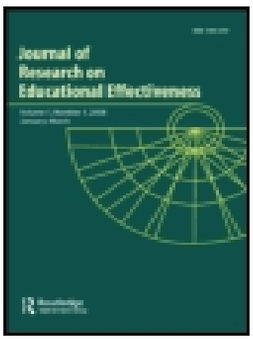

Routledge
Taylor & Francis Group

Journal of Research on Educational Effectiveness

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

Effectiveness of Large-Scale, State-Sponsored Language and Literacy Professional Development on Early Childhood Educator Outcomes

Shayne B. Piasta, Laura M. Justice, Ann A. O'Connell, Susan A. Mauck, Melissa Weber-Mayrer, Rachel E. Schachter, Kristin S. Farley & Caitlin F. Spear

To cite this article: Shayne B. Piasta, Laura M. Justice, Ann A. O'Connell, Susan A. Mauck, Melissa Weber-Mayrer, Rachel E. Schachter, Kristin S. Farley & Caitlin F. Spear (2017) Effectiveness of Large-Scale, State-Sponsored Language and Literacy Professional Development on Early Childhood Educator Outcomes, Journal of Research on Educational Effectiveness, 10:2, 354-378, DOI: 10.1080/19345747.2016.1270378

To link to this article: http://dx.doi.org/10.1080/19345747.2016.1270378

View supplementary material

Accepted author version posted online: 14 Dec 2016.
Published online: 14 Dec 2016.

Submit your article to this journal

Article views: 85

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=uree20

Download by: [74.140.109.0]

Date: 06 April 2017, At: 10:23

JOURNAL OF RESEARCH ON EDUCATIONAL EFFECTIVENESS
2017, VOL. 10, NO. 2, 354–378
http://dx.doi.org/10.1080/19345747.2016.1270378

Effectiveness of Large-Scale, State-Sponsored Language and Literacy Professional Development on Early Childhood Educator Outcomes

Shayne B. Piasta[a], Laura M. Justice[a], Ann A. O'Connell[a], Susan A. Mauck[a], Melissa Weber-Mayrer[a], Rachel E. Schachter[a], Kristin S. Farley [a], and Caitlin F. Spear[a]

ABSTRACT

The current study investigated the effectiveness of large-scale, state-sponsored language and literacy professional development (PD) intended to improve early childhood educators' knowledge, beliefs, and practices. PD was offered in a real-world context and delivered at-scale across the state, implemented by an independent contractor. Educators ($n = 535$) were randomly assigned to participate in one of three types of PD: 30 hrs of language and literacy PD presented in a workshop format, 30 hrs of language and literacy PD plus monthly coaching, or PD on alternative topics (comparison). Baseline and outcome measures were collected by an independent research team to determine PD effectiveness. Growth curve analyses showed little change in educator outcomes over an 18-month period, with no differences among conditions or evidence of moderation by educator or classroom characteristics. Explanations for lack of PD impacts and the importance of evaluating educational investments are discussed.

KEYWORDS

professional development
effectiveness trial
early childhood
language and literacy

Over the last two decades, such influential reports as those produced by the National Reading Panel (2000) and the National Early Literacy Panel (2008) have substantially increased our understanding of the skills that children need to develop to become proficient readers. For instance, both reports discuss the importance of code-focused (e.g., alphabetics and phonological awareness) and meaning-focused (e.g., vocabulary) skills as critical contributors to skilled reading, and present specific empirically validated approaches that educators can use to improve children's skills in these areas. Consequently, there has been considerable attention directed toward identifying effective avenues for increasing early educators' use of such approaches in their classrooms, typically serving 3- to 5-year-olds, through professional development (PD).

PD represents an array of strategies that are used to improve educators' knowledge, beliefs, and/or practices in key areas. These can range from attendance at brief workshops and seminars at educational conferences to participation in extended coursework, in-class coaching, and/or local communities of practice. Large-scale surveys have sought to identify the attributes of effective PD, such as those strategies most influential for affecting educators'

CONTACT Shayne B. Piasta piasta.1@osu.edu Department of Teaching and Learning, The Ohio State University, Columbus, OH 43210, USA.
[a]The Ohio State University, Columbus, Ohio, USA
Supplemental data for this article can be accessed at http://dx.doi.org/10.1080/19345747.2016.1270378.
© 2017 Taylor & Francis Group, LLC

knowledge and classroom practice. Specifically, Garet and colleagues reported that PD is most effective when it focuses on improving educators' content knowledge, features opportunities for active learning, and is integrated and coherent with other growth opportunities (Garet, Porter, Desimone, Birman, & Yoon, 2001). The results of a number of carefully designed studies suggest that PD that incorporates these strategies can positively influence a variety of educator outcomes, such as discplinary knowledge (Cunningham, Etter, Platas, Wheeler, & Campbell, 2015), beliefs about their own efficacy (Ross & Bruce, 2007), and instructional practices in the classroom (Desimone, Porter, Garet, Yoon, & Birman, 2002; Jackson et al., 2006; Piasta et al., 2010). Bringing about change in educators' knowledge, beliefs, and instructional practices is viewed as central in theories of change linking PD to improved student achievement, including early literacy (Schachter, 2015), for theoretical and empirical reasons. Theoretically, educators' knowledge and beliefs guide selection and implementation of classroom literacy practices, with all three constructs expected to influence children's learning (Schachter, Spear, Piasta, Justice, & Logan, 2016; Sheridan, Edwards, Marvin, & Knoche, 2009). Such connections also have empirical support in research studies, although recent work suggests that these associations may be more complex than initially understood (see Markussen-Brown et al., 2017; Schachter et al., 2016).

The research community has particular interest in designing and testing PD that improves early educators' capacity to help prekindergarten-aged children develop the early-literacy skills that are foundational to future proficient reading (Cunningham et al., 2015; Hamre et al., 2010; Jackson et al., 2006; Landry, Swank, Smith, Assel, & Gunnewig, 2006; Pianta, Mashburn, Downer, Hamre, & Justice, 2008; Piasta et al., 2010). Such efforts draw upon general consensus regarding those skills that are especially important predictors of skilled reading, such as vocabulary and alphabet knowledge (see National Early Literacy Panel, 2008) as well as a large volume of work identifying specific classroom practices and programs that can help young children develop these early-literacy skills (e.g., Arnold, Lonigan, Whitehurst, & Epstein, 1994; Beck & McKeown, 2007; Biemiller & Boote, 2006; Cunningham et al., 2015; Hamre et al., 2010; Jackson et al., 2006; Justice, Kaderavek, Fan, Sofka, & Hunt, 2009; Pianta et al., 2008; Piasta et al., 2010). Further, investigations of early educators' classroom practices suggest that many provide only limited opportunities for children to develop these skills (Cabell, DeCoster, LoCasale-Crouch, Hamre, & Pianta, 2013; Cunningham, Zibulsky, & Callahan, 2009; Justice, Mashburn, Hamre, & Pianta, 2008; Pelatti, Piasta, Justice, & O'Connell, 2014).

Given such circumstances, researchers have conducted well-controlled investigations of the effects of PD participation on the knowledge, beliefs, and/or practices of early educators with respect to early-literacy instruction (Assel, Landry, Swank, & Gunnewig, 2007; Hamre et al., 2010; Jackson et al., 2006; Landry, Swank, Anthony, & Assel, 2011; Landry et al., 2006; Lonigan, Farver, Phillips, & Clancy-Menchetti, 2011; Milburn et al., 2015). These investigations seek to understand the extent to which educators exhibit changes in key outcomes as a function of PD participation, such as the way they organize the classroom for instruction (e.g., Landry et al., 2011). Interestingly, several investigations have featured planned variations of PD to identify optimal ways to design and deliver PD (Jackson et al., 2006; Lonigan et al., 2011). Jackson and colleagues (2006), for instance, examined the effects of a 15-week, 44-hr national satellite broadcast PD program on the classroom literacy practices of 31 early childhood educators as compared to 24 controls. A subset of educators receiving PD also received mentoring in their classrooms over the 15-week period from a trained consultant

over four to six sessions. Given this design, the study sought to determine the value-added of including mentoring in PD, in addition to assessing the overall effects of this PD program on educators' practices. Findings indicated that educator participation in the PD had positive effects on their classroom practices, but that exposure to mentoring had no added value, with changes in educator practices in the PD and PD+ mentoring groups largely equivocal (see also Assel et al., 2007; Lonigan et al., 2011). Such work suggests that participation in an extensive PD program can have positive effects on early educators' instructional practices, but that additional in-class mentoring may not be a necessary component.

The present study was conducted to examine the effectiveness of a large-scale, state-sponsored effort to improve early educators' knowledge, beliefs, and practices, with educators participating in either a basic 30-hr workshop, the workshop plus intensive in-class coaching, or a comparison condition. Although Jackson and colleagues' work did not see a positive impact of their PD+ mentoring condition over the basic PD, coaching is often included in state-sponsored PD given that there is some evidence suggesting that it may be a critical component of effective PD in state-sponsored efforts (Landry, Anthony, Swank, & Monseque-Bailey, 2009; Landry et al., 2011). To this point, although there are a number of exemplars in the literature regarding what effective PD should look like for early educators, relatively few studies have assessed the effects of PD offerings designed and developed outside of the researcher community. Presently, early educators have numerous opportunities to engage in PD offerings available through not only their employer but also a range of for- and non-profit efforts, the latter supported via extensive state and federal commitments to elevate the quality of the early-education workforce (e.g., Early Head Start, Race to the Top). For instance, more than $1 billion in federal funding has been provided to states via the federal Race to the Top Early Learning Challenge program since 2011, which supports, in part, provision of state-sponsored PD opportunities for early educators in an effort to improve the quality of early education.

There is decidedly little evidence regarding the nature, quality, and impact of such real-world PD offerings and the value-added of such investments (cf. Landry et al., 2009, 2011). In fact, examinations of real-world PD find that a majority of educators report little to no change in their knowledge and practices as a result of participation (Horizon Research, 2002), and large-scale effectiveness research shows that impacts of real-world PD on educator practices may be conditional on the characteristics of the educators themselves. For instance, the results of Landry and colleagues' (2009) study of PD provided to 262 early childhood educators in four states suggested that effects of intensive in-class coaching (twice monthly 2-hr sessions) on overall teaching quality was conditional on educators' baseline teaching quality, such that the PD only benefited educators who began with poor to average teaching quality.

