)	0.591	0.032
TTPA versus Control	−0.008 (0.009)	0.402	−0.051
Early literacy			
TT versus control	0.015 (0.012)	0.264	0.084
TTPA versus Control	−0.009 (0.012)	0.451	−0.050
Social–emotional			
TT versus control	0.022 (0.011)	0.045	0.131*
TTPA versus Control	−0.002 (0.011)	0.982	0.012
Executive function			
TT versus control	0.018 (0.011)	0.099	0.112^{+}
TTPA versus control	−0.011 (0.011)	0.324	−0.069

Note: Sample size for TT versus control = 2,268 children nested in 296 teachers nested in 161 schools. Sample size for TTPA versus control = 2,255 children nested in 291 teachers nested in 158 schools. All impact estimates computed from 100 multiply imputed datasets. TT, teacher training condition; TTPA, teacher training plus parental-awareness training condition.
$^{+}p < 0.10$;
$^{*}p < 0.05$.

A more nuanced picture emerges when we consider impacts on the four subdomains that make up the composite school readiness scale. At the end of the intervention year, TT (but not TTPA) had positive impacts on three of the four domains: early literacy ($d = 0.11$, $p < 0.05$), early numeracy ($d = 0.11$, $p < 0.05$), and social–emotional skills ($d = 0.18$, $p < 0.05$). One year later, there were no positive impacts on early literacy or numeracy, but there were sustained positive impacts on social–emotional skills ($d = 0.13$, $p < 0.05$) and new marginally significant impacts on executive function skills ($d = 0.11$, $p < 0.10$). In short, the impacts on pre-academic skills faded out, but the impacts on 'soft' skills persisted (social–emotional skills) or emerged (executive function).

Bailey et al. (2017) discuss key features of child and adolescent interventions, as well as personal and environmental characteristics, that may help explain why some impacts persist after interventions end and some impacts fade out. They argue that interventions that target (directly or indirectly) skills in children that are malleable, fundamental, and that would not develop over time in the absence of the intervention are more likely to yield persistent impacts. They characterize (a) literacy and math skills and (b) social–emotional

TABLE 3 Impact variation on school readiness by child, household, and school characteristics

	b (*SE*)	*p*-value
Panel A: child characteristics		
Baseline school readiness		
TT versus control	0.017 (0.015)	0.263
TTPA versus control	−0.002 (0.014)	0.870
Baseline score	0.503 (0.020)	0.000***
TT × Baseline score	−0.004 (0.027)	0.886
TTPA × Baseline score	−0.011 (0.027)	0.667
Gender		
TT versus control	0.017 (0.009)	0.052+
TTPA versus control	−0.004 (0.009)	0.622
Male	0.000 (0.006)	0.956
TT × male	−0.005 (0.008)	0.529
TTPA × male	−0.006 (0.008)	0.435
Cohort during implementation year		
TT versus control	0.020 (0.010)	0.036*
TTPA versus control	0.003 (0.010)	0.795
KG2 (vs. KG1)	0.027 (0.008)	0.001***
TT × KG2	−0.010 (0.010)	0.384
TTPA × KG2	−0.011 (0.010)	0.305
Panel B: household characteristics		
Household wealth		
TT versus control	0.001 (0.019)	0.949
TTPA versus control	0.005 (0.019)	0.806
Wealth score (1–100)	0.000 (0.000)	0.286
TT × Wealth score	0.000 (0.000)	0.467
TTPA × Wealth score	0.000 (0.000)	0.480
Parent literacy		
TT versus control	0.015 (0.013)	0.250
TTPA versus control	0.004 (0.013)	0.753
Male head in household (1 = yes)	0.021 (0.012)	0.085+
Male head is literate (1 = yes)	−0.015 (0.009)	0.088+
TT × Male head in household	−0.012 (0.015)	0.397
TTPA × Male head in household	0.032 (0.015)	0.031*
TT × Male head is literate	0.015 (0.025)	0.196
TTPA × Malehead is literate	0.023 (0.011)	0.046*
Panel C: school characteristics		
Private (vs. public) school		
TT versus control	0.009 (0.011)	0.385
TTPA versus control	−0.000 (0.011)	0.997
Private school (vs. public)	0.035 (0.010)	0.000***
TT × private school	0.009 (0.013)	0.486
TTPA × private school	−0.016 (0.013)	0.235

