education sciences

Article

Social and Emotional Learning and Early Literacy Skills: A Quasi-Experimental Study of RULER

Craig S. Bailey *, Olivia Martinez and Elizabeth DiDomizio

Yale Child Study Center, Yale School of Medicine, Yale University, New Haven, CT 06511, USA
* Correspondence: craig.bailey@yale.edu; Tel.: +1-203-432-9475

Abstract: RULER is a pre-K–12 systemic approach to social and emotional learning (SEL) that supports educators in cultivating emotion skills and provides them with pedagogical principles and tools to help them explicitly teach social and emotional skills to students. The current study tests whether preschool-level access to RULER was associated with growth in early literacy skills using a sample of 1051 preschoolers in 95 classrooms across 19 community-based centers in a small urban northeastern US community. The Preschool Early Literacy Indicators were collected three times over an academic year by district staff. Multi-level growth modeling with inverse probability weighting revealed that children in preschools with access to RULER grew significantly more in their early literacy skills than children without access to RULER—an end-of-year difference equivalent to 0.25 standard deviations, 95% CI [0.14, 0.40]. The current study demonstrates that RULER may promote young children's early literacy skills.

Keywords: social and emotional learning; early literacy; preschool; growth modeling; RULER; PELI

check for updates

Citation: Bailey, C.S.; Martinez, O.; DiDomizio, E. Social and Emotional Learning and Early Literacy Skills: A Quasi-Experimental Study of RULER. *Educ. Sci.* **2023**, *13*, 397. https://doi.org/10.3390/educsci13040397

Academic Editor: Michelle Neumann

Received: 1 February 2023
Revised: 3 March 2023
Accepted: 24 March 2023
Published: 14 April 2023

1. Introduction

Social and emotional learning (SEL) is proliferating into school-based programming from preschool to high school. With that growth, measuring impact becomes critically important. Meta-analyses of preschool SEL programs have demonstrated small to medium effect sizes on the development of both social and emotional competencies and small effect sizes on early academic outcomes, including early literacy skills [1]. Similar impacts have also been found in later grades [2–5].

The development of social and emotional skills, including empathy and perspective-taking, involves many of the same skills in literacy [6,7]. Similarly, a majority of pedagogical approaches to promote SEL include similar strategies to those used for other aspects of learning, such as dialogic reading practices and vocabulary acquisition during storybook read-alouds [2]. This likely explains why many theories of change for SEL interventions and programming include outcomes beyond social and emotional skills [8–11] and why the Collaborative for Academic, Social, and Emotional Learning, the leading organization that reviews SEL programs, only recommends those that show efficacy beyond social and emotional skills [12].

Originally conceptualized by Salovey and Mayer in 1990 and updated since [13–15], emotional intelligence is comprised of a set of abilities across different aspects of functioning (i.e., perceiving and expressing emotions, facilitating emotional experience in thought, understanding emotions, managing emotions). Education-based applications of emotional intelligence organize these abilities using the acronym "RULER," which represents five key skills of emotional intelligence (i.e., recognizing, understanding, labeling, expressing, and regulating emotions); "RULER" also represents the name of the pre-K–12 approach to SEL employed in this study [8].

RULER is an evidence-based approach to SEL that has shown impact for teachers and students across domains, ages, and cultures [8,16–20]. In the classroom, teachers embed

RULER tools and practices into teacher–child interactions and routines and teach a curriculum specifically developed using English Language Arts best practices and standards to support students in building an advanced emotion vocabulary [21]. Previous studies have investigated RULER's impact on teacher–student interactions from kindergarten through grade 12 [16–18,20] as well as fifth- and sixth-grade student English Language Arts grades, engagement, attention problems, learning problems, leadership, study skills, and social skills [16,19,22]. Grade 9 students in schools and classrooms using RULER have also shown improvements in emotion understanding and grade point averages. However, RULER's impact on preschool-aged children's early literacy has yet to be explored. The purpose of this study is to investigate whether RULER is associated with growth in early literacy.