All in all, there is a need to understand whether educators participating in large-scale PD programs do, in fact, benefit from these experiences, in light of the points made earlier as well as several additional considerations. First, it is unclear whether the providers of PD in real-world conditions exemplify the quality of providers seen in empirical investigations, who typically receive extensive training and are closely monitored in their implementation of PD. In the Landry et al. (2009) study, PD was implemented by two highly trained and closely monitored facilitators. These facilitators completed a four-day intensive training workshop, participated in weekly conference calls with an overall project manager and weekly conference calls with the entire network of facilitators, and were observed on-site

several times by the manager. It is unclear whether a typical PD provider would have such ongoing support. Second, it is also plausible that the participants in real-world PD offerings may be distinct in important ways from those studied in scientific investigations of PD impacts. For instance, one large-scale investigation examined relations between participation in various aspects of PD on the early-literacy practices of 154 early educators (Hamre et al., 2010), all of whom were credentialed educators (bachelor's degree or higher) working in school-based prekindergarten programs. It is unclear whether similar PD results would be observed in a more diverse group of early educators, such as those with a more limited educational background or those working in other settings (e.g., center-based child care, home care). In fact, a high level of attrition (about 30%) was reported in one study of PD drawing participants from a variety of early-education settings, including Head Start and child care programs (Jackson et al., 2006), again suggesting that the impacts of PD may be conditional on the population of early educators being served, with some but not all participants seeming to benefit (see Landry et al., 2009).

The present study improves our understanding of PD impacts on the early-literacy knowledge and practices of early educators by examining educator outcomes as a function of participation in real-world, statewide PD. Specifically, the purpose was to examine the effectiveness of one large-scale state-sponsored PD effort for affecting early childhood educators' knowledge, beliefs, and practices when offered as a stand-alone 30-hr course or as coupled with a coaching component. This state-sponsored PD model has been offered continually for over 10 years at a cost of approximately $500,000 annually (excluding development and startup costs during initial years). As an effectiveness trial, this study differs in salient ways from many of the studies on PD conducted to date. First, the PD in which early educators participated was designed and implemented by contractors with the state, thus differing from other large-scale studies testing the effects of researcher-developer PD interventions (e.g., Assel et al., 2007; Hamre et al., 2010; Landry et al., 2009). Thus, results will improve our understanding of the effects seen in real-world PD offerings that are commonplace across the country and exemplify the PD programs in which many educators participate. Second, the early educators who elected to participate in the state-sponsored PD represented a broad range of the early education workforce and the contexts in which they work, including urban, suburban, and rural settings as well as early childhood special education settings. Examining PD impacts on early educators who work in diverse classroom settings is important, given evidence that PD impacts may be conditional on educator characteristics (see Landry et al., 2009).

To this end, the present study also examined the ways in which PD impacts may vary based on five select educator and classroom characteristics, namely educator education, educator credentialing, classroom locality (rural, urban, suburban), affiliation with Head Start, and classroom enrollment (early childhood special education vs. non-early childhood special education). With respect to educator education, we examined whether PD effects differed as a function of education level given prior work suggesting that PD may be more impactful when educators have higher levels of education (Landry et al., 2006). For educator credentialing, based on Early and colleagues' (2007) work suggesting higher quality practices for state-certified early educators, and the large variability in whether early educators hold such credentials, we assessed whether PD would differentially benefit certified and noncertified educators. Finally, the classroom characteristics anticipated to moderate effects were selected a priori because they are often emphasized in the literature as ways to represent the

large and diverse early-education sector into smaller, yet meaningful categorical groupings (e.g., Early et al., 2007; Miller & Votruba-Drzal, 2013). Because the PD examined in this study was offered across the state to educators in diverse programs, including those in rural settings and those working in early childhood special education programs, there was interest among the state sponsors of the PD in determining whether it seemed to particularly benefit educators working in various settings.

Method

Participants

Four sequential cohorts of educators participated in the study between 2010 and 2014; educators were allowed to participate in one cohort of the study only. All early childhood educators who registered for state PD opportunities were invited to participate in the study if they met the following criteria: (a) held a position as a lead, co-lead, or assistant educator, (b) directly taught prekindergarten-aged children (i.e., 3 to 5 years old), and (c) were open to a variety of PD opportunities (such that they could be randomly assigned to condition). PD opportunities are regularly offered to early childhood educators statewide, to meet regulations requiring that all early childhood educators participate in 10 hrs of PD per biennium. These opportunities are advertised by the state department of education and its contractors, such as the Early Childhood Quality Network (ecQ-net). During 2012 and 2013, research staff also contacted early childhood centers directly to share information about PD opportunities.

A total of 535 educators participated in PD/study activities and contributed data to the present analyses. Originally, 760 educators consented to participate, but 29% withdrew prior to the beginning of study activities, mainly due to employment and/or scheduling changes (e.g., no longer working in an early childhood setting, administration no longer afforded time to attend PD) that made them no longer eligible to participate. Attrition rates were not differential among the three study conditions—27%, 30%, and 31% for PD, PD+, and comparison conditions respectively; χ^2 (2, $N = 760$) = 1.24, $p = .537$. The overall attrition rate is commensurate with typical turnover rates in early childhood (Hale-Jinks & Knopf, 2006; Rhodes & Huston, 2012), similar to rates in other early childhood PD studies (e.g., Jackson et al., 2006), and falls within What Works Clearinghouse standards (Institute of Education Sciences, 2014). In accordance with these standards, analyses also accounted for educators' baseline scores.

Most participants were lead educators in the classroom (66%), with some co-educators (11%) and assistant educators (11%; note that percentages may not sum to 100% due to missing/unreported data). The majority were female (98%), White (78%), and non-Hispanic/Latino (82%); 17% were Black and 1% were of other or multiple races. Educators averaged 43 years of age ($SD = 10.86$) and 11 years of early childhood teaching experience ($SD = 7.92$). Sixty-six percent held early childhood teaching certification of some kind (i.e., state teaching license). Educators taught in a range of early childhood programs, including private center-based childcare (23%), publicly funded center-based childcare (51%), and home-based childcare (3%); 408 different early childhood centers were represented. The vast majority taught in mixed-aged classrooms serving a combination of 3-, 4-, and/or 5-year-olds (83%), with some teaching single-aged classrooms (5%) and

Table 1. Frequencies for characteristics selected as potential moderators by condition.

Moderators		Condition			K-W	
		PD[a] Freq (%)	PD+[b] Freq (%)	Comparison[c] Freq (%)	*H*	*p*
Ed level	None	26 (14.3%)	26 (14.5%)	33 (19.0%)	1.67	.434
	AA	38 (20.9%)	42 (23.5%)	35 (20.1%)		
	BA	56 (30.8%)	43 (24.0%)	59 (33.9%)		
	MA	43 (23.6%)	47 (26.3%)	33 (19.0%)		
Certificate		117 (64.3%)	124 (69.3%)	111 (63.8%)	2.33	.313
Head Start		62 (34.1%)	63 (35.2%)	68 (39.1%)	0.51	.774
ECSE		46 (25.3%)	38 (21.2%)	42 (24.1%)	1.31	.521
Center location	Rural	69 (37.9%)	60 (33.5%)	50 (28.7%)	5.89	.052
	Suburb	56 (30.8%)	53 (29.6%)	55 (31.6%)		
	Urban	43 (23.6%)	48 (26.8%)	60 (34.5%)		

Notes. PD = language and literacy professional development condition; PD+ = language and literacy professional development plus coaching condition; Percentages (%) are based on all responses, including those with missing data on the moderator variables; freq = frequency; K-W = Kruskal-Wallis test; Ed level = education level operationalized as highest degree earned. None = no college degree; AA = two-year degree; BA = bachelor's degree; MA = graduate degree; Certificate = educator had a state certificate. Head Start = educator taught in a Head Start program. ECSE = early childhood special education.
[a]n = 182 (30, 59, 40, and 53 in cohorts 1–4, respectively). [b]n = 179 (27, 54, 44, 54 in cohorts 1–4, respectively). [c]n= 174 (28, 54, 35, 57 in cohorts 1–4, respectively).

some teaching mixed-aged classrooms that also included younger children (5%). Many taught in Head Start programs (36%), given that this comprises the majority of publicly funded programs in the state. Most educators reported using Creative Curriculum (57%) or High/Scope (10%), with 6% not using any particular curriculum. Additional descriptive information is presented in Table 1. Notably, educators did not significantly differ across conditions on any reported characteristics.

Procedures

Educators were randomly assigned to one of three conditions: language and literacy PD (PD), language and literacy PD plus coaching component (PD+), or PD on an alternative topic (comparison). Assignments were made using batch random assignment restricted to force equal sample sizes within each cohort and a computerized random number generator (Shadish, Cook, & Campbell, 2002). Educators participated in PD between September and January; coaching spanned the full academic year. For study purposes, educators reported their knowledge and beliefs via questionnaires and participated in classroom observations in the fall (baseline, T1), winter (T2), and spring (T3) of the PD year as well as in the fall of the subsequent year (T4).

PD

All PD was funded by the state department of education, contracted to ecQ-net, and provided at no cost to educators. ecQ-net commissioned faculty at Ohio higher education institutions with expertise in emergent literacy, math, science, or social studies to provide PD content and delivery recommendations. The coaching model was developed collaboratively by ecQ-net staff and the faculty creating the literacy modules to ensure alignment. PD was delivered by facilitators hired by the state department of education and trained by either ecQ-net or faculty involved in PD creation. All facilitators were content specialists and held

at least a master's degree in a relevant field; language and literacy facilitators, in particular, were deemed Early Language and Literacy Specialists. All PD was offered in various locations across the state during all study years, with these selected to be geographically accessible to all participants. PD for all three study conditions was offered simultaneously in each selected location and attended by study participants as well as other educators who did not elect to participate in the study.