Note: Estimates are computed using observed scores, in three-level models: children nested in classrooms nested in schools. Effect sizes calculated accounting for the 3-level model structure (Hedges, 2009). Sample includes children present at baseline and follow-up. Sample size for TT versus control = 2,268 children nested in 296 teachers nested in 161 schools. Sample size for TTPA versus control = 2,255 children nested in 291 teachers nested in 158 schools. All impact estimates computed from 100 multiply imputed datasets. Models include the following control variables: private (vs. public)-sector status of the school, six district dummies, a dummy variable for if the school was assigned to receive teacher text messages, a dummy for if the school was assigned to receive parent flyers, a series of five dummy variables accounting for within-sample mobility in year 1, 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. TT, teacher training condition; TTPA, teacher training plus parental-awareness training condition.

+$p < 0.10$;
*$p < 0.05$;
***$p < 0.001$.

and executive function skills as both malleable and fundamental. But they specifically claim rudimentary reading and math skills are among those that grow rapidly in most counterfactual conditions. In contrast, they identify other skills that may not 'fade out', especially in more adverse environments, including some normative cognitive functions (like executive functions) and emotion regulation (part of the social–emotional skill domain). In short, Bailey et al. (2017) argue that skills that meet all three criteria (malleable, fundamental, would not develop normally in the absence of the intervention) are required if intervention impacts are to persist. Using this framework, we suspect that rudimentary literacy and math scores among preschoolers in Ghana meet two of these three criteria, and thus, children are more likely to catch up with these skills. And perhaps social–emotional skills and executive function skills meet all three criteria and represent what Bailey and colleagues call 'trifecta skills', at least in the Ghanaian preschool context.

The 1-year impact study found that adding three brief parental-awareness meetings, administered through school PTAs by local government district coordinators, to the teacher training (TTPA) actually *counteracted* the positive gains children made as a result of the teacher training (Wolf et al., 2019). Importantly, these meetings focused on increasing parental engagement and awareness of developmentally appropriate education in early childhood as opposed to parenting skills, unlike most evidence-based parent interventions that focus on parenting practices for younger children aged 0–3 (e.g. Yousafzai, Rasheed, Rizvhi, Armstrong, & Bhutta, 2014; Özler et al., 2018). Research suggests that peri-urban Ghanaian parents value preschool and view it as a way to prepare children for academic learning and socialization (Kabay et al., 2017). Follow-up interviews with teachers in the TTPA condition indicate that some parents did not agree with the messages of the teacher training program (to promote child-centered and play-based learning) and so attempted to

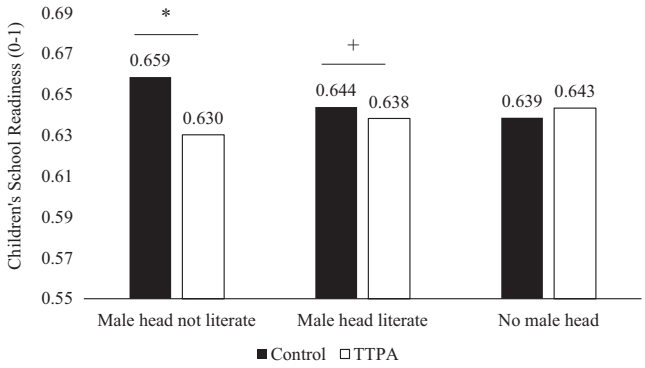

FIGURE 1 Impacts of TTPA on children's school readiness by male household head literacy status. $^{*}p < 0.05$, $^{+}p < 0.10$

counter the changes in teachers' practices at home out of concern that their children would not learn (both academically and socially) if such methods were used. For example, 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'. These experiences indicate some parents did not agree with the messages and perhaps advocated for teachers to continue to use the old methods. It also possible that certain parents compensated at home for the changes occurring in school by increasing some strategies to socialize and support their children's development in ways that were counterproductive.

The impact variation found in persistent negative effects may help explain these results. Specifically, the counteracting effect of the parental-awareness meetings persisted 1 year later, but this appears to be the case primarily for children in households with a non-literate male head. Perhaps parents with less education were more likely to disagree with the messages conveyed in the PTA meetings, and male parents felt more empowered than females to attempt to counter at home the changes in teachers' practices, or to push back on teachers not to use these methods. Further research is needed to understand how to engage parents who are not literate in an effective way that promotes their children's development, while understanding their concerns and desires for their child and helping parents foster those desires in productive ways. Ghanaian parents value early educational opportunities for their child but are critical of play in preschool because they view it as unstructured time (Kabay et al., 2017). Providing parents with opportunities to observe and participate in play- and activity-based learning in their children's classrooms is emerging as one promising approach to changing their perspectives (Krutikova et al., 2019). In addition, the Greater Accra Region is one of the great ethnic and linguistic diversities given the rapid internal migration occurring in the region (Owusu & Agyei-Mensah, 2011). Further research is needed to examine whether and how ethnic backgrounds, coupled with urbanization, may have moderated the impacts of both treatment conditions on children and their caregivers.

