Changing literacy instruction in Kenyan classrooms: Assessing pathways of influence to improved early literacy outcomes in the HALI intervention

Sharon Wolf[a],[*], Elizabeth L. Turner[b], Matthew C.H. Jukes[c], Margaret M. Dubeck[c]

[a] University of Pennsylvania, United States
[b] Duke University, United States
[c] RTI International, United States

ARTICLE INFO

ABSTRACT

This study uses data from the Health and Literacy Intervention (HALI) program evaluation, an in-service teacher training program focused on early grade literacy instruction for class one teachers. We assess how changes in classroom instructional processes impacted by the HALI teacher training were associated with improved early literacy outcomes for children. We find that experimentally induced increases in exposure to print—measured both through changes to time spent reading in class and through print displayed in the classroom—were associated with improvements in students' reading fluency and reading comprehension. Implications for global education efforts to improve learning outcomes are discussed.

1. Introduction

Literacy is more than simply the ability to read and write. Literacy skills provide the gateway to effectively connecting to, interpreting and discerning the world in which one lives. In efforts to improve the education and literacy levels of their populations, countries across Sub-Saharan Africa have made tremendous strides in increasing school access to more children over the past fifteen years (UNESCO, 2015). Yet literacy levels still remain low across the continent. At around 72%, sub-Saharan Africa as a region has the lowest youth literacy rate in the world (UNESCO, 2012). In Kenya specifically, dramatic growth in primary school enrollment—from 65.4% in 2000 to 84.9 in 2012—has shown some success, with youth literacy rates increasing slightly from 82.4% in 2007 to 85.8% in 2015 (World Bank, 2017).

Research suggests that despite being in school, a large portion of children fail to learn functional literacy skills in the first three years of primary school (Gove and Cvelich, 2010; Uwezo, 2013). With global education goals shifting from access to school to access to high quality education and learning (United Nations, 2015), early literacy skills are critical to ensuring learning and literacy outcomes. While evidence is starting to accumulate as to what school-based strategies may work to improve literacy in early grades in sub-Saharan Africa (McEwan, 2015; Ganimian and Murnane, 2016) and in Kenya in particular (Piper and Zuilkowski, 2015; Piper et al., 2014; Piper et al., 2016), most studies do not unpack the mechanisms through which different strategies work (or do not). Research on which strategies support early literacy learning

have been conducted almost exclusively in high-income countries (e.g., Snow et al., 1998; Snow et al., 1995). What specific instructional strategies are malleable to intervention and promote early literacy skills, particularly in linguistically diverse contexts in sub-Saharan Africa? More research is needed on the key elements of successful programs.

Using data from the Health and Literacy Intervention program evaluation (HALI; Jukes et al., 2017), this study assesses which key classroom instructional practices impacted by the program were associated with improved student literacy outcomes over the course of one school year. We assess instructional practices related to the medium of instruction (i.e., using a written medium), the instructional focus (i.e., explicit instruction of letters and sounds, explicit instruction of phonological blending and segmenting), and print exposure (i.e., time spent reading in class and the display of print in the classroom).

1.1. Instructional practices that promote early literacy skills

A growing literature over the past three decades has enabled a deep understanding of specific literacy instructional practices that promote children's literacy skills. This work has been conducted almost exclusively in high-income countries. However, because of the similarities across languages of learning the relationship between oral language and print that represents that language, it may provide a framework from which to begin to study similar processes in other contexts.

The provision of high-quality literacy instruction is

---

multidimensional. In reviewing the literature on effective teaching practices in early learning classrooms of children ages five to eight, Hall (2013) discusses the importance of balancing the written form of language with its meaning, uses, and purposes. In addition, providing students opportunities to engage with text – both by reading and writing – is critical to a host of early literacy skills such as phonological awareness, word recognition, spelling patterns, vocabulary, punctuation, grammar, and text structure (Topping and Ferguson, 2005; Wray et al., 2001; Wilkinson and Townsend, 2000; Parr and Limbrick, 2010; Mazzoli and Gambrell, 2003).

In settings with low-income and ethnically diverse students, effective teachers devote more time to reading activities (including independent reading and writing in response to reading) than the moderately and least effective teachers in those same settings (Taylor et al., 1999, 2000; Taylor and Pearson, 2002; Topping and Ferguson, 2005). Time spent in reading activities is shown to be particularly beneficial for poor readers (Mol and Bus, 2011). In addition, successful literacy practices involve direct teaching of skills that can be applied in reading and writing. Unlike learning to speak, which children learn through exposure, reading and writing skills do not happen naturally (Lyon, 1998). And it has been found in *meta*-analyses and research reviews that most children benefit from explicit instruction to acquire basic literacy skills (Lyon, 1998; Stanovich, 2000) followed by systematic practice of those skills (NICHHD, 2000; Snow et al., 1998). In other words, not only teaching children word recognition through memorization, but through teaching students explicit skills to recognize words (e.g., blending sounds and syllables; Taylor et al., 2000; Topping and Ferguson, 2005).

Exposure to print (in addition to time spent reading) is another critical predictor of literacy skills. A *meta*-analytic review found that from infancy through early adulthood, across outcome domains of reading comprehension and technical reading and spelling, there were moderate to strong correlations with print exposure (Mol and Bus, 2011). Exposure to print was an important predictor of literacy at all ages, with an escalating affect for older ages. Specifically, in preschool and kindergarten, print exposure explained 12% of the variance in oral language skills, in primary school 13%, in middle school 19%, in high school 30%, and in college and university 34% (Mol and Bus, 2011).

1.2. The HALI intervention

The HALI intervention focused on two aspects of improving Grade 1 children's capacity to learn: health (through anti-malarial intervention) and quality literacy education (through in-service teacher training). Since there were no impacts of the health intervention on children's learning outcomes (Halliday et al., 2014), in this study we focus only on the literacy intervention. The literacy intervention consisted of three components: (1) in-service training, which comprised of a 3-day workshop at the start of the school year that included developmental theories on literacy acquisition in a subsequent language (i.e., not the child's first language) and guided opportunities to learn and use provided instructional materials, followed by a 1-day problem-solving and instructional materials development workshop four months after the commencement of the school year; (2) a manual containing 140 sequential, semi-scripted lesson plans for literacy sessions either for Kiswahili or English, and (3) weekly interactive text messages providing brief instructional tips and motivation to implement lesson plans. The intervention is described in greater detail elsewhere (Dubeck et al., 2015).