Language and Literacy PD. Educators assigned to the PD and PD+ conditions experienced 30 hrs of language and literacy PD delivered in a workshop format; all PD and PD+ educators attended the same workshops facilitated by the same Early Language and Literacy Specialists. The PD was intended to improve early childhood educators' knowledge, beliefs, and practices, with an ultimate goal of supporting young children's emergent literacy development through use of evidence-based practices and intentional teaching strategies. Content was derived from the research literature, aligned with state early learning standards, and covered five major domains related to providing high-quality language and literacy learning experiences: (a) Environment, (b) Play, (c) Oral Language, (d) Early Reading, and (e) Early Writing. Each domain was covered in two 3-hr sessions over one full or two half days with manualized content to facilitate consistent delivery across the state; notably, although domain-specific content was emphasized during these two sessions, content was also revisited in other sessions to provide an integrated PD experience (e.g., physical literacy environment information was introduced during Environment sessions but also discussed in Play sessions with respect to literacy props). All sessions followed a similar three-part format: (a) *Explorations*, during which facilitators introduced the session's content and goals and worked with educators to explore their existing knowledge, beliefs, and practices relevant to session content; (b) *Implications and Demonstrations*, during which facilitators provided specific content knowledge, introduced particular language and literacy practices as supported by examples and demonstrations, and provided opportunities for educators to try new activities and practices (see online supplementary material for sample practices); and (c) *Connections to Teaching and Learning*, during which facilitators summarized the session's content and goals, discussed ways of integrating new content and practices into classroom practices, and explained Into Practice activities. In the latter, educators made commitments to implement a specific new practice in their own classrooms, document their experiences, and share these experiences in subsequent PD sessions. In these ways, the PD not only emphasized building early childhood educators' language and literacy content knowledge but also the ways in which this knowledge could be applied to instruction.

Coaching. In addition to the language and literacy PD above, educators assigned to the PD+ condition also experienced coaching designed to facilitate integration of workshop content into their classroom practice. Educators were to receive a minimum of 4 to 6 hrs of coaching per month; coaching was to begin simultaneously with the language and literacy workshops and continue for the duration of the academic year. The coaching was aligned with the workshops and focused on the same five domains. Coaches were experienced early childhood educators who volunteered to participate in the state's coaching program, also implemented by ecQ-net. Each worked under an Early Language and Literacy Specialist, from whom they received an additional 24 hrs of coaching training plus ongoing support. Coaches also received a coaching manual outlining the structure and content of the coaching

program. Coaching was designed to follow a cycle involving goal setting, planning, observation, and feedback and reflection. Educators worked with their coaches to determine particular areas to improve and planned ways to accomplish these goals. Coaches observed educators implementing new practices, followed by critical reflection and coach feedback on implementation. The coaching cycle then repeated with subsequent goal setting. Educators completed portfolios documenting these cycles and their progress, and coaches completed logs documenting coaching activities (Schachter, Weber-Mayrer, Piasta, & O'Connell, 2015).

Comparison. A variety of state PD opportunities were also offered by ecQ-net on other topics, and educators randomly assigned to the comparison condition experienced PD on these alternative topics to provide a rigorous counterfactual and decrease threats to validity (Shadish et al., 2002). In 2010–2011, educators in the comparison condition completed two 12-hr PD offerings on math, science, or social studies, which served to closely equate educators on the total amount of PD experienced. The format and delivery of these offerings were analogous to that of the language and literacy PD. In 2012–2013 ecQ-net replaced these offerings with a new 30-hr PD opportunity that directly paralleled the language and literacy PD, focused on young children's cognition, with math, science, and social studies content embedded. The comparison PD did not emphasize any content covered in the language and literacy PD.[1]

Outcome Measures

Outcome measures were carefully selected based on the stated intent that the PD would improve early childhood educators' knowledge, beliefs, and practices, all of which are commonly targeted as PD outcomes (Schachter, 2015). Researchers met with representatives from the state department of education and ecQ-net to select measures that aligned with PD content and/or represented important intended outcomes (e.g., general benefits to instructional quality). To the extent possible, measures had existing psychometric information from use in previous empirical studies of early childhood PD.

Knowledge. Three measures assessed educators' language and literacy knowledge. The first was a *proximal knowledge* measure constructed by ecQ-net. Eighteen multiple-choice items were developed to assess knowledge of key content from the PD (e.g., "Repeated reading enriches children's: [a] comprehension, [b] ability to see and hear how written language works, [c] vocabulary development, [d] all of the above"); these items directly aligned with content expressly included in the PD and represented essential knowledge to be gained from the PD as determined by ecQ-net. Findings for this measure should be interpreted with caution given low internal consistency (Cronbach's $\alpha = .45$), perhaps due to covering content across all five domains of the PD. The second was a measure of language and literacy *disciplinary knowledge* adapted for early childhood educators by Cunningham and colleagues (Cunningham et al., 2009) from the widely used Moats (1994) survey. The measure includes 19 multiple-choice and short-answer items and assesses educators' understanding of oral

[1]For purposes of this study, ecQ-net attempted to collect survey data from PD facilitators and participants to document the extent to which PD was implemented as intended across the three conditions. Response rates on these measures were highly variable (less than 50% for facilitators and less than 25% for educators), and data exhibited strong negative skew and limited variability. For these reasons, implementation data were not considered in analyses of educator outcomes.

and written language structures relevant to young children's language and literacy development (e.g., counting syllables, manipulating phonemes). Internal consistency for the current sample was .77. The third was a measure of educators' knowledge of language and literacy as it informs practice (*knowledge for practice*) developed by Neuman and Cunningham (2009). The measure includes 50 multiple-choice and 20 true/false items that address eight core language and literacy competencies (i.e., letter knowledge, literacy assessments, literacy curriculum, oral language, comprehension, parental role in language and literacy development, phonological awareness, print conventions, strategies for working with second-language learners) and additional items concerning child development and best practices as aligned with National Association for the Education of Young Children standards. Items were reviewed for face validity by early childhood literacy experts and subjected to initial item analysis (Neuman & Cunningham, 2009); internal consistency in the current sample was .76. For each of the three knowledge measures, the number of items answered correctly was summed to create the raw scores used in analyses.

Beliefs. Two aspects of educators' beliefs related to language and literacy were assessed. Educators' feelings of *self-efficacy* were assessed via a scale specific to supporting young children's language and literacy development (Justice et al., 2008) as adapted from the Bandura Teacher Self-Efficacy Scale (1997). Educators rate their perceived efficacy on nine items (e.g., "How much can you do to promote children's alphabet knowledge?") using a 5-point Likert scale. Psychometric work indicated that five of the nine items comprised a reliable and unidimensional self-efficacy factor (Arthur, McCormick, & Bovaird, 2012). In the present study, educators' ratings for these five items were averaged to create the score used in analyses; higher scores indicate greater feelings of efficacy. Internal consistency was .94. Educators' beliefs about language and literacy development and instruction (*language and literacy beliefs*) were assessed using the Preschool Teacher Literacy Beliefs Questionnaire (Hindman & Wasik, 2008). This measure uses 30 items to represent educators' beliefs regarding code-focused skills, oral language and vocabulary, shared reading, and writing within the early childhood context (e.g., "As a teacher I believe preschool children should learn to identify beginning and ending sounds in words"). Educators rate the extent to which they agree with each item on a 5-point Likert scale. Ratings were averaged to create an overall score used in analyses; higher scores reflect beliefs more aligned with research and best practices. Internal consistency was .78.

Practice. Two aspects of practice were assessed via one-day classroom observations conducted at baseline, T2, T3, and T4. The observation was conducted on a day that the educator indicated would be representative of typical instruction; educators confirmed representativeness at the end of the observation by rating "How typical of a day was today?" on a scale of 1 (not typical at all) to 5 (very typical; $M = 4.26$, $SD = 0.63$). The *quality of the physical literacy environment* was measured via the 27-item Classroom Literacy Observation Profile (CLOP; McGinty & Sofka, 2009). The CLOP was completed by research staff who demonstrated greater than 90% item-level agreement on three consecutive administrations with a master observer. It measures the availability and use of literacy-related materials (e.g., various genres of books), environmental print (e.g., child-generated writing samples, labels), and literacy-related technology (e.g., audio center) in the classroom. Each item is rated with respect to presence (1 = yes, 0 = no) or frequency, with scales for the latter item-dependent

(e.g., "How many word/letter puzzles are accessible?" scored as 0 = none, 1 = one to three, 2 = four to six, and 3 = seven or more). Frequency scales were rescored to range from 0 to 1 by dividing the item score by the number of scale categories in order to have all CLOP items contribute equally to the summed composite score used in analyses. Internal consistency was .78.

Baseline, T2, T3, and T4 classroom observations were also video-recorded for purposes of measuring the *quality of classroom instruction*. Recordings captured all classroom activities taking place during classroom instructional time, as defined by the educator (M = 1 hr 33 min, SD = 33 min). Observations were parsed into 20-minute segments (i.e., "cycles"), and research staff randomly selected three 20-minute cycles for coding using the Classroom Assessment Scoring System, Pre-K version (CLASS; Pianta, La Paro, & Hamre, 2006). CLASS assesses educator–child interactions to measure three domains of quality, namely instructional support, emotional support, and organizational support. For our purposes, the instructional support domain served as the outcome of interest, as the three dimensions comprising the CLASS instructional support domain (i.e., language modeling, concept development, and quality of feedback) aligned well with PD content (e.g., having extended conversations with children, use of language-facilitating strategies and diverse vocabulary). Moreover, research indicates associations between CLASS instructional support and children's early literacy skill development (Hamre, Hatfield, Pianta, & Jamil, 2014), and this has served as an outcome of interest in other studies of language- and literacy-focused PD (Mashburn, Downer, Hamre, Justice, & Pianta, 2010; Yoshikawa et al., 2015). Finally, CLASS instructional support was deemed an important outcome of interest by the state, given its use as a metric for assessing early childhood classrooms participating in the state's Quality Rating Improvement System. For each cycle, trained research staff used a 7-point Likert scale to rate each dimension, with higher scores representing better quality. The instructional support score was calculated as the average across cycles and dimensions. Training consisted of completing a CLASS workshop led by a CLASS-certified instructor and achieving the scoring reliability standards set by CLASS developers (i.e., 80% scoring agreement on five gold-standard CLASS training videos that includes agreement for any given dimension on at least two out of five videos; CLASS procedures define agreement as scoring within one point of the master-coded score). Research staff completed annual CLASS recertification and demonstrated 88% agreement for the instructional support domain across the 20% of cycles randomly selected for double coding across the baseline, T2, T3, and T4 observations. Internal consistency was .85.