We also considered moderation of impacts by select child, household, and school characteristics to examine sources of heterogeneity of impacts. Previous studies of ECE intervention impacts have often found that relatively more disadvantaged children, lower-functioning children (at baseline), and younger children are more responsive to interventions than are more advantaged, higher-functioning, and older children (Yoshikawa et al., 2013). In this study, we did not find that household wealth, baseline skills, grade level, gender, or whether children were in a public versus private school moderated treatment impacts. This is encouraging from an equity perspective – that is that inequalities did not increase among the targeted, relatively disadvantaged population. But it also means the program did not close equity gaps among children. Attention to who benefits from educational programs, as well as how to ensure that the most marginalized benefit, is critical areas for research to continue to consider.

5.1 | Strengths and limitations

This study has numerous strengths: a randomized experimental design with sufficient power to detect small effects, the use of culturally adapted measures collected by Ghanaian data collectors, longitudinal tracking of children for a year after the end of the one school-year intervention, and assessment of multiple sub domains of children's school readiness. But there also are important limitations that should be acknowledged. First, there was significant attrition of the children in the sample (about one-fifth of the baseline sample) and significant missing responses for about one-third of the caregivers due to a difficulty obtaining correct phone numbers. The use of multiple imputation and multiple controls probably limits any bias due to attrition. And notably, the pattern of results is very similar for the sample of children who are present both baseline and the second follow-up (see Appendix Table 5). Second, we were not able to collect systematic data on the types of classroom environments and teaching practices all children were exposed to, which would have allowed us to understand whether and how different educational environments support persistence of impacts. Third, the study's sample is limited to six peri-urban and semi-rural districts in the Greater Accra Region in Ghana, with implications for generalizability of the findings. Rural communities, which make up 45% of Ghana's population (World Bank, 2016) and include the most educationally disadvantaged children in the country (Ghana Ministry of Education, 2016), were not included in this study, and the findings are not necessarily generalizable to these children. Furthermore, the Greater Accra Region is the fastest growing region in Ghana, with significant diversity in terms of cultural and ethnic backgrounds (Owusu & Agyei-Mensah, 2011). Unfortunately, we did not collect data on the ethnic background of the families in the study, beyond the main language spoken at home. This restricted our ability to assess the role that ethnicity and culture may have played in how families responded to the interventions. Fourth, due to time and resource constraints, we collected very little 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.

5.2 | Implications for developmental science and global early childhood policy

As ECE continues to advance as a way to improve early childhood development in low- and middle-income countries (e.g. McCoy et al., 2018), concerns about quality have risen. Our findings suggest that a brief in-service teacher training, built into existing governmental systems and implemented over the course of one school year, can lead to some longer-term gains in children's early learning and development, particularly in developmental domains that may not otherwise be promoted by Ghanaian teachers. How to increase the size of these longer-term improvements is an important area for future work, in particular if ECE 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 is affordable and can feasibly be implemented at scale is one with which the field must continue to grapple.

This study suggests several future directions for policy-relevant research to explore, including what are the mechanisms of parents' roles in child development for preschoolers? how can parents' interests and activities be harnessed to be more complementary with improved teacher training? to what extent are the persistent effects that we find 1 year after program completion going to persist over longer time horizons? what characteristics of primary schools are likely to reinforce the gains in preschool?; and to what extent or with what modification are the effective aspects of teacher training in peri-urban Ghana transferable to other contexts? In our ongoing research, we are attempting to explore some of these questions. For example, we are in the field now assessing impacts after another year and collecting much more extensive information about the primary classroom environment in which children are learning to assess whether variation in future educational environments helps explain persistence or fade-out of ECE effects. With such ongoing research, we hope to contribute further to knowledge about what makes ECE most effective in contexts such as in Ghana.