1.1. Early Literacy

Literacy and its acquisition are fundamental during early childhood. When preschoolers build their alphabet knowledge and phonemic awareness, and enhance their expressive and receptive language and comprehension, they have the building blocks for meeting the academic demands of kindergarten [23,24]. Literacy—decoding, oral reading fluency, reading comprehension, writing, and spelling—and its emergence is conceptualized along a continuum. It begins in infancy with the development of necessary skills (e.g., pointing and referencing, book engagement, holding books, and flipping through pages), extends through preschool (e.g., awareness that books tell stories, personally connecting to stories, telling their own stories), and continues into elementary school (e.g., text and illustrations conveying meaning, beginning of reading [25]). Early literacy, according to a 2008 Report from the National Early Literacy Council, consists of: (a) alphabet knowledge; (b) phonological awareness; (c) rapid, automatic naming of letters, digits, objects, and colors; (d) writing one's name, and (e) phonological memory—skills which were found to be associated with indicators of later literacy [26].

Previous research has shown that the preschool period is critical for the development of early literacy skills [26,27]. Starting in preschool, individual differences in early literacy and literacy itself remain stable [28–30]. Even with this stability, early literacy interventions and general pedagogical approaches still demonstrate impact [31–33], with some research even pointing to the school context as ripe for early literacy promotion [34].

Universal SEL programming with an integrative, whole-child approach that enhances how teachers talk and read to children (e.g., RULER) may impact social–emotional development and literacy acquisition [35,36]. Zins and colleagues (2004) note that many evidence-based SEL programs promote school readiness skills including school engagement, school liking, and a positive approach towards learning [36]. Indeed, whole-child perspectives in education also acknowledge the developmental overlap between social–emotional and pre-academic domains and call for integrative pedagogical approaches [37]. For example, Curby and colleagues (2015) found that young children's expressiveness and regulation patterns, as well as their emotion knowledge, were associated with early literacy skills [38]. Similarly, Medford and McGeown (2016) found that inattention was the strongest predictor of early literacy skills, followed by conduct and peer relationships [39], which is bolstered further by previous research [23,40–42]. But other research found little evidence supporting an association between social and emotional behaviors and skills and later literacy [23,42].

Research has shown that young children with more developed emotion skills are more engaged in the learning process and perform better academically than children who struggle with these skills [43,44]. Deficits in SEL skills are associated with anti-social behavior, peer rejection, grade retention, school dropout, substance abuse, arrest records, and negative relationships with peers, teachers, and parents [45,46]. In general, a growing body of research documents the importance of emotional intelligence and related social and emotional competencies for short- and long-term academic success [19,43,47,48]. Among kindergartners, social and emotional skills predict literacy scores even after controlling for IQ [49].

It is, therefore, reasonable to hypothesize that interventions that enhance teacher–child interactions as well as the pedagogy that teachers use as the focus of SEL programming, such as storybook read-alouds along with personal and social narratives, may result in changes in children's early literacy skill acquisition over time. Early childhood educator professional development focused on enhancing young children's early literacy skills has been shown to yield small to moderate effects with stronger effects when programming enhances teacher–child interactions [32,50,51]. SEL interventions, too, have shown small effects on these same outcomes. Some researchers hypothesize that the reported effects on children's early literacy skills may be due to the role of role of language and reading within SEL programming [52,53].

1.2. RULER

RULER is a whole-school, universal approach to SEL that provides professional development for educators, a curriculum for students, and resources for families. It involves four tools. The Mood Meter is a color-coded grid with axes representing energy (arousal) and pleasantness (valence) to facilitate the development of the RULER skills. The Charter is a process for creating an emotionally-safe climate. Building a Charter encourages open communication about feelings and helps groups answer the questions "How do we want to feel?" and "How will we ensure everyone has these feelings?". The Blueprint is a process for building empathy and perspective-taking skills, including a guide for managing past, present, or future conflicts. The Meta-Moment supports effective emotion regulation with an emphasis on disrupting automatic responding with intentional strategies aligned with ones' vision of their best self [8]. Staff learning and development include routines that accompany the four tools to enhance staff's own emotional intelligence [8]. RULER also is intended to support educators and their students in cultivating an "emotions matter" mindset [8].