Characteristics Selected a Priori as Potential Moderators

Prior to beginning the study, select educator and classroom characteristics were chosen as potential moderators to determine whether effects would be similar across the diverse array of educators participating in the state-sponsored PD. Characteristics were self-reported by educators on a questionnaire administered at fall baseline; frequencies by condition are indicated in Table 1. *Educators' highest level of education* was represented in moderator analyses via a series of dummy codes representing no college degree, two-year/associate's degree, four-year/bachelor's degree, or graduate degree. *Certification* was represented as a dummy code indicating whether or not the educator held stated-issued teaching certification. *Classroom type* was originally intended to represent three categories (i.e., private center-based, publicly funded center-based, home based), but few home-based educators participated, and

most classrooms associated with publicly funded centers were affiliated with Head Start. Thus, we reconceptualized this characteristic to reflect whether the educator taught in a Head Start classroom, represented as a dummy code in moderator analyses. *Early Childhood Special Education* (ECSE) was represented as a dummy code based on whether the educator indicated teaching in an ECSE classroom or was a special educator teaching within a preschool special education program. Finally, the *location* in which educators taught was represented as rural, suburban, or urban via dummy codes.

Results

Descriptive statistics for all outcomes are presented in Table 2. With alpha set at .05, conditions showed one significant baseline difference: Educators in the PD condition exhibited higher quality classroom environments than those in the comparison condition ($d = .25$; $p = .031$). Prior to beginning analyses, we investigated patterns of missing data for all variables included in analyses. The proportion of missing data ranged from 3.7% to 34.1% across the four time points, with the greatest amounts of missing data on the outcome variables

Table 2. Descriptive statistics for educator outcomes.

		Condition			ANOVA	
Outcome measures		PD *M (SD)*	PD+ *M (SD)*	Comparison *M (SD)*	*F*	*p*
Proximal knowledge	Baseline	12.85 (2.17)	13.29 (2.00)	12.78 (2.11)	2.99	.051
	T2	13.38 (2.14)	13.21 (2.33)	12.96 (2.23)		
	T3	13.05 (1.94)	13.19 (2.09)	12.85 (2.26)		
	T4	13.38 (2.13)	13.31 (2.07)	13.03 (2.34)		
Disciplinary knowledge	Baseline	12.48 (3.28)	12.60 (3.14)	12.13 (3.16)	0.97	.379
	T2	12.90 (3.00)	12.76 (3.24)	12.28 (3.30)		
	T3	12.79 (3.12)	13.07 (2.92)	12.38 (3.54)		
	T4	13.03 (3.33)	13.08 (3.15)	12.17 (3.63)		
Knowledge for practice	Baseline	44.67 (6.54)	44.71 (6.92)	44.86 (6.06)	0.04	.958
	T2	45.38 (6.62)	45.01 (6.94)	44.64 (6.65)		
	T3	46.45 (6.66)	44.51 (7.30)	44.86 (6.74)		
	T4	44.93 (7.64)	45.60 (6.62)	44.38 (7.19)		
Self-efficacy	Baseline	3.26 (0.57)	3.29 (0.61)	3.28 (0.57)	0.15	.862
	T2	3.31 (0.53)	3.38 (0.57)	3.36 (0.55)		
	T3	3.29 (0.58)	3.36 (0.58)	3.41 (0.52)		
	T4	3.30 (0.62)	3.31 (0.57)	3.29 (0.56)		
Language & literacy beliefs	Baseline	2.41 (0.25)	2.45 (0.25)	2.45 (0.24)	1.06	.346
	T2	2.47 (0.25)	2.44 (0.28)	2.47 (0.25)		
	T3	2.47 (0.30)	2.46 (0.30)	2.46 (0.27)		
	T4	2.47 (0.25)	2.47 (0.28)	2.48 (0.25)		
Quality of literacy environment	Baseline	14.49 (3.30)	13.81 (3.49)	13.48 (3.53)	3.34	.036
	T2	14.78 (3.65)	14.44 (3.89)	13.85 (3.80)		
	T3	14.55 (3.40)	14.20 (3.83)	13.73 (3.80)		
	T4	14.28 (3.04)	13.71 (3.62)	13.67 (3.23)		
Quality of classroom instruction	Baseline	2.25 (0.66)	2.27 (0.63)	2.24 (0.66)	0.13	.880
	T2	2.10 (0.60)	2.07 (0.56)	2.03 (0.63)		
	T3	2.15 (0.60)	2.09 (0.59)	2.09 (0.60)		
	T4	2.20 (0.58)	2.14 (0.59)	2.12 (0.59)		

Notes. PD = language and literacy professional development condition; PD+ = language and literacy professional development plus coaching condition; Baseline = fall/T1, T2 = winter, T3 = spring, T4 = fall follow-up. Maximum possible scores for each outcome are as follows: 18 for proximal knowledge, 19 for disciplinary knowledge, 70 for knowledge for practice; 4 for self-efficacy, 4 for language and literacy beliefs, 27 for quality of literacy environment, and 7 for quality of classroom instruction.

measured at T4 (30.6% to 34.1%) and the variable reflecting whether educators held a state certificate (30.6%). Separate variance *t* tests indicated that there was a systematic association between missingness and other variables included in the models, suggesting the missing data could not be treated as missing completely at random (Enders, 2010; Graham, 2012). Thus, the issue of missing data was addressed using full information maximum likelihood for all analyses. Analyses were conducted using SPSS v22 and HLM v7.

Impact of Professional Development

To determine the effects of the language and literacy PD on educators' knowledge, beliefs, and classroom practice, we used individual growth models for each of the seven outcome measures. We began by fitting unconditional growth models to capture the patterns of change observed over time (O'Connell, Logan, Pentimonti, & McCoach, 2013). As shown in Figure 1, a similar pattern was observed for most outcome measures, in that there was some evidence of change between baseline fall and spring (T3) of the PD year, and a slight drop-off by fall follow-up (T4). Thus, a series of nested quadratic models of growth were fit to the data. In these models, patterns of change over time are captured by the linear and quadratic components, with the linear component representing instantaneous rate of change at the intercept point, and the quadratic component representing degree of curvature. Based on AIC criteria and chi-square difference tests, nested models excluding the quadratic component provided significantly worse fit relative to models including that component; thus, the quadratic component was retained in all models. Chi-square tests of the variance components indicated that the intercepts and linear components should be allowed to vary at random, with the quadratic component fixed. Post hoc power analyses based on the parameters of these growth models indicated sufficient power ($> .75$) to detect small changes between conditions (less than two-tenths of a point).

To determine whether PD condition could account for patterns of change over time, we next used a set of dummy codes to capture differences in the randomly varying intercepts and linear components for each of the two language and literacy PD conditions (PD, PD+) versus the comparison PD. Models with time centered at baseline were constructed first and were recentered in subsequent analyses at T3 and T4 to examine post-PD differences. Consequently, our interpretation is focused on the effects observed in the baseline models and differences in the intercepts in the recentered models. Results are shown in Table 3 for knowledge outcomes, Table 4 for belief outcomes, and Table 5 for practice outcomes. Due to the large number of models and variables that we examined, we used .01 as the a priori alpha level to determine statistical significance for all models.

For each of the three knowledge outcomes, there were no significant effects of PD on patterns of growth (Table 3); thus, change did not vary by PD received. In fact, the general pattern of change was flat (Figure 1). The quadratic component was not significantly different from zero (β_{20}), and effects for the linear component for the type of PD received were also not significantly different from zero (β_{10}, β_{11}, β_{12}, respectively). For all three outcomes, the intercepts were significantly different from zero at baseline, T3, and T4 (these represent estimated means for the comparison condition at baseline, T3, and T4). Participation in the language and literacy PD tended to yield larger means than the comparison condition over time (β_{01}, β_{02}), but these increases did not attain statistical significance at $\alpha = .01$.

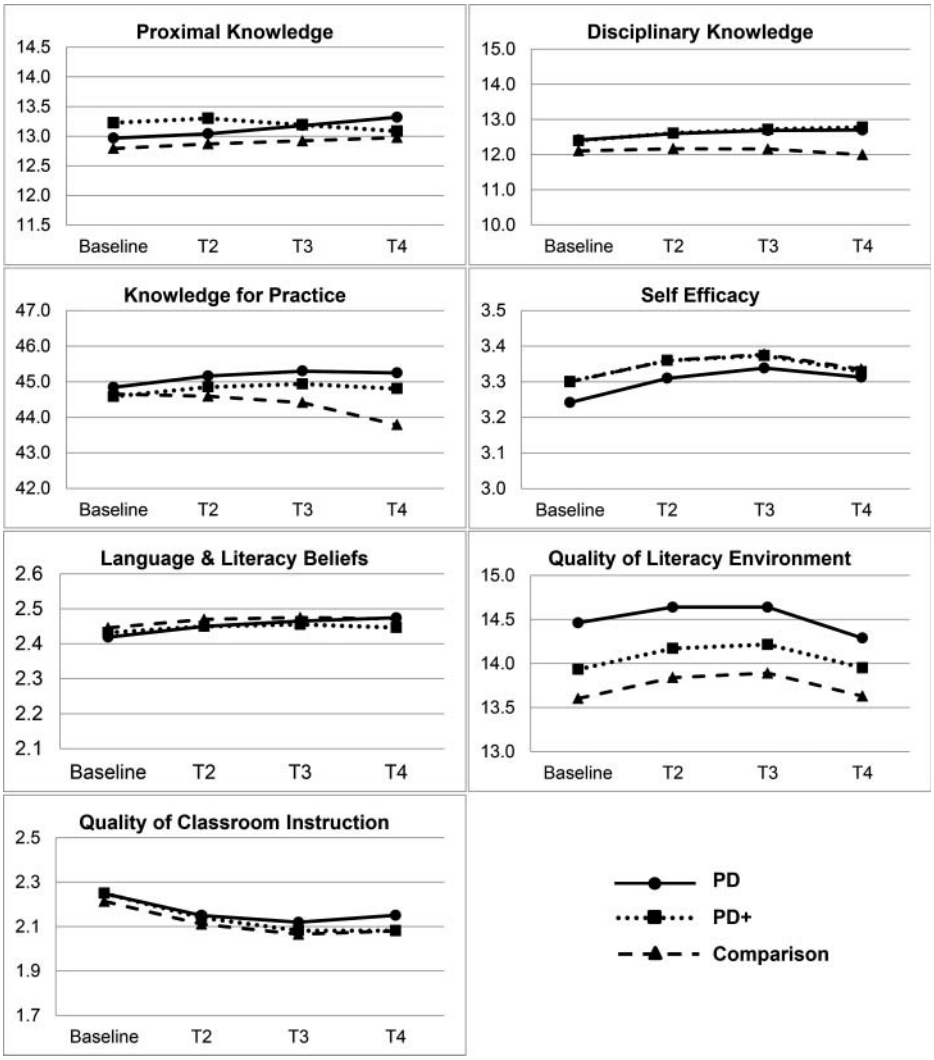

Figure 1. Growth curves for the average educator by condition for each outcome estimated using coefficients from quadratic models at baseline with fixed quadratic component added to the random intercept and random linear component. Baseline = fall/T1; T2 = winter; T3 = spring; T4 = fall follow-up. PD = language and literacy professional development condition; PD+ = language and literacy professional development plus coaching condition; Comparison = nonliteracy professional development.