The literacy intervention sought to increase children's exposure to print and improve teachers' instructional practice by building on effective practices that were already in use locally. For example, Dubeck et al. (2012) found that during Kiswahili instruction, some teachers were explicitly teaching the relationship between sounds and syllables. The HALI intervention sought to expand this practice in Kiswahili and to encourage explicit blending of sounds in English and in general, to

help teachers use literacy skills in one language to aid literacy acquisition in another. The program also aimed to increase children's engagement with print by encouraging teachers to use written modes of instruction during common practices such as song and oral reading, and providing teachers with opportunities to create materials during the trainings to increase the amount of print in the classroom.

Main impacts of the program on teaching practices and children's early literacy skills are reported in Jukes et al. (2017). This study found that the literacy intervention improved classroom practices with effect sizes from 0.57 to 1.15 standard deviations, including greater time spent using written modes instruction and focusing on letters and sounds. In addition, the program improved three of four primary measures of children's early literacy skills at the end of both the first and second years of the program with effect sizes ranging from 0.12 to 0.64. The program also reduced student dropout from school from 5.3 to 2.1%. This study builds upon the reports of the direct effects of the HALI intervention.

### 1.3. The present study – a multiple mediation analysis

The present study builds on analyses of the direct impacts of the HALI intervention on classroom instruction and children's literacy outcomes reported to assess the relations between specific literacy practices and early literacy skills. Our analysis is also guided by an earlier observational study in the region (Dubeck et al., 2012) and considers both practices that were already in use locally before the intervention (e.g., connecting sounds and syllables), and practices that were deficient and encouraged through the intervention (e.g., interaction with text, blending and segmenting letters to form words and sounds). Using multi-level structural equation modeling (MSEM; Preacher et al., 2010), we consider the pathways between experimentally induced impacts on classroom literacy instruction and practice on Kenyan children's Kiswahili early literacy outcomes (Fig. 1).

This study contributes to the present body of knowledge in a number of ways. First, while several impact evaluations to improve students' academic achievement and literacy outcomes have been conducted over the last few years in sub-Sahara Africa in general (see McEwan, 2015 for a meta-analysis) and in Kenya in particular (Piper et al., 2014, 2016), this is the first to examine the mediating pathways of experimentally induced classroom instructional practice and learning outcomes in this context. Second, to our knowledge this is one of the first studies to assess which specific instructional practices are associated with children's literacy outcomes in Kenya. We guide our analysis with literature on successful literacy instructional practices that have been identified in other contexts to assess their applicability to teaching in the lingua franca of the region (i.e., English and Kiswahili) in Kenyan schools. The findings allow for more targeted approaches for training teachers to provide high quality literacy instruction.

2. Methods

2.1. Sampling procedures and participants

Data come from the first school year of the HALI literacy intervention impact evaluation, which occurred between January 2010 and March 2012 together with a program of screening and treatment for malaria. The evaluation involved a cluster randomized trial (Brooker et al., 2010), in which 101 public primary schools were randomly assigned to one of four arms receiving either: (i) the malaria intervention alone; (ii) the literacy intervention alone; (iii) both interventions combined; or (iv) neither intervention. There was no impact of the malaria intervention on any literacy outcome (Halliday et al., 2014). Thus, in this study, as with previous studies (Jukes et al., 2017), we collapse both treatment arms that received the literacy intervention (ii and iii) as treatment schools (*N* = 51 schools), and both arms that did not receive the literacy intervention (i and iv) as the control arm

Treatment status
(Level 2)

Classroom mediators
(Level 2)

Child outcome
(Level 1)
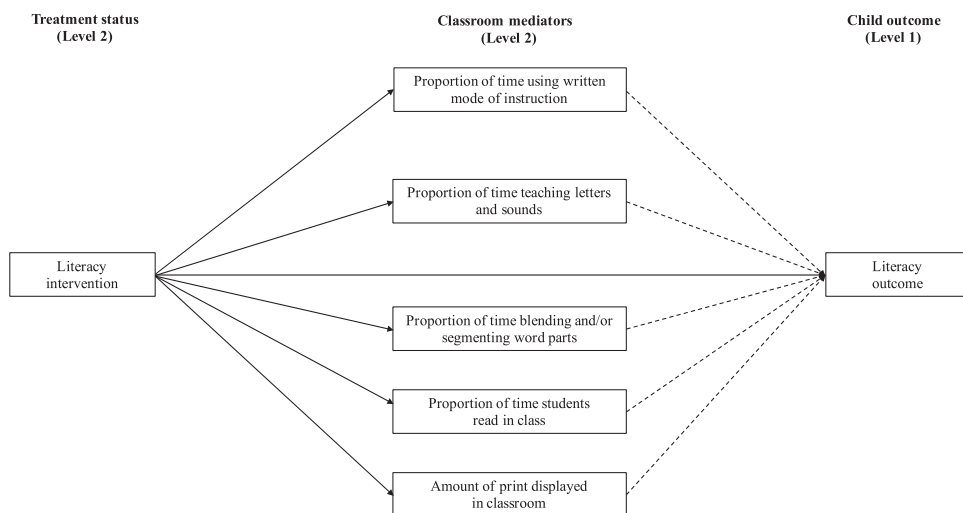
Fig. 1. Conceptual Model Testing the Relations of Experimentally Induced Classroom Changes with Child Early Literacy Outcomes.

Notes. All tested paths shown. Solid lines indicate causal paths; dashed lines indicate non-causal paths. Child age, gender, and baseline score (Level 1) included in all paths to child outcomes. All mediators are correlated. All exogenous variables (i.e., treatment status, child covariates) are correlated.

(N = 50 schools). At the time of the study, schools in Kenya were divided into Teacher Advisory Center (TAC) tutor zones, which meet to discuss curriculum and progress, supported by a TAC tutor. In our sample, there were 26 TAC tutor zones with between three and six schools each.