ACKNOWLEDGEMENTS

This paper reflects contributions from many organizations and individuals. We would like to thank the dedicated staff and thought partners at Innovations for Poverty Action, including Loic Watine, Bridget Gyamfi, Amma Aboagye, Kwabena Asamonah, Edward Tsinigo, Henry Atimone, Maham Farhat, Shams Muzaffar, Renaud Comba, 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 provided intramural support to the first and second authors for research activities that directly contributed to this paper.

CONFLICT OF INTEREST

There are no conflicts of interest to report.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in 'World Bank Data Catalogue' at https://datacatalog.worldbank.org/dataset/ghana-quality-preschool-impact-evaluation-2017, reference number [GHA_2017_QPIE-EL_v01_M].

ENDNOTES

[1] d_{WT} represents a standardized mean difference between treatment and control schools accounting for nesting of children within classrooms and classrooms within schools, as calculated following Hedges (2009).

[2] Each model also included two dummy variables for if the school was randomly assigned to receive weekly teacher text messages or text messages and parent flyers. These were tested as ways to reinforce key messages from the intervention. No impacts of these reinforcements were detected on any outcome in the first follow-up (see Authors, 2018), and thus, we do not interpret these coefficients.

[3] We conducted the imputation in two steps. 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 datasets. All impact estimates on teacher and classroom-level data were computed on these 20 datasets (using Stata's "mi estimate" command). Second, we randomly selected ten of these teacher datasets. We merged each individual dataset with the child outcomes data, caregivers data, and basic children demographic characteristics from all waves of data. For each of the ten teacher datasets, we imputed ten child datasets, resulting in 100 child-level datasets.

REFERENCES

Ansari, A. (2018). The persistence of preschool effects from early childhood through adolescence. *Journal of Educational Psychology*, *110*, 952. https://doi.org/10.1037/edu0000255

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., & Schady, N. (2016). Teacher quality and learning outcomes in kindergarten. *Quarterly Journal of Economics*, *131*(13), 1415–1453. https://doi.org/10.1093/qje/qjw016

Asare, K., & Nti, S. K. (2014). Teacher education in Ghana: A contemporary synopsis and matters arising. *SAGE Open*, *4*(2), 1–8. https://doi.org/10.1177/2158244014529781

Bailey, D., Duncan, G. J., Odgers, C. L., & Yu, W. (2017). Persistence and fadeout in the impacts of child and adolescent interventions. *Journal of Research on Educational Effectiveness*, *10*(1), 7–39. https://doi.org/10.1080/19345747.2016.1232459

Barnett, W. S. (1995). Long-term effects of early childhood programs on cognitive and school outcomes. *The Future of Children*, 25–50. https://doi.org/10.2307/1602366

Barnett, W. S., Jung, K., Friedman-Krauss, A., Frede, E. C., Nores, M., Hustedt, J. T., … Daniel-Echols, M. (2018). State prekindergarten effects on early learning at kindergarten entry: An analysis of eight state programs. *AERA Open*, *4*, 2332858418766291. https://doi.org/10.1177/2332858418766291

Baum, D. R., Abdul-Hamid, H., & Wesley, H. T. (2018). Inequality of educational opportunity: The relationship between access, affordability, and quality of private schools in Lagos. *Nigeria. Oxford Review of Education*, *44*(4), 459–475.

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.

Berlinski, S., Galiani, S., & Gertler, P. (2009). The effect of pre-primary education on primary school performance. *Journal of Public Economics*, *93*(1–2), 219–234. https://doi.org/10.1016/j.jpubeco.2008.09.002

Bidwell, K., Parry, K., & Watine, L. (2014). *Exploring early education programs in peri-urban settings in Africa: Accra, Ghana*. New Haven, CT: Innovations for Poverty Action.

Black, M. M., Walker, S. P., Fernald, L. C., Andersen, C. T., DiGirolamo, A. M., Lu, C., … Devercelli, A. E. (2017). Early childhood development coming of age: Science through the life course. *The Lancet*, *389*(10064), 77–90.

Bornstein, M. H., & Putnick, D. L. (2012). Cognitive and socioemotional caregiving in developing countries. *Child Development*, *83*, 46–61. https://doi.org/10.1111/j.1467-8624.2011.01673.x

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, *53*(1), 371–399. https://doi.org/10.1146/annurev.psych.53.100901.135233

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), 1–31.