For self-efficacy and language and literacy beliefs, there were no significant effects of PD on patterns of growth over time (Table 4). The general pattern of change for both outcomes was quadratic (β_{20}), with self-efficacy and language and literacy beliefs increasing through T3 and declining slightly by T4 (Figure 1). However, the values for the estimated means in the recentered models at T3 and T4 show little substantive difference from baseline, indicating that the actual amount of change was minimal.

For the two practice outcomes, quality of literacy environment and quality of classroom instruction, a similar quadratic pattern was observed (Figure 1) with no significant effects of PD

Table 3. Conditional quadratic models for educator knowledge outcomes, centered at baseline, T3, and T4.

	Proximal knowledge			Disciplinary knowledge			Knowledge for practice		
Parameters	Baseline	T3	T4	Baseline	T3	T4	Baseline	T3	T4
Intercept π_0									
Intercept β_{00}	**12.792**	**12.873**	**12.962**	**12.109**	**12.068**	**11.911**	**44.655**	**44.209**	**43.483**
PD β_{01}	0.157	0.256	0.344	0.308	0.532	0.754	0.185	0.930	1.635
PD+ β_{02}	0.470	0.255	0.021	0.290	0.572	0.851	−0.072	0.558	1.137
Linear π_1									
Intercept β_{10}	0.029	−0.096	−0.020	0.034	−0.039	−0.062	0.007	−0.089	−0.229[a]
PD β_{11}	0.019	0.019	0.015	0.040	0.044	0.038	0.127	0.124	0.123
PD+ β_{12}	−0.036	−0.036	−0.041	0.049	0.057	0.049	0.109	0.107	0.101
Quadratic π_2									
Intercept β_{20}	−0.001	0.019	0.003	−0.004	0.002	0.004	−0.009	−0.002	0.011
Variance									
Intercept τ_{00}	**1.936**	**2.314**	**2.596**	**7.321**	**8.258**	**9.364**	**30.823**	**36.660**	**45.402**
Linear τ_{11}	0.001	0.001	0.001	0.002	0.001	0.002	**0.055**	**0.062**	**0.051**
Residual σ^2	**2.375**	**2.390**	**2.336**	**2.730**	**2.757**	**2.735**	**11.355**	**11.207**	**11.409**
Reliability									
Intercept β_0	.531			.783			.775		
Linear β_1	.017			.211			.049		

Notes. Reliabilities estimated from unconditional model. PD = language and literacy professional development condition; PD+ = language and literacy professional development plus coaching condition. Baseline = fall/T1; T3 = spring; T4 = fall follow-up. Linear component = number of months from baseline. Quadratic component = number of months since baseline squared.
[a]$p < .05$; Boldface $p < .01$.

(Table 5). As with the self-efficacy and language and literacy beliefs outcomes, the estimated means in the recentered models at T3 and T4 show little substantive difference from baseline.

Outcomes for the two language and literacy PD conditions, PD and PD+, were compared to determine if there was any advantage of coupling PD with a coaching component. Models

Table 4. Conditional quadratic models for educator beliefs, centered at baseline, T3, and T4.

	Self-efficacy			Language and literacy beliefs		
Parameters	Baseline	T3	T4	Baseline	T3	T4
Intercept π_0						
Intercept β_{00}	**3.300**	**3.334**	**3.308**	**2.446**	**2.464**	**2.466**
PD β_{01}	−0.058	−0.040	−0.028	−0.027	−0.009	0.007
PD+ β_{02}	0.001	−0.005	−0.017	−0.014	−0.021	−0.028
Linear π_1						
Intercept β_{10}	**0.027**	−0.012	**−0.024**	**0.009**	−0.000	−0.003
PD β_{11}	0.004	0.002	0.002	0.003	0.003	0.003
PD+ β_{12}	−0.001	−0.001	−0.002	−0.001	−0.002	−0.001
Quadratic π_2						
Intercept β_{20}	**−0.002**	0.002	**0.024**	**−0.001**	0.000	0.000
Variance						
Intercept τ_{00}	**0.170**	**0.164**	**0.206**	**0.040**	**0.049**	**0.059**
Linear τ_{11}	**0.001**	**0.001**	**0.001**	0.000	0.000	0.000
Residual σ^2	**0.149**	**0.151**	**0.149**	**0.021**	**0.022**	**0.022**
Reliability						
Intercept β_0	.609			.718		
Linear β_1	.194			.039		

Notes. Reliabilities estimated from unconditional model. PD = language and literacy professional development condition; PD+ = language and literacy professional development plus coaching condition. Baseline = fall/T1; T3 = spring; T4 = fall follow-up. Linear component = number of months from baseline. Quadratic component = number of months since baseline squared. Boldface $p < .01$.

Table 5. Conditional quadratic models for practice outcomes, centered at baseline, T3, and T4.

	Quality of literacy environment			Quality of classroom instruction		
Parameters	Baseline	T3	T4	Baseline	T3	T4
Intercept π_0						
Intercept β_{00}	**13.602**	**13.637**	**13.425**	**2.213**	**2.130**	**2.138**
PD β_{01}	0.859[a]	0.737[a]	0.660	0.035	0.053	0.068
PD+ β_{02}	0.331	0.307	0.313	0.036	0.021	−0.016
Linear π_1						
Intercept β_{10}	0.113[a]	**−0.275**	**−0.170**	**−0.044**	0.040[a]	**0.035**
PD β_{11}	−0.020	−0.019	−0.016	0.004	0.002	0.003
PD+ β_{12}	−0.001	−0.000	−0.003	−0.004	−0.005	−0.005
Quadratic π_2						
Intercept β_{20}	**−0.011**	**0.046**	**0.014**	**0.003**	**−0.008**	**−0.004**
Variance						
Intercept τ_{00}	**10.184**	**8.157**	**7.900**	**0.132**	**0.090**	**0.100**
Linear τ_{11}	**0.027**	**0.024**	**0.026**	0.001[a]	0.001	0.001
Residual σ^2	**3.820**	**3.855**	**3.816**	**0.262**	**0.267**	**0.262**
Reliability						
Intercept β_0	.410			.788		
Linear β_1	.143			.268		

Notes. Reliabilities estimated from unconditional model. PD = language and literacy professional development condition; PD+ = language and literacy professional development plus coaching condition. Baseline = fall/T1; T3 = spring; T4 = fall follow-up. Linear component = number of months from baseline. Quadratic component = number of months since baseline squared.
[a]$p < .05$; Boldface $p < .01$.

akin to all above analyses were estimated, with PD+ (rather than comparison) as the reference group. No significant differences were detected between PD and PD+ on any of the outcomes.

Moderator Analyses

To investigate potential moderation effects of our five a priori educator and classroom characteristics, we examined growth curve models that included each focal characteristic and its interaction with PD conditions on both the intercept and linear component. Models were created separately for each characteristic and each of the seven outcomes.

None of the focal characteristics or their interactions with the PD dummy codes significantly predicted the linear components of growth curve models. Thus, these effects were removed from models. The final series of models testing moderation included the two dummy-coded PD variables on both intercepts and linear components along with each of the focal characteristics and their interactions with PD as predictors of intercepts only. When interactions were not significant, these were removed, and the model was rerun to estimate main effects.

Only one potential moderator, educators' highest degree earned, significantly interacted with PD and only when knowledge for practice served as the outcome ($t_{\text{BA x PD}} = 2.70$, $p = .007$; $t_{\text{MA x PD}} = 2.54$, $p = .001$; $t_{\text{AA x PD}} = 1.75$, *ns*). Probing this interaction yielded the pattern exhibited in Figure 2, whereby educators in the PD condition with a BA had higher means for knowledge for practice than those in the comparison condition at baseline, T3, and T4; the same was true when contrasting educators in the PD condition with an MA and those in the comparison condition. Although the interaction effects were in the same positive direction for the PD+ condition at baseline, T3, and T4 (Figure 2), none of these

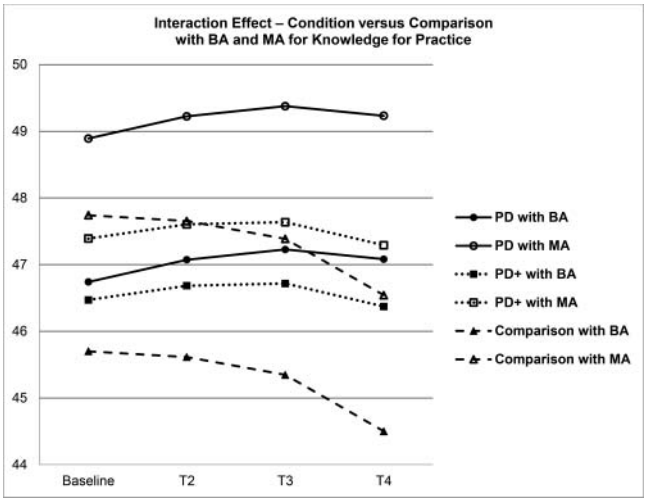

Figure 2. Growth curves of estimated means across time for educators with a bachelor's degree (BA) or graduate degree (MA) in the PD condition relative to educators in the comparison group with a BA or MA degree on the measure of knowledge for practice. Also shown are growth curves for educators with a BA or MA in the PD+ condition, though these interactions were not significantly different from the comparison. Baseline = fall/T1; T2 = winter; T3 = spring; T4 = fall follow-up. PD = language and literacy professional development condition; PD+ = language and literacy professional development plus coaching condition; Comparison = nonliteracy professional development.

interaction terms were statistically significant. Notably, given the consistent difference in level of knowledge for practice from baseline through T4, favoring educators assigned to the PD condition with a BA or MA, these results do not indicate that effects of PD were moderated by educators' education level.