At the national level, the study was approved by the Division of Malaria Control, Ministry of Public Health and Sanitation and the Ministry of Education. At provincial and district levels, meetings were held with the Provincial Medical Officer and the Provincial Director of Education in Mombasa, as well as district health and education officials in Kwale and Msambweni. Consent was also sought from school head-teachers and Teacher Advisory Council (TAC) tutors. Prior to school randomization, enumeration of children was conducted and school meetings were held with parents and guardians to explain the study and seek written informed consent. Children provided verbal informed assent prior to assessment in the school. The study was approved by the Kenya Medical Research Institute and National Ethics Review Committee (SSC number 1543), the London School of Hygiene & Tropical Medicine Ethics Committee (5503), and the Harvard University Committee on the Use of Human Subjects in Research (F17578-101).

2.1.1. School and teacher sample

The study was conducted in rural government primary schools in Kwale and Msambweni districts, approximately 50 km south from Mombasa on the Kenyan coast. Schools that were more than 70 km from the project office and those with another literacy intervention under analysis were excluded, leaving 101 schools, 21 and 80 schools in Kwale and Msambweni districts respectively. All Grade 1 teachers were invited to participate in the study. At baseline, each school had one Grade 1 teacher (i.e., 1 classroom). By follow up, eight schools had split their Grade 1 class to two streams, resulting in 117 classrooms between treatment and control. In these instances, every effort was made to train the new grade 1 teacher in the treatment condition. All treatment school classrooms were analyzed in the treatment arm, even if the new teacher was not trained. Classroom observations were not conducted for three of those classrooms (in one school) due to teacher refusals. Thus, the present analysis included 114 classrooms in 100 schools. In our analysis, we include a dummy variable for classrooms in schools that had more than one first grade class.

2.1.2. Child sample

Children were randomly selected from the class roster. Parents of participating schools consented to enroll their child in the study at the community meetings that occurred before the study began. If a child was chosen whose parents did not give consent, the parents were visited

in their home to obtain consent. It the parent still did not consent, the child was withdrawn from the study and another child was selected in his/her place. This only occurred a few times; nearly all parents consented to enroll their child in the study. All participating Grade 1 children were assessed at baseline and 9-months later at the end of the academic year. The final sample for the present analysis consisted of 2206 Grade 1 children who had outcome measurements at the 9-month follow-up. Within classrooms, there were 19.5 students, on average (SD = 6.1).

2.2. Measures

2.2.1. Classroom mediators

During the second school term of the evaluation, before the follow-up outcome measurements, two unannounced visits to each school were carried out to conduct teacher interviews and observations. In intervention and control schools, an assessor observed English and Kiswahili lessons on two separate visits. In intervention schools only, the same assessor also observed a third lesson on each visit in which the HALI intervention materials were taught. Classroom scores were averaged across the two, or three, lessons observed. The classroom environment was rated on two dimensions: four types of instructional practices and print display in the classroom.

2.2.1.1. Direct observations. Teachers were observed by trained assessors. Classroom observers achieved adequate reliability levels on all classroom dimensions included (see Appendix A Table A1). Assessors rated teachers on an adapted version of the Classroom Language Arts Systematic Sampling and Instructional Coding (CLASSIC) classroom observation tool (Scanlon et al., 2003). Every 90 s a 'snapshot' of the classroom was taken to provide detailed information about student and teacher behavior in 5 categories. Teachers' scores were computed as the proportion of class time when the instructional focus was on a particular behavior. Teachers' scores were averaged across both lessons.

The following description summarizes these categories with emphasis on behaviors highlighted in our results: (1) The material to which the teacher is referring or to which the pupils are attending. It could be a visible object such as written text or something spoken like a rhyme. The key classification used in our analyses was whether the *material was written* (e.g., on a book or chalkboard); (2) the language part to which the teacher is referring or to which the pupils are attending. These were categorized as story or rhyme, sentence, word, word part, letter or sound, and in this analysis, we focus on *letters or sounds*; (3) the teacher's specific instructional focus (e.g., the meaning of the word or phonological segmentation). We focus on time spent *phoneme segmenting and blending* (i.e., breaking words down into

individual sounds); (4) the student response or activity. The key student response category of interest was the proportion of time *students spent reading in class*, either out loud, in groups, or individually.

2.2.1.2. Classroom print environment. Observers also rated the classroom using a classroom inventory checklist developed for this study to assess the richness of the display of print in the classroom environment in both English and Kiswahili. Assessors completed a checklist of how many of 39 different types of print displays, including a pocket chart and string for displaying print that were provided for intervention classrooms, were visible in the classroom (e.g., poems or rhymes displayed English alphabet; Kiswahili alphabet; days of the week; months of the year; labelled body parts; labelled shapes; maths numbers written with words). On average, at the classroom level, there was a mean of 18.5 (*SD* = 10.0) displays of print.

2.2.2. Kiswahili early literacy outcomes

Children's educational outcomes were assessed at baseline, and at 9 months after baseline. Literacy tests were adapted from the Early Grade Reading Assessment (EGRA; RTI International, 2009) which has been widely used in sub-Saharan Africa and elsewhere and from the Phonological Awareness Literacy Screening (Invernizzi et al., 2007). All tests were conducted at both time-points. We assessed four local language literacy outcomes that were impacted by the intervention (Jukes et al., 2017): Kiswahili letter sounds, Kiswahili word identification, Kiswahili reading comprehension, and Kiswahili reading fluency. Table 1 presents descriptives at the nine-month follow-up.

2.2.2.1. Kiswahili letter reading fluency. Children were given 60 s to read 100 letters presented in random order. The Kiswahili alphabet is presented in lower case and consists of 23 letters and seven digraphs (e.g., dh, sh) and children give the associated letter sound. The score was computed as the rate of letters identified correctly per minute. The task was discontinued if children could not identify a single letter in the first row of ten letters.

2.2.2.2. Kiswahili word reading fluency. Students were presented with a table of 50 words typically found in beginning reading materials. The score is the rate of words read correctly per minute. The task was discontinued if children could not read a single word in the first row.