Campbell, F. A., Pungello, E. P., Burchinal, M., Kainz, K., Pan, Y. I., Wasik, B. H., … Ramey, C. T. (2012). Adult outcomes as a function of an early childhood educational program: An Abecedarian Project follow-up. *Developmental Psychology*, *48*(4), 1033. https://doi.org/10.1037/a0026644

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

Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from Head Start. *American Economic Journal: Applied Economics*, *1*(3), 111–134. https://doi.org/10.1257/app.1.3.111

Dixon, P., & Tooley, J. (2012). A case study of private schools in Kibera. An update . *Educational Management Administration & Leadership*, *40*(6), 690–706.

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. Available from: https://issuu.com/sabretom/docs/10_12_12_final_version_of_narrative_op. [last accessed 14 August 2016].

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

Ghana Statistical Service (2013). *2010 Population & housing census: National analytical report*. Ghana Statistics Service.

Gormley, W. T. Jr, Phillips, D., & Anderson, S. (2018). The effects of Tulsa's Pre-K program on middle school student performance. *Journal of Policy Analysis and Management*, *37*(1), 63–87. https://doi.org/10.1002/pam.22023

Greenfield, P. (2009). Linking social change and developmental change: Shifting pathways of human development. *Developmental Psychology*, *45*(2), 401–418. https://doi.org/10.1037/a0014726

Halpin, P. H., Wolf, S., Yoshikawa, H. Y., Rojas, N., Kabay, S., Pisani, L., & Dowd, A. J. (2019). Measuring early learning and development across cultures: Invariance of the IDELA across five countries. *Developmental Psychology*, *55*(1), 23–37. https://doi.org/10.1037/dev0000626

Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child Development*, *76*(5), 949–967. https://doi.org/10.1111/j.1467-8624.2005.00889.x

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P., & Yavitz, A.. (2010). *A new cost-benefit and rate of return analysis for the Perry Preschool Program: A summary* (No. w16180). National Bureau of Economic Research.

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: Russell Sage Foundation.

Hindman, A., & Bustamente, A. (2019). Teacher depression as a dynamic variable: Exploring the nature and predictors of change over the head start year. *Journal of Applied Developmental Psychology*, *61*, 43–55. https://doi.org/10.1016/j.appdev.2018.09.004

Institute for Education Science (IES) (2017). *WWC version 4.0 standards handbook*. Washington, D.C.: IES.

Jenkins, J. M., Watts, T. W., Magnuson, K., Gershoff, E. T., Clements, D. H., Sarama, J., & Duncan, G. J.. (2018). Do high-quality kindergarten and first-grade classrooms mitigate preschool fadeout? *Journal of Research on Educational Effectiveness*, *11*(3), 339–374. https://doi.org/10.1080/19345747.2018.1441347

Jennings, P. A. (2016). CARE for teachers: A mindfulness-based approach to promoting teachers' social and emotional competence and well-being. In K. A. Schonert-Reichl & R. A. Roeser (Eds.), *Handbook of mindfulness in education: Integrating theory and research into practice* (pp. 133–148). New York: Springer-Verlag Publishing.

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. https://doi.org/10.1016/j.ijedudev.2017.08.013

Klusmann, U., Kunter, M., Trautwein, U., Lüdtke, O., & Baumert, J. (2008). Teachers' occupational well-being and quality of instruction: The important role of self-regulatory patterns. *Journal of Educational Psychology*, *100*(3), 702–715.

Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, *88*(4), 547–588.

Krutikova, S., Leighton, P., Caeyers, B., Attanasio, O., Cattan, S., Sosa, L. C., & Bornstein, M. (2019). Strategies to improve pre-school quality: Experimental impacts of "Lively Minds" low-cost community-run playschemes in rural Ghana. Paper presented at *Society for Research in Child Development*. Baltimore, MD.

Lewis, L. (2013). *Is there a role for the private sector in education? Education for Global Development*. Available from: https://blogs.worldbank.org/education/europeandcentralasia/there-role-private-sector-education. [last accessed 29 September 2017].