For all other outcomes, there were no significant interactions between characteristics and the type of PD received, although many characteristics showed main effects. In general, educators with higher levels of education (BA or MA) tended to have higher levels of knowledge, self-efficacy, language and literacy beliefs, and quality of the literacy environment. Educators possessing a certificate to teach young children tended to have higher levels of knowledge, language and literacy beliefs, and quality of the literacy environment. Educators teaching in Head Start classrooms tended to have lower levels of disciplinary knowledge and self-efficacy but provided higher quality literacy environments. Educators in ECSE classrooms exhibited higher levels on all outcomes except the quality of classroom instruction. Main effects of location were more mixed but, in general, educators in urban settings tended to have lower knowledge, lower language and literacy beliefs, and lower quality of literacy environments. Full moderator results are available via online supplementary materials.

Discussion

The purpose of the present study was to examine the effectiveness of one state's language and literacy PD on an array of educator outcomes. An advantage of the study was its randomized controlled trial design to examine the impacts of PD as typically experienced by educators outside of researcher-controlled settings and thus exemplifying the types of PD that most educators are utilizing. Unlike previous studies in which PD has been created,

implemented, and monitored by researchers, this PD was developed and implemented by contractors of the state department of education as an ongoing effort to improve educators' knowledge, beliefs, and practices to support the early literacy development of young children. We, as researchers, had no personal investment in the PD, nor did we help to design it. Indeed, the effectiveness trial was designed as a rigorous, independent, and objective evaluation of long-standing, real-world PD that represented a significant investment of state funds. In part, this evaluation tests the return on this investment. Given its scale, the study had sufficient statistical power for detecting even small changes in educator outcomes and examining potential moderators of PD effects. Importantly, this study provides key insights to stakeholders about the extent to which this PD was successful in realizing intended impacts as implemented in the real-world early childhood context.

Results of this effectiveness trial showed no impacts of state-sponsored language and literacy PD, whether offered with or without a coaching component, on intended educator outcomes. In fact, most outcomes showed no or very little evidence of change over the 18-month period during which educators were involved in the study. Moreover, results showed that the lack of impacts was similar across educators regardless of education levels, state certification, affiliation with Head Start, ECSE status, or location of their classrooms. Before discussing these findings in greater depth, it is important to note that many of the data patterns evident in the current study mirror those reported in the extant literature. For example, educators' scores on outcome measures used in extant studies are similar to previously reported findings (Burchinal et al., 2008; Cunningham et al., 2009; Hindman & Wasik, 2008; Neuman & Cunningham, 2009; Pianta et al., 2005), as are main effects for a priori-selected characteristics examined as moderators (e.g., educators with degrees tended to score higher on measures of knowledge [Goldschmidt & Phelps, 2010; Kelcey, 2011] and classroom quality [Early et al., 2007; LoCasale-Crouch et al., 2007]). The demographic and other background characteristics of the sample are also fairly typical of the early childhood educator workforce. This is true for gender and race (Clifford et al., 2005; Institute of Medicine and National Research Council, 2012) as well as the use of curricula (Clifford et al., 2005; Lonigan et al., 2011). There were, however, more educators in our sample with a bachelor's degree or higher and with state certifications than are typically reported (Institute of Medicine and National Research Council, 2012), perhaps due to the large number of participants affiliated with Head Start, which has implemented policy changes recently regarding educator credentials. Overall, these similarities to previous reports attest to the validity of our outcome data and also indicate that our sample is fairly representative of educators who receive PD in real-world, state-sponsored contexts.

Our results concerning the impact of PD, or lack thereof, can also be interpreted within the context of previous research. Whereas several researchers have found positive effects for more controlled, researcher-led PD (Assel et al., 2007; Hamre et al., 2010; Landry et al., 2011; Lonigan et al., 2011) and closely monitored coaching provided by experts (Hindman & Wasik, 2012; Milburn et al., 2015; Powell, Diamond, Burchinal, & Koehler, 2010; Weiland & Yoshikawa, 2013), others have found no or limited effects of PD. For example, Neuman and Wright (2010) found no impacts on educator outcomes for PD based on coursework only, and coursework plus coaching PD impacted the quality of educators' classroom environments but not educators' knowledge. In their large-scale study of PD plus coaching offered across four states, Landry et al. (2009) did not find impacts of PD on all measures of educator practice, and, similarly, in a study of PD targeting educators' use of language

supports within the classroom, Piasta et al. (2012) showed effects on only one of two targeted sets of supports. Moreover, other researchers have also found no added value of coaching (Assel et al., 2007; Jackson et al., 2006; Lonigan et al., 2011). Additionally, with respect to real-world PD, state and district offerings have been reported as ineffective (Horizon Research, 2002), and educators often report that PD reinforces rather than reforms their practice (Hill, 2009). Thus, although disappointing, the lack of impacts is not altogether unexpected based on the extant literature, coupled with the fact that this state-sponsored PD had no previous evidence of efficacy and was implemented with limited checks on implementation. Below, we offer two possible explanations for the lack of impacts.

Lack of PD Impacts on Educator Outcomes

There are a number of reasons that the state would have anticipated the PD to have positive impacts on educator outcomes, with the design of the PD perhaps the most compelling. The content of the PD delivered in this study was developed by early childhood experts across the state and assumed to reflect current and accurate knowledge of best practices in language and literacy instruction. Moreover, the PD was designed to be intensive, embedded, and ongoing to maximize the likelihood of affecting change in educators' knowledge, beliefs, and practice, consistent with best PD practices (Garet et al., 2001; Yoon, Duncan, Lee, Scarloss, & Shapley, 2007). For example, PD provided 30 hrs of coursework over multiple months, which exceeds the minimum recommendation of 14 hrs (Yoon et al., 2007), and was further supplemented by monthly coaching across the entire academic year for some participants. Moreover, both the coursework and coaching included components to link PD content to classroom practice.

Given these strengths in design of the PD, one explanation for the lack of impacts is that our measures were not well aligned with the content or intended outcomes of the PD. Construct validity is always an important design consideration (Borko, 2004; Desimone, 2009), and it is possible that the PD affected educator outcomes in ways that were not measured in the current study. For example, the PD may have resulted in more subtle changes in the quantity or quality of specific classroom language and literacy practices (e.g., providing opportunities to build both code- and meaning-focused emergent literacy skills associated with later reading success; utilizing evidence-based best practices such as print referencing or explicit vocabulary instruction; regularly engaging children in shared book reading; Piasta, 2016) than could be captured in the selected outcome measures. We must point out, however, that we took great care in selecting study measures, choosing these in collaboration with key stakeholders to reflect content targets of PD and the intended outcomes as represented in the theory of change evident in PD documentation. Further, with the exception of the proximal knowledge measure, one of the strengths of this study was our use of existing knowledge, beliefs, and practice measures that have been validated in previous research, both to ensure construct validity and to increase interpretability of our findings relative to available literature. To this end, the educators in our sample scored similarly on these measures compared to previous reports (Cunningham et al., 2009; Hindman & Wasik, 2008; Neuman & Cunningham, 2009; Pianta et al., 2005), and baseline scores were sufficiently low to allow for growth (i.e., results were not due to ceiling effects). We also utilized multiple measures to more fully represent our broad constructs of interest (knowledge, beliefs, practice) and capture potential changes as a result of PD. Yet, despite this attention to construct

validity, it remains a possibility that our measures were not nuanced enough to detail more specific changes in educators' outcomes. More generally, more sensitive and psychometrically strong knowledge, beliefs, and practice measures are needed, particularly those that can be flexibly used across multiple settings (Schachter, 2015) as well as those that have been designed to reliably capture longitudinal change. A related measurement limitation worth noting concerns the proximal knowledge measure developed by ecQ-net to directly align with PD content. This measure had high face validity but had not been used in previous research or undergone rigorous psychometric evaluation; although results for this measure paralleled those for the other knowledge measures, these results should be interpreted with caution particularly given the low internal consistency.

Alternatively, a second explanation for the lack of PD impacts is that the state-sponsored PD was ineffective in changing educators' knowledge, beliefs, or practice. Although reasons for this cannot be definitively identified, we can conjecture that aspects of PD content and format, as well as challenges inherent to implementing PD in real-world contexts, may have contributed to lack of effects. First, it may be that the content addressed within the PD was too broad to effect change; it is unclear to what depth content must be addressed in order to achieve effects, especially for practice outcomes (Landry et al., 2009; Neuman & Cunningham, 2009). Moreover, although the PD was manualized to ensure consistent content, facilitators had flexibility in how content was conveyed and linked with existing knowledge, beliefs, and practice in an effort to meet recommendations that PD be individualized for adult learners (Weber-Mayrer, Piasta, & Pelatti, 2015). This may have influenced exposure to content as well as uptake. Furthermore, despite attempts to ensure that the PD was embedded and connected to educators' practice, the manner in which the PD made these connections (e.g., reflection, Into Practice activities, coaching) may not have been sufficient. In general, there is little empirical evidence as to how to best achieve connections to practice or individualization within the design of PD (Peterson, Taylor, Burnham, & Schock, 2009; Wayne, Yoon, Zhu, Cronen, & Garet, 2008) and more research on these topics is necessary. Relatedly, although coaching was expected to provide additional embedded, ongoing, and practice-based support, it may not have realized these aims. We have emerging evidence that coaching interactions targeted a wide variety of topics beyond the practices emphasized in PD (Schachter et al., 2015), suggesting that coaching may not have been as tightly aligned with PD content as intended. Additionally, although most educators attended PD sessions at the level expected (attendance ranged from 80% to 100% [$M = 0.98$, $SD = 0.06$] based on the state database used to document compliance with PD regulations), exposure to coaching was much more variable; not all educators in the PD+ condition experienced 4 to 6 hrs of coaching per month (range of 1 to 78 total hrs [$M = 28.62$, $SD = 19.99$] based on logs submitted by coaches). Notably, ideal levels of PD exposure or intensity have not been established in the literature, and the nature of associations with educator outcomes, whether incremental or threshold-dependent, remains unclear (Weber-Mayrer et al., 2016). The quality of coaching also likely varied, given that coaches were volunteers who, anecdotally, differed in their experience levels, expertise, and amount of support received. The lack of information about coaches and the coaching process, although common in large-scale coaching studies, is a limitation that deserves greater attention in the future.