2.2.2.3. Kiswahili passage reading fluency. Students read a grade 2 level narrative passage in Kiswahili. The score was computed as the rate of words read correctly per 60 s. The task was discontinued if children could not read a single word in the first row.

2.2.2.4. Kiswahili passage reading comprehension. After the passage reading fluency assessment, students were asked questions that corresponded to the text they read. It included four explicit comprehension and one inferential question. The score was computed as the total correct out of five.

2.2.2.5. Covariates. Four covariates were included in the models: child age, gender, and baseline spelling score; and a dummy variable for classrooms that were in schools with two classrooms.

2.3. Analytic plan

Guided by our theory of change, we analyzed indirect relationships between treatment status, classroom literacy instructional practice, and children's Kiswahili literacy outcomes. We used a multiple mediator, multilevel structural equation modeling (MSEM) framework as outlined by Preacher et al. (2010), using a 2-2-1 model where the predictor and mediators are at level 2 (classroom) and the outcomes are at level 1 (child) (Fig. 1). Given small intra-class correlations (ICCs) at the school-cluster level (see Table 3), we do not model a third level of nesting. Similarly, given that only 8 schools each had 2 classrooms, we do not model this level of nesting. In all models, mediators were correlated with each other, and all exogenous predictor variables (i.e., treatment statuts, child covariates) were correlated with each other.

In the analysis in the present study, we examined five potential classroom mediators. Causal impacts of the intervention on the classroom mediators were reported in Table 7 in Jukes et al. (2017). Four represented the proportion of classroom time spent on activities that have been established to be important to teaching early literacy skills: focus on letters and sounds, using a written mode of instruction, a focus on blending/segmenting word parts, and student reading. The fifth was a measure of the amount of print displays in the classroom. Covariates were included on all modeled pathways. We followed the approach recommended by Preacher et al. (2010) and include all five mediators in the analysis for each outcome. All analyses were conducted in MPlus version 6.12. For all models, adequate model fit was considered using the following criteria: a root mean square error of approximation (RMSEA) of < 0.06 (Hu and Bentler, 1995), a confirmatory fit index (CFI) of > 0.90 (Bentler, 1990), and the χ^2/df ratio, with values between 1 and 3 reflecting acceptable fit (Arbuckle, 2013).

Table 1
Descriptive statistics of key study variables.

	Full sample	Treatment	Control
	M (SD) or%		
Classroom mediators (6-months)			
Percent of class time observed on letters/sounds	3.2 (4.9)	5.7 (5.5)	0.0 (0.0)
Percent of class time observed using written mode	59.4 (16.0)	67.6 (11.4)	0.5 (0.2)
Percent of class time observed on blending/segmenting	19.5 (16.9)	26.7 (16.6)	0.1 (0.1)
Percent of class time observed with students reading	34.0 (14.1)	39.0 (11.9)	0.3 (0.1)
Print displayed on wall	18.52 (10.03)	21.41 (8.41)	15.69 (10.66)
Primary outcomes (9-months)			
Kiswahili letter sounds (lpm)	7.55 (11.48)	10.39 (13.11)	4.78 (8.79)
Kiswahili word identification (wpm)	6.04 (9.37)	7.14 (10.43)	4.98 (8.06)
Kiswahili reading comprehension (0–5)	0.34 (0.87)	0.40 (0.93)	0.29 (0.80)
Kiswahili reading fluency (wpm)	3.68 (8.28)	4.42 (9.98)	2.96 (7.22)
Select child characteristics			
Child is male	50.9%	52.2%	49.7%
Child age at baseline	7.83 (1.68)	7.74 (1.67)	7.92 (1.69)
Sample size	2238	1107	1131

Notes. lpm = letters per minute; wpm = words per minute.

Table 2
Bivariate correlations for mediating and dependent variables.

		1	2	3	4	5	6	7	8
1	Proportion of class time observed on letters/sounds	1.00							
2	Proportion of class time observed using written mode	0.44	1.00						
3	Proportion of class time observed on blending/segmenting	0.59	0.55	1.00					
4	Proportion of class time observed on with students reading	0.25	0.43	0.24	1.00				
5	Print displayed on wall	0.17	0.12	*0.01*	0.13	1.00			
6	Kiswahili letter sounds (lpm)	0.19	0.13	0.09	0.17	0.19	1.00		
7	Kiswahili word identification (wpm)	0.10	0.09	0.05	0.11	0.16	0.51	1.00	
8	Kiswahili reading comprehension (0–5)	0.05	0.06	*0.02*	0.11	0.15	0.40	0.79	1.00
9	Kiswahili reading fluency (wpm)	0.06	0.07	*0.02*	0.12	0.15	0.46	0.90	0.89

Note. Italicized numbers indicated correlation is not statistically significant at p < .05.

3. Results

3.1. Descriptive Statistics

Descriptive statistics for the full sample, and by treatment and control arm, are presented in Table 1. Overall, averaged across the lessons observed in this study, teachers spent time on teaching letters and sounds (3.2% of class time on average), and on blending/segmenting (19.5% of class time on average). Students spent about one-third of class time reading (34.0% on average), and more than half of class time using a written mode of instruction (e.g., blackboard or textbook; 59.4% on average). These five practices were moderately correlated, with correlations among the four practices ranging from 0.24 to 0.59 (see Table 2).

For child outcomes, we estimated the proportion of variance attributable to classrooms and children (nested in classrooms) by fitting a linear mixed model for each outcome with no covariates and only a random intercept for TAC, a random intercept for classroom and a residual error for students (see Table 3). For three of the four outcomes (word identification, reading comprehension, and reading fluency), nearly all (84–90%) of variance in children's literacy was accounted for by differences across children, with only approximately one percent of the variance explained by differences in TACs. Given small intra-class correlations (ICCs) at the TAC level (see Table 3), we do not model a third level of nesting in any subsequent analyses, such as the multiple mediator analyses. For the letter identification outcome, the majority of variance (65%) was explained by differences across children, 13% by differences across classrooms, and the remaining 12% by differences across TAC zones.