McCartney, K., Dearing, E., Taylor, B. A., & Bub, K. L. (2007). Quality child care supports the achievement of low-income children: Direct and indirect pathways through caregiving and the home environment. *Journal of Applied Developmental Psychology*, *28*(5–6), 411–426. https://doi.org/10.1016/j.appdev.2007.06.010

McClelland, M. M., Cameron, C. E., Duncan, R., Bowles, R. P., Acock, A. C., Miao, A., & Pratt, M. E. (2014). Predictors of early growth in academic achievement: The head-toes-knees-shoulders task. *Frontiers in Psychology*, *5*, 599. https://doi.org/10.3389/fpsyg.2014.00599

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 Med*, *13*, e1002034. https://doi.org/10.1371/journal.pmed.1002034

McCoy, D. C., Salhib, C., Yoshikawa, H., Black, M., Britto, P., & Fink, G. (2018). Home- and center-based learning opportunities for preschoolers in low- and middle-income countries. *Child and Youth Services Review*, *88*, 44–56. https://doi.org/10.1016/j.childyouth.2018.02.021

McCoy, D. C., & Wolf, S. (2018). Changes in classroom quality predict Ghanaian preschoolers' gains in academic and social-emotional skills. *Developmental Psychology*, *54*(8), 1582–1599. https://doi.org/10.1037/dev0000546

McLean, L., & Connor, C. M. (2015). Depressive symptoms in third-grade teachers: Relations to classroom quality and student achievement. *Child Development*, *86*, 945–954. https://doi.org/10.1111/cdev.12344

Miedel, W. T., & Reynolds, A. J. (2000). Parent involvement in early intervention for disadvantaged children: Does it matter? *Journal of School Psychology*, *37*, 379–402.

Ministry of Education (2016). *Education sector performance report*. Ghana.

Morris, P. A., Connors, M., Friedman-Krauss, A., McCoy, D. C., Weiland, C., Feller, A., … Yoshikawa, H. (2018). New findings on impact variation from the Head Start Impact Study: Informing the scale-up of early childhood programs. *AERA Open*, *4*(2), 2332858418769287. https://doi.org/10.1177/2332858418769287

Neuman, M. J., Josephson, K., & Chua, P. G. (2015). *A review of the literature: Early childhood care and education (ECCE) personnel in low- and middle-income countries*. Early childhood care and education (Working Paper Series 4).

Osei, G. M. (2006). Teachers in Ghana: Issues of training, remuneration and effectiveness. *International Journal of Educational Development*, *26*, 38–51. https://doi.org/10.1016/j.ijedudev.2005.07.015

Owusu, G., & Agyei-Mensah, S. (2011). A comparative study of ethnic residential segregation in Ghana's two largest cities, Accra and Kumasi. *Population and Environment*, *32*(4), 332–352. https://doi.org/10.1007/s11111-010-0131-z

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. https://doi.org/10.1016/j.jdeveco.2018.04.004

Pesando, L. M., Wolf, S., Behrman, J. R., & Tsinigo, E. (manuscript under revision). Are private kindergartens really better? Examining school choice, parental resources, and children's school readiness in Ghana. *Comparative Education Review*.

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

Pisani, L., Borisova, I., & Dowd, A. J. (2018). Developing and validating the International Development and Early Learning Assessment (IDELA). *International Journal of Educational Research*, *91*, 1–15. https://doi.org/10.1016/j.ijer.2018.06.007

Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., …Downer, J. (2012). Third grade follow-up to the head start impact study: Final report. OPRE report 2012–45. *Administration for Children & Families*.

Puma, M., Bell, S., Cook, R., Heid, C., Shapiro, G., Broene, P., …Ciarico, J. (2010). Head start impact study. Final report. *Administration for Children & Families*.

Ramey, C. T., & Campbell, F. A. (1979). Compensatory education for disadvantaged children. *The School Review*, *87*(2), 171–189. https://doi.org/10.1086/443468

Republic of Ghana (2004). *Early Childhood Care and Development Policy*. Accra, Ghana: Ministry of Women and Children's Affairs. Available from: https://new-ndpcstatic.s3.amazonaws.com/pubication/Early+Childhood+Care+and+Development+Policy.pdf. [last accessed 1 June 2018].

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

Sandefur, J. (2016). *Internationally comparable mathematics scores for fourteen African Countries (working paper no. 444)*. Washington, DC: Center for Global Development.

Schreiner, M., & Woller, G. (2010). A simple poverty scorecard for Ghana. *Consultative Group to Assist the Poorest*. Available from: http://SimplePovertyScorecard.com/GHA_2005_ENG.pdf.

Schwartz, K., Cappella, E., Aber, J. L., Scott, M. A., Wolf, S., & Behrman, J. R. (2019). Early childhood teachers' lives in context: Implications for professional development in under-resourced areas. *American Journal of Community Psychology*. https://doi.org/10.1002/ajcp.12325

Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., & Nores, M. (2005). *Lifetime effects: The high/scope perry preschool study through age 40 (monographs of the high/scope educational research foundation, 14)*. Ypsilanti, MI: High Scope Educational Research Foundation.