When implemented with children, RULER in early childhood settings features a classroom curriculum with 13 units and three of the four tools (i.e., Mood Meter, Charter, Meta-Moment), scaled developmentally, as well as routines to promote and teach emotional intelligence in the classroom [8]. Elements of the curriculum are based upon best practices for the natural acquisition of language [25]. Each classroom unit is completed over several weeks and focuses on a new emotion word. A unit provides teachers with lesson plans for using their own personal stories and key storybooks, as well as small and large group activities to teach the content knowledge associated with the emotion words. Daily routines for children, such as checking in on the Mood Meter, classroom check-ins with their peers using the Charter, and taking deep breaths with developmentally-tailored regulation strategies using the Meta-Moment, help them develop the five key emotion skills. RULER has been field tested for efficacy with young children and found to be developmentally appropriate [54].

Given RULER's approach to building social and emotional skills, children's early literacy skills may be enhanced by the activities in RULER's curriculum (e.g., storybook read-alouds and incorporating personal and social narratives; [8,16]). Evaluations of RULER have demonstrated small-to moderate-sized effects of the approach on fifth and sixth graders' social and emotional competence as well as English language arts grades [16]. Beyond RULER, meta-analyses looking at 286 studies examining the effects of SEL programming spanning pre-K–12 found that students performed significantly better on standardized tests of reading and math as well as on early literacy and pre-math outcomes compared with students in non-participating schools and classrooms [1,3].

1.3. Enhancing Early literacy Skills through RULER's Approach to SEL: Two Mechanisms

As shown in Table 1, RULER purports to enhance children's early literacy skills, encompassing their alphabet knowledge, comprehension, phonological awareness, and vocabulary–oral language through two interrelated mechanisms associated with their learning: (1) enhancing educator skills and teacher–child interactions and (2) a classroom curriculum focused on directly teaching emotion vocabulary, knowledge, and skills and indirectly promoting attention to and engagement in learning.

Table 1. Enhancing early literacy skills through RULER's approach to SEL.

		Early Literacy Skills			
		Alphabet Knowledge	Comprehension	Phonological Awareness	Vocabulary-Oral Language
RULER Components	Emotions matter mindset and emotional intelligence (i.e., recognizing, understanding, labeling, expressing, and regulating)		✔		✔
	Teacher–child interactions (i.e., language modeling, quality feedback)		✔	✔	✔
	Curricular focus on emotion vocabulary, knowledge, and skills (i.e., storybook read-alouds, play-based activities)	✔	✔	✔	✔
	Core routines (i.e., Charter, Mood Meter, Meta-Moment)		✔		✔

Children begin to develop early literacy skills through the adults in their lives by way of joint attention during interactions and linguistic inputs and exposure to print, including storybooks [55,56]. RULER emphasizes the role of the adult through enhancing their curiosity and reflectiveness about emotion and fostering a deeper understanding of the causes and consequences of emotion. With RULER, adults are expected to cultivate a growth mindset about children's development of emotional competence. As a result of shifting their mindsets and improving their own skills, teachers exposed to RULER exhibit enhanced teacher–child interactions (e.g., quality feedback and language modeling), which have been shown across numerous studies to impact children's learning, including their early literacy [16,20,50,57]. Best practices for how teachers interact with children are considered by many as indicative of high-quality teaching (e.g., in emotionally, organizationally, and instructionally supportive ways) as well as the key ingredients of high-quality instruction (e.g., variability of language modeling, richness of discussion, and complexity of concepts; [51]). In parallel to language and literacy programs, some SEL programs have demonstrated impact on children's early reading and math skills by way of improvements in teacher–child interactions [58]. Studies by Howes and colleagues (2008) and Wasik and Hindman (2011) reveal that enhancing teacher–child interactions may have the most benefit for children's phonological awareness, expressive vocabulary, and receptive vocabulary skills but may do little to support alphabet awareness [50,57].