3.2. Mediation analysis

We next conducted a multiple mediation analysis to examine whether observed classroom practices and classroom displays of text mediated the relation between treatment status and children's early literacy outcomes. Results of the SEM model confirmed findings from Jukes et al. (2017), which used generalized linear modeling: positive and statistically significant direct relations were found between treatment status and all five classroom processes in each of the models. To avoid redundancy, we present the results of direct effects of treatment

Table 3
Intraclass correlations of dependent variables.

	Proportion of variance		
	Child	Classroom	TAC
Kiswahili Letter Sounds	0.651	0.229	0.120
Kiswahili Word Identification	0.837	0.156	0.007
Kiswahili Reading Comprehension	0.901	0.086	0.013
Kiswahili Reading Fluency	0.872	0.120	0.008

Notes. TAC = Teacher Advisory Centers.

on classroom processes in the text for the first outcome only. The results for the remaining four outcomes are all presented in Table 4.

3.2.1. Kiswahili letter sounds

The SEM model testing the relations between treatment status, classroom processes, and Kiswahili letter sounds was found to have excellent model fit: $\chi^2(18) = 50.83$; RMSEA = 0.029; CFI = 0.91; SRMR Within = 0.017; SRMR Between = .066. Standardized coefficients of the direct paths of this model, as well as the direct and indirect effects, are presented in the first column of Table 4.

Receipt of treatment status positively impacted: increased class time on letters and sounds ($\beta = 0.49$, S.E. = 0.06, $p < .001$); increased class time using a written model of instruction ($\beta = 0.50$, S.E. = 0.07, $p < .001$); increased time on blending/segmenting ($\beta = 0.39$, S.E. = 0.08, $p < .001$); increased time spent with students reading ($\beta = 0.36$, S.E. = 0.08, $p < .001$); and a display of more print/text in the classroom ($\beta = 0.30$, S.E. = 0.09, $p < .001$). Treatment status was still a positive predictor of letter sound knowledge scores in this model ($\beta = 0.19$, S.E. = 0.09, $p < .05$).

In addition to the direct effects, the main interest of this study was to assess indirect effects among the study variables. There were no significant relations between classroom processes and Kiswahili letter sounds, or significant indirect effects between treatment status and Kiswahili letter sounds.

3.2.2. Kiswahili word identification

The SEM model testing the relations between treatment status, classroom processes, and Kiswahili word identification was found to have excellent model fit: $\chi^2(18) = 45.75$; RMSEA = 0.026; CFI = 0.98; SRMR Within = 0.017; SRMR Between =.071. Standardized coefficients of the direct paths of this model, as well as the direct and indirect effects, are presented in the second column of Table 4. As in the model described above, treatment status positively impacted all mediating classroom processes. In addition, treatment status was a marginally statistically significant predictor of word identification scores in this model ($\beta = 0.15$, S.E. = 0.09, $p < .10$). There were no significant indirect effects between treatment status and Kiswahili word identification.

3.2.3. Kiswahili reading comprehension

The SEM model testing the relations between treatment status, classroom processes, and Kiswahili reading comprehension was found to have excellent model fit: $\chi^2(18) = 49.66$; RMSEA = 0.028; CFI = 0.94; SRMR Within = 0.017; SRMR Between = 0.071. Standardized coefficients of the direct paths of this model, as well as the direct and indirect effects, are presented in the third column of Table 4. As in the model described above, treatment status positively impacted all mediating classroom processes. Significant and direct relations were also found between some classroom processes and children's spelling scores. Specifically, display of print/text in the classroom was positively and significantly related to reading comprehension outcomes with a similar magnitude ($\beta = 0.38$, S.E. = 0.09, $p < .001$). In addition, the

Table 4
SEM Standardized Coefficients and Direct and Indirect Effects.

	Letter ID			Word ID			Comprehension			Fluency		
	Coeff	(S.E.)		Coeff	(S.E.)		Coeff	(S.E.)		Coeff	(S.E.)	
Path coefficients												
Within-level (L2)												
Intervention → Written mode	0.498	(0.065)	***	0.499	(0.064)	***	0.499	(0.064)	***	0.499	(0.064)	***
Intervention → Blending/segmenting	0.386	(0.084)	***	0.389	(0.083)	***	0.389	(0.083)	***	0.389	(0.083)	***
Intervention → Letters/sounds	0.488	(0.062)	***	0.488	(0.062)	***	0.488	(0.062)	***	0.361	(0.096)	***
Intervention → Student reading time	0.363	(0.084)	***	0.362	(0.095)	***	0.362	(0.095)	***	0.361	(0.096)	***
Intervention → Print displayed	0.304	(0.086)	***	0.304	(0.086)	***	0.304	(0.086)	***	0.304	(0.086)	***
Between-level												
Written mode → Outcome	−0.040	(0.079)		−0.009	0.077		−0.030	0.094		−0.044	0.081	
Blending/segmenting → Outcome	−0.149	(0.105)		−0.016	0.069		−0.011	0.123		−0.016	0.088	
Letters/sounds → Outcome	0.137	(0.082)	+	0.053	(0.089)		−0.022	0.107		−0.038	0.088	
Student reading time → Outcome	0.085	(0.085)		0.042	(0.068)		0.251	(0.081)	**	0.165	(0.068)	*
Print displayed → Outcome	0.049	(0.070)		0.111	(0.080)		0.376	(0.093)	***	0.251	(0.089)	**
Intervention → Outcome	0.188	(0.088)	*	0.145	(0.089)	+	−0.096	(0.133)		0.104	(0.105)	
Direct and Indirect Effects												
Total	0.224	(0.077)	**	0.209	(0.080)	**	0.079	(0.128)		0.194	(0.091)	*
Total indirect	0.035	(0.056)		0.064	(0.057)		0.176	(0.074)	*	0.089	(0.063)	
Specific indirect												
Intervention → Written mode → Outcome	−0.020	(0.040)		−0.004	(0.038)		−0.015	(0.047)		−0.022	(0.040)	
Intervention → Blending/segmenting → Outcome	−0.058	(0.042)		−0.006	(0.027)		−0.004	(0.047)		−0.006	(0.035)	
Intervention → Letters/sounds → Outcome	0.067	(0.043)		0.026	(0.044)		−0.011	(0.052)		−0.018	(0.043)	
Intervention → Student reading time → Outcome	0.031	(0.031)		0.015	(0.026)		0.091	(0.039)	*	0.060	(0.031)	*
Intervention → Print displayed → Outcome	0.015	(0.021)		0.034	(0.027)		0.114	(0.045)	*	0.076	(0.037)	*
Direct												
Intervention → Outcome	0.188	(0.088)	*	0.145	(0.088)	+	−0.096	(0.133)		0.104	(0.105)	

Notes: +p < .10, * p < .05, **p < .01.

proportion of time students were reading was positively statistically significantly related to spelling outcomes ($\beta = 0.25$, S.E. = 0.08, $p < .01$). In this model, treatment status was no longer a significant direct predictor of reading comprehension.