Shager, H. M., Schindler, H. S., Magnuson, K. A., Duncan, G. J., Yoshikawa, H., & Hart, C. M. (2013). Can research design explain variation in Head Start research results? A meta-analysis of cognitive and achievement outcomes. *Educational Evaluation and Policy Analysis*, *35*(1), 76–95. https://doi.org/10.3102/0162373712462453

Shonkoff, J. P., & D. A. Phillips (Eds.) (2000). *From neurons to neighborhoods: The science of early childhood development*. Washington, DC: National Academy of Sciences.

UNESCO (2014). School enrollment, primary and secondary (gross), gender parity index (GPI). United Nations Educational, Scientific, and Cultural Organization Institute for Statistics. http://data.worldbank.org/indicator/SE.ENR.PRSC.FM.ZS [last accessed 14 August 2016].

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

UNGEI (2012). What works in girls' education in Ghana. Ghana for the Ministry of Education and the Girls' Education Unit, Ghana Education Service. Available from: www.ungei.org/files/What_Works_in_Girls_Education_in_Ghana.pdf [last accessed 14 August 2016].

UNICEF (2015). *Ghana's District League Table 2015*. Available from: https://www.unicef.org/ghana/media_9697.html.

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

Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development*, *84*(6), 2112–2130. https://doi.org/10.1111/cdev.12099

Wolf, S., Aber, J. L., & Behrman, J. R. (2019) Ghana – quality for preschool impact evaluation 2017, endline survey; world bank data catalogue; Ref. GHA_2017_QPIE-EL_v01_M.

Wolf, S., Aber, J. L., Behrman, J. R., & Tsinigo, E. (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. https://doi.org/10.1080/19345747.2018.1517199

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. https://doi.org/10.1086/682902

Wolf, S., Halpin, P. F., Yoshikawa, H., Down, A., Pisani, L., & Borisova, I. (2017). Measuring school readiness globally: Assessing the construct validity and measurement invariance of the International Development and Early Learning Assessment (IDELA) in Ethiopia. *Early Childhood Research Quarterly*, *40*(4), 21–36. https://doi.org/10.1016/j.ecresq.2017.05.001

Wolf, S., & McCoy, D. C. (2019). Household socioeconomic status and parental investments: Direct and indirect relations with school readiness in Ghana. *Child Development*, *90*(1), 160–278. https://doi.org/10.1111/cdev.12899

World Bank (2016). *The world bank data – Ghana*. June 1, 2018 Available from: https://data.worldbank.org/indicator/SP.RUR.TOTL.ZS?locations=GH.

World Bank (2018). Data repository: World bank country and lending groups. Available from: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups [last accessed 24 April 2019].

Wuermli, A. J., Tubbs, C. C., Peterson, A., & Aber, J. L. (2015). Children and youth in low- and middle-income countries: Toward an integrated developmental and intervention science. *Child Development Perspectives*, *9*(1), 61–66. https://doi.org/10.1111/cdep.12108

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. https://doi.org/10.1037/a0038785

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.

Yoshikawa, H., Wuermli, A., Raikes, H., Kim, S., & Kabay, S. B. (2018). Toward high-quality early childhood development programs and policies at national scale: Directions for research in global contexts. *SRCD Social Policy Report*, *31*(1), 1–36. https://doi.org/10.1002/j.2379-3988.2018.tb00091.x

Young, R., & Johnson, D. R. (2015). Handling missing values in longitudinal panel data with multiple imputation. *Journal of Marriage and Family*, *77*, 277–294. https://doi.org/10.1111/jomf.12144

Yousafzai, A. K., Rasheed, M., Rizvhi, A., Armstrong, R., & Bhutta, Z. A. (2014). Effect of integrated responsive stimulation and nutrition interventions in the Lady Health Worker programme in Pakistan on child development, growth, and health outcomes: A cluster-randomised factorial effectiveness trial. *The Lancet*, *384*(9950), 1282–1293. https://doi.org/10.1016/S0140-6736(14)60455-4

Zaslow, M., Anderson, R., Redd, Z., Wessel, J., Daneri, P., Green, K., … Martinez-Beck, I. (2016). Quality thresholds, features, and dosage in early care and education: Introduction and literature review. *Monographs of the Society for Research in Child Development*, *81*(2), 7–26. https://doi.org/10.1111/mono.12236

Zelazo, P. D., & Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. *Child Development Perspectives*, *6*(4), 354–360.