Overall, we cannot be sure that the PD was implemented as intended, such that all participating educators experienced high-quality PD. Examining the quality of the PD was not a main goal of this investigation, and quality of implementation is a difficult construct to measure, rarely accounted for in other intervention studies (Durlak & DuPre, 2008). Implementation is particularly challenging to assess within large-scale, real-world contexts. In the present study, we had hoped to examine multiple indicators of implementation fidelity, including ratings of quality, adherence, and participant responsiveness, as potential moderators or mediators of PD impacts (Desimone, 2009) but were restricted in measuring implementation, due to parameters of our partnership with the state and ecQ-net, limited resources, and the need to remain independent from implementation. When coupled with low response rates and limited variability, little information could be gained from analysis of implementation data. Thus, the issue of quality, in terms of maintenance, measurement, and links with outcomes, remains a possible explanation for lack of PD effects and is a crucial issue to consider in future PD endeavors and research.

Challenges and Importance of Effectiveness Trials

The need to examine effects of programs, policies, and practices within real-world contexts has received growing attention (Gottfredson et al., 2015; James-Burdumy et al., 2012; Sarama, Clements, Starkey, Klein, & Wakeley, 2008). Such studies remain relatively scarce in the education literature, however, perhaps due to the challenges inherent in effectiveness trials (Gottfredson et al., 2015). These include more heterogeneous samples, greater potential for attrition, variability in implementation fidelity, and the need to coordinate and standardize implementation and data collection across geographically dispersed sites.

Notwithstanding such challenges, effectiveness trials are critically important in understanding the success, or lack thereof, of educational approaches when adopted and implemented at large scale. Ideally, smaller, more controlled studies would have been conducted to demonstrate the efficacy of this specific PD model prior to the large state investments and being brought to scale. Such evidence, coupled with better implementation data, would help disentangle alternative explanations for the current null findings. Yet, it is not uncommon for policies or programs to be enacted with substantial cost prior to investigating their effectiveness (Kennedy, 2010; Rogers, 2014), and, regardless of scale, evidence that PD or other interventions can improve intended outcomes is critical for key stakeholders when considering whether to continue investments in educational practices (Chen, 2010). The operating costs of the current PD were over $500,000 annually. This estimate does not include costs associated with venues for PD offerings (provided in-kind by early childhood agencies and organizations), initial creation of PD content and design, coverage/travel for educators to attend PD sessions, or compensation for coaches, who were volunteers. This figure also does not include the hundreds of hours invested by early childhood educators participating in the PD. In the face of such investments, it is imperative that we understand how initiatives at the state and federal level are enacted and, as importantly, how effective they are in achieving their goals. The current results suggest the need to strongly consider the potential challenges inherent in large-scale implementation of PD and also the need to consult the research literature to discern evidence-based means of supporting

educator change when designing and enacting state- or district-wide PD models, such as those proposed in federal initiatives (e.g., Race to the Top) and other quality-improvement systems. For example, studies suggest that effective coaching may require highly qualified and well-trained expert coaches, a strong emphasis on practice, tailoring of content for individual educators, and use of particular evidence-based strategies linked to educator change (e.g., Bean, Draper, Hall, Vandermolen, & Zigmond, 2010; Carlisle & Berebitsky, 2011; Matsumura, Garnier, & Resnick, 2010; Sailors & Price, 2015). The extent to which such key components can be integrated and their integrity maintained within large-scale PD remains an open question and likely delineates whether such PD can actualize intended goals.

In sum, the results of the current study do not provide evidence that this particular state-sponsored PD was effective in improving educators' language and literacy knowledge, beliefs, or practice. Rather, results highlight a continued need to attend to issues of design, quality, and delivery of PD for early childhood educators, especially for large-scale implementations, and serve as a caution against investing in large-scale PD models without initial evaluations of efficacy and implementation fidelity. Given that some early childhood educators are not adequately prepared to provide the high-quality language- and literacy-rich learning opportunities that place children on the path to continued literacy success (Cabell et al., 2013; LoCasale-Crouch et al., 2007; Pelatti et al., 2014), investments are needed to develop and test PD programs that are both viable and effective in real-world contexts.

Acknowledgments

This research would not have been possible without collaboration from the Early Childhood Quality Network, including Dennis Sykes, Sharon Sullivan, Kathryn Rider, and Melissa Ross, or the cooperation of the Ohio Department of Education's Office of Early Learning and School Readiness. We also express deep appreciation for the early childhood agencies, administrators, educators, and PD facilitators involved, as well as the research staff members responsible for coordinating and executing data collection and other research activities.

Funding

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305E100030 to The Ohio State University. The opinions expressed are those of the authors and do not necessarily represent the views of the Institute of Education Sciences, U.S. Department of Education, Early Childhood Quality Network, or Ohio Department of Education.

ARTICLE HISTORY

Received 4 December 2015
Revised 8 November 2016
Accepted 6 December 2016

ORCID

Kristin S. Farley http://orcid.org/0000-0002-0507-8945

References

Arnold, D. H., Lonigan, C. J., Whitehurst, G. J., & Epstein, J. N. (1994). Accelerating language development through picture book reading: Replication and extension to a videotape training format. *Journal of Educational Psychology*, *86*, 235–243.

Arthur, A. A., McCormick, C., & Bovaird, J. (2012). *Psychometric properties of LARRC scales: Recommendations for the Family Background Questionnaire and Teacher Background Questionnaire*. Lincoln, NE: Language and Reading Research Consortium.

Assel, M. A., Landry, S. H., Swank, P. R., & Gunnewig, S. (2007). An evaluation of curriculum, setting, and mentoring on the performance of children enrolled in pre-kindergarten. *Reading and Writing*, *20*, 463–494.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: W. H. Freeman.

Bean, R. M., Draper, J. A., Hall, V., Vandermolen, J., & Zigmond, N. (2010). Coaches and coaching in Reading First schools. *The Elementary School Journal*, *111*, 87–114.

Beck, I., & McKeown, M. (2007). Increasing low-income children's oral vocabulary repertoires through rich and focused instruction. *The Elementary School Journal*, *107*, 251–271.

Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology*, *98*, 44–62.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, *33*, 3–15. doi:10.3102/0013189x033008003

Burchinal, M., Howes, C., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher–child interactions and instruction. *Applied Development Science*, *12*, 140–153.

Cabell, S. Q., DeCoster, J., LoCasale-Crouch, J., Hamre, B. K., & Pianta, R. C. (2013). Variation in the effectiveness of instructional interactions across preschool classroom settings and learning activities. *Early Childhood Research Quarterly*, *28*, 820–830. doi:10.1016/j.ecresq.2013.07.007

Carlisle, J., & Berebitsky, D. (2011). Literacy coaching as a component of professional development. *Reading and Writing*, *24*, 773–800. doi:10.1007/s11145-009-9224-4

Chen, H. T. (2010). The bottom-up approach to integrative validity: A new perspective for program evaluation. *Evaluation and Program Planning*, *33*, 205–214.

Clifford, R. M., Barbarin, O., Chang, F., Early, D., Bryant, D., Howes, C., … Pianta, R. (2005). What is pre-kindergarten? Characteristics of public pre-kindergarten programs. *Applied Developmental Science*, *9*, 126–143.

Cunningham, A. E., Etter, K., Platas, L., Wheeler, S., & Campbell, K. (2015). Professional development in emergent literacy: A design experiment of Teacher Study Groups. *Early Childhood Research Quarterly*, *31*, 62–77. doi:10.1016/j.ecresq.2014.12.002

Cunningham, A. E., Zibulsky, J., & Callahan, M. D. (2009). Starting small: Building preschool teacher knowledge that supports early literacy development. *Reading and Writing*, *22*, 287–510.

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, *38*, 181–199.

Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis*, *24*, 81–112.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, *41*, 327–350. doi:10.1007/s10464-008-9165-0

Early, D. M., Maxwell, K. L., Burchinal, M., Alva, S., Bender, R. H., Bryant, D., … Griffin, J. A. (2007). Teachers' education, classroom quality, and young children's academic skills: Results from seven studies of preschool programs. *Child Development*, *78*, 558–580.

Enders, C. K. (2010). *Applied missing data analysis* (2nd ed.). New York, NY: Guilford Press.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, *38*, 915–945.

Goldschmidt, P., & Phelps, G. (2010). Does teacher professional development affect content and pedagogical knowledge: How much and for how long? *Economics of Education Review*, *29*, 432–439.
Gottfredson, D., Cook, T., Gardner, F. M., Gorman-Smith, D., Howe, G., Sandler, I., & Zafft, K. (2015). Standards of evidence for efficacy, effectiveness, and scale-up research in prevention science. *Prevention Science*, *16*, 893–926. doi:10.1007/s11121-015-0555-x
Graham, J. W. (2012). *Missing data: Analysis and design*. New York, NY: Springer.
Hale-Jinks, C., & Knopf, H. (2006). Tackling teacher turnover in child care: Understanding causes and consequences, identifying solutions. *Childhood Education*, *82*, 219–226.
Hamre, B. K., Hatfield, B., Pianta, R., & Jamil, F. (2014). Evidence for general and domain-specific elements of teacher–child interactions: Associations with preschool children's development. *Child Development*, *85*, 1257–1274. doi:10.1111/cdev.12184
Hamre, B. K., Justice, L. M., Pianta, R. C., Kilday, C., Sweeney, B., Downer, J. T., & Leach, A. (2010). Implementation fidelity of MyTeachingPartner literacy and language activities: Association with preschoolers' language and literacy growth. *Early Childhood Research Quarterly*, *25*, 329–347.
Hill, H. C. (2009). Fixing teacher professional development. *Phi Delta Kappan*, *90*, 470–476. doi:10.1177/003172170909000705
Hindman, A. H., & Wasik, B. A. (2008). Head Start teachers' beliefs about language and literacy instruction. *Early Childhood Research Quarterly*, *23*, 479–492.
Hindman, A. H., & Wasik, B. A. (2012). Unpacking an effective language and literacy coaching intervention in Head Start: Following teachers' learning over two years of training. *The Elementary School Journal*, *113*, 131–154. doi:10.1086/666389
Horizon Research. (2002). *The 2000 National Survey of Science and Mathematics Education: Compendium of tables*. Chapel Hill, NC: Horizon Research.
Institute of Education Sciences. (2014, March). *What Works Clearinghouse procedures and standards handbook (version 3.0)*. Retrieved from http://ies.ed.gov/ncee/wwc/Handbooks
Institute of Medicine and National Research Council. (2012). *The early childhood care and education workforce: Challenges and opportunities: A workshop report*. Washington, DC: The National Academies Press.
Jackson, B., Larzelere, R., St. Clair, L., Corr, M., Fichter, C., & Egertson, H. (2006). The impact of HeadsUp! Reading on early childhood educators' literacy practices and preschool children's literacy skills. *Early Childhood Research Quarterly*, *21*, 213–226.
James-Burdumy, S., Deke, J., Gersten, R., Lugo-Gil, J., Newman-Gonchar, R., Dimino, J., … Liu, A. Y.-H. (2012). Effectiveness of four supplemental reading comprehension interventions. *Journal of Research on Educational Effectiveness*, *5*, 345–383. doi:10.1080/19345747.2012.698374
Justice, L. M., Kaderavek, J. N., Fan, X., Sofka, A., & Hunt, A. (2009). Accelerating preschoolers' early literacy development through classroom-based teacher–child storybook reading and explicit print referencing. *Language, Speech, and Hearing Services in Schools*, *40*, 67–85.
Justice, L. M., Mashburn, A. J., Hamre, B. K., & Pianta, R. C. (2008). Quality of language and literacy instruction in preschool classrooms serving at-risk pupils. *Early Childhood Research Quarterly*, *23*, 51–68. doi:10.1016/j.ecresq.2007.09.004
Kelcey, B. (2011). Assessing the effects of teachers' reading knowledge on students' achievement using multilevel propensity score stratification. *Educational Evaluation and Policy Analysis*, *33*, 458–482.
Kennedy, M. M. (2010). Against boldness. *Journal of Teacher Education*, *61*, 16–20.
Landry, S. H., Anthony, J. L., Swank, P. R., & Monseque-Bailey, P. (2009). Effectiveness of comprehensive professional development for teachers of at-risk preschoolers. *Journal of Educational Psychology*, *101*, 448–465. doi:10.1037/a0013842
Landry, S. H., Swank, P., Anthony, J., & Assel, M. (2011). An experimental study evaluating professional development activities within a state funded pre-kindergarten program. *Reading and Writing*, *24*, 971–1010. doi:10.1007/s11145-010-9243-1
Landry, S. H., Swank, P. R., Smith, K. E., Assel, M. A., & Gunnewig, S. B. (2006). Enhancing early literacy skills for preschool children: Bringing a professional development model to scale. *Journal of Learning Disabilities*, *39*, 306–324.
LoCasale-Crouch, J., Konold, T., Pianta, R., Howes, C., Burchinal, M., Bryant, D., … Barbarin, O. (2007). Observed classroom quality profiles in state-funded pre-kindergarten programs and