We next assessed the indirect relations among the study variables. Given inconsistent mediation findings (i.e., some negative and some positive indirect effects; see MacKinnon et al., 2007), we follow the recommendations of Alwin and Hauser (1975) to use the absolute values of the direct and indirect effects prior to calculating the proportion mediated. The proportion of class time students read was found to be a statistically significant mediator, $\beta = 0.091$, S.E. = 0.039, $p < .05$, which is 27.5% of the total effects ($\beta = 0.331$) and 38.7% of the total indirect effects ($\beta = 0.235$). In addition, display of print/text in the classroom was found to be a statistically significant mediator of the relation between treatment status and children's Kiswahili reading comprehension, $\beta = 0.114$, S.E. = 0.045, $p < .05$. This is 34.4% of the total effects, and 48.5% of the total indirect effects. Together these two classroom mediators make up 87.1% of the total indirect effects and well over half (61.9%) of the total effects of the intervention on Kiswahili reading comprehension.

3.2.4. Kiswahili reading fluency

The SEM model testing the relations between treatment status, classroom processes, and Kiswahili reading fluency was found to have excellent model fit: $\chi^2(18) = 48.09$; RMSEA = 0.027; CFI = 0.96; SRMR Within = .017; SRMR Between = .072. Standardized coefficients of the direct paths of this model, as well as the direct and indirect effects, are presented in the fourth column of Table 4. As in the model described above, treatment status positively impacted all mediating classroom processes. In addition, significant and direct relations were also found between some classroom processes and children's spelling scores. Specifically, display of print/text in the classroom was positively and significantly related to spelling outcomes with a similar magnitude ($\beta = 0.25$, S.E. = 0.09, $p < .01$). In addition, the proportion of time student read was positively statistically significantly related to reading

fluency ($\beta = 0.17$, S.E. = 0.07, $p < .05$). In this model, treatment status was no longer a significant direct predictor of reading comprehension.

In addition to the direct effects described above, we assessed indirect effects among the study variables. The proportion of class time devoted to students' reading was found to be a statistically significant mediator, $\beta = 0.060$, S.E. = 0.031, $p < .05$, which is 21.0% of the total effects ($\beta = 0.286$) and 33.0% of the total indirect effects ($\beta = 0.182$). In addition, display of print/text in the classroom was found to be a statistically significant mediator of the relation between treatment status and children's Kiswahili reading comprehension, $\beta = 0.076$, S.E. = 0.037, $p < .05$. This is 26.6% of the total effects, and 41.8% of the total indirect effects. Together these two classroom mediators make up 47.6% of the total effects of the intervention on Kiswahili reading fluency.

4. Discussion

This study assessed the components of teachers' instructional practices that were associated with improvements in class one children's early literacy skills in coastal Kenya using data from the HALI program evaluation. As noted above, the HALI intervention was designed to improve children's early literacy skills through increasing children's exposure to literacy instructional quality by providing teachers with specific tools and lessons to incorporate in to their pedagogical practice. A recent review of in-service teacher training evaluations in LMICs that shows trainings are more effective if they focus on specific pedagogical techniques (Popova et al., 2016). As global efforts to provide high quality education and learning for all children continue, understanding which pedagogical skills should be targeted when resources are limited is a critical area where research can inform practice. We provide evidence that the HALI intervention was indeed effective in improving learning because of specific, observable changes made in the classroom. Specifically, experimentally induced changes in children's exposure to print (both through increased time spent reading and the classroom print environment) measured mid-way through the school year were associated with improvements in children's early literacy skills at the

end of the year.

A question that arises is why experimentally induced changes to the other hypothesized classroom mediators − time spent in class on phoneme blending/segment, teaching letters and sounds, and using a written mode of instruction − were not predictive of children's early learning outcomes. These are considered important teaching practices to promote early learning skills (Hall, 2013). But these may be important only to the extent that they are implemented with high quality (e.g., LaParo et al., 2009). It is possible that while the intervention increased the *quantity* of time teachers spent on these practices, it did not improve the *quality* of instruction. Thus, while teachers focused more on these key areas as a result of the intervention, they might not have done so in a way that promoted learning. The two significantly predictive mediators of time spent reading and exposure to print are more student-directed and less dependent on high quality instruction. Notably, these are two aspects of instructional practice that were found to be lacking in classrooms in the region before the intervention (Dubeck et al., 2012). If this is the case, it suggests that teachers need more ongoing support, in addition to three days of training, semi-scripted lessons and weekly text message support, in order to improve the quality, and not just the quantity, of their instructional practice. Research in the United States points to the promising role of in-class monitoring and coaching in promoting improvements to teaching quality (Allen et al., 2011). And indeed, a few recent studies indicate that this approach holds promise in sub-Saharan Africa as well. Integrating coaching to teacher in-service support in Kenya showed teacher coaching improved children's literacy outcomes, and was more effective when coaches had fewer schools to cover. The USAID Early Grade Reading Assessment (EGRA) Plus: Liberia program showed large impacts of student achievement of coach-to-school ratios of 4:1 and saw large impacts on student achievement (Piper and Korda, 2011). Another recent study in pre-primary education in Ghana found that in-service teacher training paired with coaching and monitoring by district education staff had moderately sized impacts on observed teacher-child interaction quality, as well as small impact on children's learning outcomes (Wolf et al., 2017). Finding ways to provide such supports for teachers within existing systems – for example, through the local government education sector – may be one way to improve teaching quality.