Zuilkowski, S.S., Piper, B., Ong'ele, S., & Kiminza, O. (2018). Parents, quality, and school choice: Why parents in Nairobi choose low-cost private schools over public schools in Kenya's free primary education era. *Oxford Review of Education*, *44*(2), 258–274.

How to cite this article: Wolf S, Aber JL, Behrman JR, Peele M. Longitudinal causal impacts of preschool teacher training on Ghanaian children's school readiness: Evidence for persistence and fade-out. *Dev Sci*. 2019;22:e12878. https://doi.org/10.1111/desc.12878

APPENDIX

TABLE A1 Means and mean differences in select school, teacher, and child characteristics at baseline, by treatment condition

	Control	TT	TTPA		
	Mean or %			*F*-statistic	*p*-value
School characteristics					
Private school status	55.7%	56.1%	53.2%	0.08	0.923
No. of years school has been established	23	23	19	0.95	0.389
Total number of KG children in school	54	63	60	0.64	0.529
Total number of KG teachers on the payroll	2.0	2.3	2.2	0.98	0.376
Main language of instruction in KG1					
English only	10.5%	13.5%	7.5%	0.68	0.509
Mother tongue only	4.5%	1.4%	1.5%	0.90	0.407
Mixture of English and mother tongue	85.1%	85.1%	91.0%	0.70	0.496
Sample size (total = 240)	79	82	79		
Teacher characteristics					
Female	97.9%	97.4%	97.3%	0.05	0.953
Age	35.3	35.7	35.2	0.07	0.933
Years as a teacher	6.55	6.16	6.64	0.22	0.801
Years as a teacher in current school	3.37	3.47	3.21	0.17	0.842
Has any post-secondary training	60.0%	62.3%	58.7%	0.22	0.804
Has training in ECD	65.7%	72.1%	64.0%	1.25	0.288
Sample size (total = 444)	143	153	148		
Household and child characteristics					
Primary caregiver marital status (%)					
Married	56.2	64.7	62.9	2.07	0.039
Never married	20.4	17.7	18.5	−0.74	0.457
Widowed/separated/divorced	23.3	17.6	18.7	−1.72	0.085
Primary caregiver is female (%)	51.5	50.9	52.3	0.34	0.735
Age of primary caregiver (years)	41.5	40.6	40.4	−1.86	0.064
Household wealth (0–100)	56.0	57.1	55.7	0.79	0.430
Male household head literacy status (%)					
Male head is literate	14.7	15.5	17.3		
Male head is not literate	69.6	70.3	67.7	1.19	0.232
No male head	15.7	14.2	15.1	−0.08	0.935
Child is female (%)	50.0	48.5	49.0	0.27	0.764
Child age	5.25	5.17	5.25	1.02	0.361
KG1 (vs. KG2)	53.5	52.1	52.6	0.24	0.789
School readiness composite (% correct)	50.9	51.8	52.2	1.66	0.190
Sample size (total = 3,435)	1,088	1,180	1,167		

TABLE A2 Impacts on child school readiness for children present at baseline and second follow-up only

	b	*SE*	*p*-value	d_{wt}
Composite IDELA score				
TT versus control	0.014	0.008	0.071	0.091[a]
TTPA versus control	−0.010	0.008	0.202	−0.063
Early numeracy				
TT versus control	0.005	0.009	0.600	0.058
TTPA versus control	−0.011	0.009	0.194	−0.123
Early literacy				
TT versus control	0.014	0.011	0.224	0.075
TTPA versus control	−0.011	0.011	0.324	−0.061
Social–emotional				
TT versus control	0.021	0.011	0.049	0.126[b]
TTPA versus control	−0.005	0.011	0.661	−0.028
Executive function				
TT versus control	0.017	0.010	0.079	0.106[a]
TTPA versus control	−0.011	0.010	0.246	−0.070
Sample size = 2,657				

Note: Estimates are computed using observed scores, in three-level models: children nested in classrooms nested in schools. Effect sizes calculated accounting for the 3-level model structure (see Hedges, 2009). TT, teacher training condition; TTPA, teacher training plus parental-awareness condition. Sample includes only children present at baseline and at both follow-up points.

$^{+}p < 0.10$;

$^{*}p < 0.05$.