The second way RULER may enhance children's early literacy skills is through its sequenced and development-focused curriculum featuring numerous storybook read-alouds, family involvement, and play-based activities in each unit (e.g., music and movement focused on emotion; art as a form of emotional expression; guided free play as a means to explore emotion; and small and large group activities that require sharing personal experiences about emotion; [8,21]). RULER's classroom curriculum supports teachers with embedding the teaching of emotion vocabulary using these key activities and practices. For example, learning about the word "disappointed" in a variety of experiences and contexts provides children with new ways of recognizing, understanding, labeling, expressing, and regulating (RULER) disappointment [8].

This multi-faceted approach to learning aligns with standards from the International Reading Association and the National Council of Teachers of English. These standards emphasize providing a range of content and modalities for children to engage in, with supports for them to understand and apply their knowledge and skills in a variety of contexts for a variety of purposes, not only by themselves, but also their with teachers, peers, and family [59]. In addition, the National Association for the Education of Young Children along with the International Reading Association provide a set of developmentally appropriate practices that are enhanced when using RULER's classroom curriculum. These practices include repetition along with ample multi-modal opportunities for child-led engagement, expression, and play that connect children's experiences to the narratives within storybooks [59]. RULER's classroom curriculum also promotes teachers' modeling of enjoyment and inquiry about others within these narratives. Best practices for early literacy promotion within RULER's classroom curriculum also include a focus on specific vocabulary (i.e., emotion vocabulary) with an alignment to the details of the story to help children understand the meaning of emotions. Studies show that these practices yield moderate to large effects on children's early literacy skills, including expressive vocabulary, receptive vocabulary, phonological awareness, print concepts, and expressive language [60–63]. Whether these effects are specifically associated with RULER in early childhood settings has yet to be explored.

1.4. Current Study

We argue that children attending preschools with access to RULER may score higher on assessments of early literacy skills than their peers in preschools without access to RULER. We anticipate a small to moderate effect based on previous research on RULER and other SEL programs [1,3,16,18–20].

2. Methods

2.1. Participants

This study includes 1,051 children in 95 classrooms across 19 community-based early childhood centers in a small, urban, northeastern US community with an estimated population of about 90,000 [64]. These community-based programs are full-time and receive a mix of federal and state funding to subsidize tuition. Based on US census data, our sample represents 20% of all children under five in the community at the time of data collection. Children were mostly preschool-aged 3–5-year-olds but also included some two-year-olds due to school enrollment cutoff criteria: $M_{\text{age in months}} = 45.15$, $SD_{\text{age in months}} = 7.03$, $min–max_{\text{age in months}} = 27.73–66.73$. Families were low-income on average (*median* = $27,870, *SD* = $50,196, *range* = $0–733,443) and tended to have lower incomes compared with the broader community, $Median_{\text{2017 Census}}$ = $81,546, $Z_{\text{Wilcoxon matched-pairs signed-ranks test}} = 15.29$, $p < 0.001$. Families identified their children as Hispanic/Latine (53%), White (22%), Black/African American (17%), Asian (4%), or multiracial/other (4%). Our sample had proportionally more Hispanic/Latine children than the overall community but similar rates for other races and ethnicities, which was likely due to recent increases in Hispanic/Latine families in the community: $\text{White}_{\text{2017 Census}}$ = 52%, Hispanic American or $\text{Latine}_{\text{2017 Census}}$ = 27%, $\text{Black/African American}_{\text{2017 Census}}$ = 16%, $\text{Asian}_{\text{2017 Census}}$ = 5%, $\text{multiracial/other}_{\text{2017 Census}}$ = < 1%. There were approximately the same number of girls (48%) as boys (52%), which contrasts with the broader community; however, the difference between our sample and the broader community was not statistically significant: $\text{female}_{\text{2017 Census}}$ = 52%, $Z_{e\text{quality of proportions}} = 0.24$, $p = 0.810$. No information was provided by the district about the teaching staff at these programs. Home language, multilingual exposure and proficiency, immigration status, and other aspects of children's language context and skills were also not collected by the district.