This study provides applicable tools for future educational research on teacher training. First, while observations of instructional practice and teacher-child interactions is an intensive process, the checklist measure of the classroom print environment is a simple tool to both adapt and to use reliably to capture meaningful classroom differences. What is needed is simply familiarity with the types of classroom materials that are used in a particular context for literacy instruction. Second, using multiple mediation methods, as employed in this study, allows for the assessment of multiple changes to the classroom environment to be examined simultaneously as they relate to children's learning outcomes. Future education intervention research should

hypothesize and collect data on classroom mediators to understand key mechanisms of change. This type of research is necessary for replicability and scalability of successful programs.

This study has limitations that must be considered when interpreting the results. The first is the skewed distributions of the mediator's variables. Specifically, the proportion of class time spent on letters and sounds was very small − around three percent. While there were treatment impacts on this dimension of the classroom environment, the variation in this mediating variable was extremely limited. The other four mediators were observed with more variability across classrooms, and none of these four were statistically skewed. Second, we are limited in our ability to draw causal inferences between the classroom mediators and children's outcomes. While the intervention was randomized, and thus causal claims can be made regarding changes in the classroom environment, our conclusions about the relations between changes in classrooms and children's early literacy outcomes are purely descriptive.

5. Conclusions

While education spending levels and enrollment rates in schools have increased across the developing world, a variety of research studies and datasets show that learning levels remain low. A recent surge in rigorous education research has shed light on possible interventions and policies that could be employed to improve teaching and learning. This study demonstrates *how* an evidence-based approach to instruction can make a significant impact on the literacy development of Kenyan children in the early grades of primary school by unpacking key changes to teachers' practice. This focus on unpacking of mechanisms of action is critical for applying the lessons of this evaluation in other contexts (White, 2009), and it is important that future research in other settings assess pathways of influence in addition to main impacts to continue to build the evidence base of both "what works" and "how".

While this study was conducted in coastal districts in Kenya, the findings may be applicable to other contexts facing similar challenges in literacy instruction where teachers currently place little emphasis on students interacting with written text. While global efforts to train and support teachers and improve high quality education and literacy skills have increased, scarce resources must be targeted to the key education levers that promote learning. This study points to the importance of children's exposure to print as one of those key levers, and suggests that training teachers to support children's interaction with text (both actively and passively by displaying text in the classroom) is of central importance in areas where children have little interaction with text. Future teacher training programs, that focus on teacher classroom behaviors, should consider adding a component on exposure to print. Additional research should be conducted to guide the allocation of resources to support high quality teaching practice and learning for all children.

Appendix A

Table A1
Inter-rater reliability for classroom observation of behaviors

	Fleiss Kappa	Reliability level[a]
Proportion of class time observed on letters/ sounds	0.45	Fair
Proportion of class time observed using written mode	0.78	Moderate
Proportion of class time observed blending/ segmenting	0.45	Fair
Proportion of class time observed with students reading	0.88	Substantial

Note. Inter-rater reliability for observation of behaviors in 10 classes by 10 raters.
[a] See Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159–174.

References

Allen, J.P., Pianta, R.C., Gregory, A., Mikami, A.Y., Lun, J., 2011. An interaction-based approach to enhancing secondary school instruction and student achievement. Science 333 (6045), 1034–1037.

Alwin, D.F., Hauser, R.M., 1975. The decomposition of effects in path analysis? Am. Sociol. Rev. 40 (1), 37–47.

Arbuckle, J.L., 2013. IBM SPSS AMOS 22 User's Guide. Amos Development Corp., Spring House, PA.

Bentler, P.M., 1990. Comparative fit indexes in structural models. Psychol. Bull. 107, 238–246. http://dx.doi.org/10.1037/0033-2909.107.2.238.

Brooker, S., Okello, G., Njagi, K., Dubeck, M.M., Halliday, K.E., Inyega, H., Jukes, M.C.H., 2010. Improving educational achievement and anaemia of school children: design of a cluster randomised trial of school-based malaria prevention and enhanced literacy instruction in Kenya. Trials 11, 1–14.

Dubeck, M.M., Jukes, M.C.H., Okello, G., 2012. Early primary literacy instruction in Kenya. Comp. Educ. Rev. 56 (1), 48–68.

Dubeck, M.M., Jukes, M.C.H., Brooker, S.J., Drake, T.L., Inyega, H.N., 2015. Designing a program of teacher professional development to support beginning reading acquisition in coastal Kenya. Int. J. Educ. Dev. 41, 88–96.

Ganimian, A.J., Murnane, R.J., 2016. Improving education in developing countries: lessons from rigorous impact evaluations. Rev. Educ. Res. 1–37.

Gove, A., Cvelich, P., 2010. Early Reading: Igniting Education for All. Research Triangle International, Research Triangle Park, NC.

Hall, K., 2013. Effective literacy teaching in the early years of school: a review of evidence. In: Larson, J., March, J. (Eds.), The SAGE Handbook of Early Childhood Literacy. SAGE Publications Ltd, London, UK.

Halliday, K.E., Okello, G., Turner, E.L., Njagi, K., Mcharo, C., Kengo, J., et al., 2014. Impact of intermittent screening and treatment for malaria among school children in Kenya: a cluster randomised trial. PLoS Med. 11 (1).

Invernizzi, M., Juel, C., Swank, L., Meier, J., 2007. Phonological Awareness Literacy Screening (PALS): Kindergarten. University of Virginia, Charlottesville, VA.

Jukes, M.C.H., Turner, E.L., Dubeck, M.M., Halliday, K.E., Inyega, H.N., Wolf, S., Zuilkowski, S.S., Brooker, S.J., 2017. Improving literacy instruction in Kenya through teacher professional development and text messages support: a cluster randomized trial. J. Res. Educ. Eff. 10 (3), 449–481.

LaParo, K.M., Hamre, B.K., Locasale-Crouch, J., Pianta, R.C., Bryant, D., Early, D., et al., 2009. Quality in kindergarten classrooms: observational evidence for the need to increase children's learning opportunities in early education classrooms. Early Educ. Dev. 20 (4), 657–692.