associations with teacher, program, and classroom characteristics. *Early Childhood Research Quarterly*, *22*, 3–17.
Lonigan, C. J., Farver, J., Phillips, B., & Clancy-Menchetti, J. (2011). Promoting the development of preschool children's emergent literacy skills: A randomized evaluation of a literacy-focused curriculum and two professional development models. *Reading and Writing*, *24*, 305–337. doi:10.1007/s11145-009-9214-6
Markussen-Brown, J., Juhl, C. B., Piasta, S. B., Bleses, D., Højen, A., & Justice, L. M. (2017). The effects of language- and literacy-focused professional development on early educators and children: A best-evidence meta-analysis. *Early Childhood Research Quarterly*, *38*, 97–115. doi:10.1016/j.ecresq.2016.07.002
Mashburn, A. J., Downer, J. T., Hamre, B. K., Justice, L. M., & Pianta, R. C. (2010). Consultation for teachers and children's language and literacy development during pre-kindergarten. *Applied Developmental Science*, *14*, 179–196.
Matsumura, L. C., Garnier, H. E., & Resnick, L. B. (2010). Implementing literacy coaching: The role of school social resources. *Educational Evaluation and Policy Analysis*, *32*, 249–272.
McGinty, A. S., & Sofka, A. E. (2009). *Classroom Literacy Observation Profile*. Charlottesville, VA: University of Virginia.
Milburn, T. F., Hipfner-Boucher, K., Weitzman, E., Greenberg, J., Pelletier, J., & Girolametto, L. (2015). Effects of coaching on educators' and preschoolers' use of references to print and phonological awareness during a small-group craft/writing activity. *Language, Speech, and Hearing Services in Schools*, *46*, 94–111.
Miller, P., & Votruba-Drzal, E. (2013). Early academic skills and childhood experiences across the urban–rural continuum. *Early Childhood Research Quarterly*, *28*, 234–248.
Moats, L. C. (1994). The missing foundation in teacher education. *Annals of Dyslexia*, *44*, 81–102.
National Early Literacy Panel. (2008). *Developing early literacy*. Washington, DC: National Institute for Literacy.
National Reading Panel. (2000). *Report of the National Reading Panel: Teaching children to read*. Washington, DC: National Institute of Child Health and Human Development.
Neuman, S. B., & Cunningham, L. (2009). The impact of professional development and coaching on early language and literacy instructional practices. *American Educational Research Journal*, *46*, 532–566.
Neuman, S. B., & Wright, T. S. (2010). Promoting language and literacy development for early childhood educators: A mixed-methods study of coursework and coaching. *The Elementary School Journal, 111*, 63–86.
O'Connell, A., Logan, J., Pentimonti, J., & McCoach, B. (2013). Linear and quadratic growth models for continuous and dichotomous outcomes. In Y. Petscher & C. Schatschneider (Eds.), *Applied quantitative analysis in the social sciences* (pp. 125–168). New York, NY: Routledge.
Pelatti, C. Y., Piasta, S. B., Justice, L. M., & O'Connell, A. A. (2014). Language- and literacy-learning opportunities in early childhood classrooms: Children's typical experiences and within-classroom variability. *Early Childhood Research Quarterly*, *29*, 445–456. doi:10.1016/j.ecresq.2014.05.004
Peterson, D. S., Taylor, B. M., Burnham, B., & Schock, R. (2009). Reflective coaching conversations: A missing piece. *The Reading Teacher*, *62*, 500–509.
Pianta, R. C., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science*, *9*, 144–159.
Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2006). *Classroom Assessment Scoring System: Pre-K*. Charlottesville, VA: Center for Advanced Study of Teaching and Learning.
Pianta, R. C., Mashburn, A. J., Downer, J. T., Hamre, B. K., & Justice, L. (2008). Effects of web-mediated professional development resources on teacher–child interactions in pre-kindergarten classrooms. *Early Childhood Research Quarterly*, *23*, 431–451.
Piasta, S. B. (2016). Current understandings of what works to support the development of emergent literacy in early childhood classrooms. *Child Development Perspectives*, *10*, 234–239. doi:10.1111/cdep.12188
Piasta, S. B., Dynia, J. M., Justice, L. M., Pentimonti, J. M., Kaderavek, J. N., & Schatschneider, C. (2010). Impact of professional development on preschool teachers' print references during shared

read alouds: A latent growth curve analysis. *Journal of Research on Educational Effectiveness*, *3*, 343–380.
Piasta, S. B., Justice, L. M., Cabell, S. Q., Wiggins, A. K., Pence Turnbull, K., & Curenton, S. M. (2012). Impact of professional development on preschool teachers' conversational responsitivity and children's linguistic productivity and complexity. *Early Childhood Research Quarterly, 27*, 387–400. doi:10.1016/j.ecresq.2012.01.001
Powell, D. R., Diamond, K. E., Burchinal, M. R., & Koehler, M. J. (2010). Effects of an early literacy professional development intervention on Head Start teachers and children. *Journal of Educational Psychology*, *102*, 299–312. doi:10.1037/a0017763
Rhodes, H., & Huston, A. (2012). Building the workforce our youngest children deserve. *Social Policy Report* (Vol. 26). Ann Arbor, MI: Society for Research in Child Development.
Rogers, R. (2014). Coaching literacy teachers as they design critical literacy practices. *Reading & Writing Quarterly*, *30*, 241–261.
Ross, J., & Bruce, C. (2007). Professional development effects on teacher efficacy: Results of randomized field trial. *The Journal of Educational Research*, *101*, 50–60.
Sailors, M., & Price, L. (2015). Support for the Improvement of Practices through Intensive Coaching (SIPIC): A model of coaching for improving reading instruction and reading achievement. *Teaching and Teacher Education*, *45*, 115–127.
Sarama, J., Clements, D. H., Starkey, P., Klein, A., & Wakeley, A. (2008). Scaling up the implementation of a pre-kindergarten mathematics curriculum: Teaching for understanding with trajectories and technologies. *Journal of Research on Educational Effectiveness*, *1*, 89–119.
Schachter, R. E. (2015). An analytic study of the professional development research in early childhood education. *Early Education and Development*, *26*, 1057–1085. doi:10.1080/10409289.2015.1009335
Schachter, R. E., Spear, C. F., Piasta, S. B., Justice, L. M., & Logan, J. A. R. (2016). Early childhood educators' knowledge, beliefs, education, experiences, and children's language- and literacy-learning opportunities: What is the connection? *Early Childhood Research Quarterly*, *36*, 281–294. doi:10.1016/j.ecresq.2016.01.008
Schachter, R. E., Weber-Mayrer, M. M., Piasta, S. B., & O'Connell, A. A. (2015). *What do coaches and teachers talk about during coaching?: Coaches' reports of the content of coaching interactions*. Manuscript submitted for publication.
Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. New York, NY: Houghton Mifflin.
Sheridan, S. M., Edwards, C. P., Marvin, C. A., & Knoche, L. L. (2009). Professional development in early childhood programs: Process issues and research needs. *Early Education and Development*, *20*, 377–401. doi:10.1080/10409280802582795
Wayne, A. J., Yoon, K. S., Zhu, P., Cronen, S., & Garet, M. S. (2008). Experimenting with teacher professional development: Motives and methods. *Educational Researcher*, *37*, 469–479.
Weber-Mayrer, M. M., Piasta, S. B., Ottley, J. R., Justice, L. M., & O'Connell, A. A. (2016). *Early childhood literacy coaching: An examination of coaching intensity and changes in educators' literacy knowledge and practice*. Manuscript submitted for publication.
Weber-Mayrer, M. M., Piasta, S. B., & Pelatti, C. Y. (2015). State-sponsored professional development for early childhood educators: Who participates and associated implications for future offerings. *Journal of Early Childhood Teacher Education*, *36*, 44–60.
Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development*, *84*, 2112–2130. doi:10.1111/cdev.12099
Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007). *Reviewing the evidence on how teacher professional development affects student achievement*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, Regional Educational Laboratory Southwest.
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*, 309–322. doi:10.1037/a0038785