2.2. Procedures

2.2.1. Data Collection and Access

Data were accessed by the first author by contacting school district staff and their community stakeholder partners, including early childhood council representatives and non-profit community organizations. This study is therefore a secondary data analysis and was reviewed by community stakeholders and approved by the IRB of the first author's university as well as the State's Office of Early Childhood. During the 2017–2018 school year, families consented to have their children assessed by the district. Children's demographic data were collected via family surveys, whereas children's early literacy skills were assessed three times by trained school district staff at the start of the school year in September, during the middle of the school year in January, and at the end of the school year in May (i.e., T1, T2, and T3). We consider T1 as the baseline and assume unequal time intervals given variability in when each child was assessed. Note that other aspects of school readiness, including social and emotional skills, were not collected by the district; only early literacy skills were collected and made available.

2.2.2. RULER Access and Implementation

With funding from a local family foundation and through a community stakeholder, the study's first author used an early childhood community meeting of preschool directors to provide an overview of RULER. Funding only allowed for about half of the community preschools to be trained, and 8 of the 19 preschools expressed interest and subsequently were offered access to RULER, constituting a quasi-experimental design. Each RULER-enrolled preschool created a RULER Implementation Team consisting of four to eight members, including administrators, classroom and family support staff, and teachers, who were subsequently trained in RULER using a train-the-trainer approach across four four-hour sessions. In RULER's train-the-trainer approach, the preschool-based RULER Implementation Teams are trained first—these teams go on to lead implementation by using RULER's tools and activities to train the remaining staff members (see Brackett et al., 2019 [8], for an overview). In general, anonymous surveys revealed that RULER Implementation Teams felt confident using RULER (M = 3.40, SD = 0.56, *range possible* = 1–4 (1 = "Strongly disagree", 2 = "Disagree", 3 = "Agree", 4 = "Strongly agree")) and believed that RULER resulted in significant impact, M = 3.24, SD = 0.68.

2.3. Measures

2.3.1. Child and Family Demographics

Covariates included gender (i.e., boy or girl), age in months, estimated yearly family income, race and ethnicity (i.e., Hispanic/Latine, Asian, Black/African American, White, or multiracial/other), and early literacy assessment version (i.e., 3–4-year-old books for young children or 4–5-year-old books for older children).

2.3.2. Early Literacy Skills

Early English literacy skills were assessed using the Preschool Early Literacy Indicators (PELI) [24], a 15-min direct child assessment comprising four subtests: Alphabet Knowledge, Comprehension, Phonological Awareness, and Vocabulary–Oral Language. Each subtest of the PELI was embedded within a storybook format intended to engage the child in active participation. Six books were used, with three books for 3–4-year-olds (i.e., children two years away from kindergarten) and three books for 4–5-year-olds (i.e., children going into kindergarten). Children were assessed in English.

For the Alphabet Knowledge subtest, letters were embedded into a picture related to the book's theme; the child was asked to identify as many letters as they could. The assessor prompted the child when needed and would continue until the child missed three consecutive letters. The alphabet knowledge score was totaled at the end and could range from 0 to 26. For the Comprehension subtest, the assessor read a short story to the child, taking time during and at the end of the story to ask the child questions ranging from simple to complex. Once the story was finished, the child was shown a series of pictures while being read a short version of the story, with the assessor pausing throughout to have the child fill in each missing word in the story. The child received up to 2 points for answering questions correctly during the story and 1 point for each word they filled in correctly during the short version of the story; scores ranged from 0 to 23.