Lyon, R.G., 1998. Why reading is not a natural process? Educ. Leadersh. 55 (6), 14–18.

MacKinnon, D.P., Fairchild, A.J., Fritz, M.S., 2007. Mediation analysis. Annu. Rev. Psychol. 58, 593–614.

Mazzoli, S.S., Gambrell, L.B., 2003. Principles of best practice: finding the common ground. In: Morrow, L.M., Gambrell, L.B., Pressley (Eds.), Best Practices in Literacy Instruction, 2nd edn. Guilford Press, New York.

McEwan, P.J., 2015. Improving learning in primary schools of developing countries: a meta-analysis of randomized experiments. Rev. Educ. Res. 85, 353–394.

Mol, S.E., Bus, A.G., 2011. To read or not to read: a meta-analysis of print exposure from infancy to early adulthood. Psychol. Bull. 137 (2), 267–296.

National Institute of Child Health and Human Development (NICHHD), 2000. Report of the National Reading Panel. Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction (NIH Publication No. 00-4769). U.S. Government Printing Office, Washington, DC.

Parr, J.M., Limbrick, L., 2010. Contextualising practice: hallmarks of effective teachers of writing. Teach. Teach. Educ. 26 (3), 583–590.

Piper, B., Korda, M., 2011. EGRA plus: Liberia. understanding the causal mechanisms: EGRA plus's effectiveness. Prepared for USAID/Liberia Under the Education Data for Decision Making (EdData II) Project, Early Grade Reading Assessment (EGRA): Plus Project, Task Order No. EHC-E-06-04-00004-00 (RTI Task 6). RTI International, Research Tri- angle Park. NC, USA Retrieved from. https://www.eddataglobal.org/countries/index.cfm?fuseaction1/4pubDetail&ID1/4540.

Piper, B., Zuilkowski, S.S., 2015. Teacher coaching in Kenya: examining instructional support in public and nonformal school. Teach. Teach. Educ. 47, 173–183.

Piper, B., Zuilkowski, S.S., Mugenda, A., 2014. Improving reading outcomes in Kenya: first-year effects of the PRIMR Initiative. Int. J. Educ. Dev. 37, 11–21.

Piper, B., Zuilkowski, S.S., Kwayumba, D., Strigel, C., 2016. Does technology improve classroom reading outcomes: comparing the effectiveness and cost-effectiveness of ICT interventions for early grade literacy in Kenya. Int. J. Educ. Dev. 49, 204–214.

Popova, A., Evans, D.K., Arancibia, V., 2016. Training Teachers on the Job. World Bank, Washington, D.C Retrieved Apr 10, 2017 from: https://openknowledge.worldbank.org/handle/10986/25150.

Preacher, K.J., Zyphur, M.J., Zhang, Z., 2010. A general multilevel SEM framework for assessing multilevel mediation. Psychol. Methods 15, 209–233.

RTI International, 2009. Early grade reading assessment toolkit. Prepared for the World Bank. Office of Human Development, Research Triangle Park, NC. Author. https://www.eddataglobal.org/documents/index.cfm?fuseaction=pubDetail&ID=149.

Scanlon, D.M., Gelzheiser, L., Fanuele, D., Sweeney, J., Newcomer, L., 2003. Classroom Language Arts Systematic Sampling and Instructional Coding (CLASSIC). Child Research and Study Center, The University at Albany, Albany, NY.

Snow, C.E., Tabors, P.O., Nicholson, P.A., Kurland, B.F., 1995. SHELL: Oral language and early literacy skills in kindergarten and first-grade children. Journal of Res. Child. Educ. 10 (1), 37–48.

Snow, C.E., Burns, M.S., Griffin, P. (Eds.), 1998. Preventing Reading Difficulties in Young Children. National Academies Press, Washington, DC.

Stanovich, K.E., 2000. Matthew effects in reading: some consequences of individual differences in the acquisition of literacy. Read. Res. Q. 21, 360–406.

Taylor, B.M., Pearson, P.D. (Eds.), 2002. Teaching Reading: Effective Schools, Accomplished Teachers. Erlbaum, Mahwah, NJ.

Taylor, B.M., Pearson, D.P., Clark, K.F., Walpole, S., 1999. Effective schools/accomplished teachers. Read. Teach. 53 (2), 156–159.

Taylor, B.M., Pearson, P.D., Clark, K., Walpole, S., 2000. Effective schools and accomplished teachers: lessons about primary-grade reading instruction in low-income schools'. Elementary School J. 101 (2), 121–165.

Topping, K., Ferguson, N., 2005. Effective literacy teaching behaviours. J. Res. Read. 28 (2), 125–143.

UNESCO, 2012. EFA Global Monitoring Report: Youth and Skills, Putting Education to Work. Author, Paris, France.

UNESCO, 2015. Education for All 2000–2015: Achievements and Challenges. Global, pp. 2015.

United Nations, 2015. Transforming Our World: the 2030 Agenda for Sustainable Development. Author, New York.

Uwezo, 2013. Are Our Children Learning? Literacy and Numeracy Across East Africa. Retrieved from. http://www.uwezo.net/wp-content/uploads/2012/08/2013-Annual-Report-Final-Web-version.pdf.

Wilkinson, I.A.G., Townsend, M., 2000. From rata to rimu: grouping for instruction in best practice New Zealand classrooms. Read. Teach. 53, 460–471.

Wolf, S., Aber, J.L., Behrman, L., 2017. Experimental Evaluation of the 'Quality Preschool for Ghana' Intervention on Teacher Professional Well-Being, Classroom Quality and Children's School Readiness. (Manuscript under revision).

World Bank, 2017. Literacy rate, youth total (% of people ages 15–24). Retrieved July 1, 2017 from: https://data.worldbank.org/indicator/SE.ADT.1524.LT.ZS?locations=KE.

Wray, D., Medwell, J., Fox, R., Poulson, L., 2001. The teaching practices of effective teachers of literacy. Educ. Rev. 52 (1), 75–84.