The Phonological Awareness subtest involved a game where a child was given a picture of a scene from the book and ten pictures related to that scene. The child was asked to say the sound or first part of the word indicated by each picture, receiving a total score ranging from 0 to 15. For the Vocabulary–Oral Language subtest, the child was shown a picture of a scene from the book with 10 objects in the scene related to a specific theme. The assessor asked the child to name the 10 objects and to tell the assessor as much as they could about at least five of the objects. Words ranged from easy to medium to difficult, and the child received a point for each picture they correctly named. The child could receive up to five points for quality of response given about the five objects, and scores could range from 0 to 35.

The PELI composite score provides the best overall estimate of a child's early literacy skills and is a weighted composite of alphabet knowledge × 2, comprehension × 4, phonological awareness × 4, and vocabulary–oral language × 3 in such a way that each subtest contributes equally. PELI validation studies have revealed concurrent and predictive criterion validity between various PELI subtests, standardized norm-referenced tests of language development and other indicators of early literacy, including the Peabody Picture Vocabulary Test–IV, Clinical Evaluation of Language Fundamentals PreK, Expressive Language Index, Clinical Evaluation of Language Fundamentals PreK, Dynamic Indicators of Basic Early Literacy Skills, Test of Preschool Early Literacy, and Get Ready to Read (i.e., αs = 0.28–0.92; see Kaminski et al., 2014 [24] for a summary). Kaminski and colleagues (2014) also report that the PELI weighted composite score has demonstrated reliability of between 0.75 to 0.96 across several studies [24].

For the current study, internal consistencies of the PELI weighted composite score at each of the three timepoints were high (α_{T1} = 0.84, α_{T2} = 0.86, and α_{T3} = 0.86). Associations among the timepoints were also high (rs(780–838) = 0.80–0.90, ps < 0.001). However, the PELI weighted composite score tended to be significantly more reliable when 4–5-year-old books were used with older children than when 3–4-year-old books were used with younger children ($\alpha_{T1\ 3\text{–}4}$ = 0.67, $\alpha_{T\ 14\text{–}5}$ = 0.82, ω(496, 348) = 1.83, p < 0.001; $\alpha_{T2\ 3\text{–}4}$ = 0.76, $\alpha_{T2\ 4\text{–}5}$ = 0.85, ω(419, 384) = 1.60, p < 0.001; $\alpha_{T3\ 3\text{–}4}$ = 0.81, $\alpha_{T3\ 4\text{–}5}$ = 0.86, ω(525, 421) = 1.36, p < 0.001).

2.4. Analytic Strategy

Covariate-adjusted multi-level growth models with inverse probability of treatment and missing data weighting were run in Stata 16.1; these models explored nested variance components of children's early literacy growth with early literacy occasions (i.e., T1–T3) at Level 1, children at Level 2, classrooms at Level 3, and preschools at Level 4. Multi-level growth models in a quasi-experimental design are conceptually similar to the difference-in-differences approach [65] and provide an estimate of the difference in the rate of change (i.e., average treatment effect) between the group of preschools with access to RULER and the group without access. In addition to the nested structure of educational data (see Table 2 for intra-class correlation coefficients and the degree of non-independence) and the unequal time intervals of early literacy data collection, there were many factors that were considered with the repeated measures design, leading to the use of multi-level growth modeling [66].

Modeling change is fraught with controversy and confusion, especially in quasi-experimental designs [67]. In general, multi-level growth modeling is robust against partially missing data patterns and is generally more powerful than traditional repeated-measures approaches [68]. Multi-level modeling allows for modeling baseline–growth covariance in an intercept-slope random effects model. Note that although gain score analyses are technically superior to baseline-as-a-covariate models with baseline inequivalence [69], a simple gain score analysis would ignore T2 in our case; three time points are almost always better—that is, more powerful—than two. Another benefit of multi-level modeling is higher power as a result of lowering standard errors compared with gain score analysis due to modeling of autoregressive variation in the outcome [70]. Missing data, baseline equivalency, and selection bias were the most prominent factors weighing on the analytical model specifications given the quasi-experimental design.

Table 2. Descriptive statistics for analysis variables.

Variable	N	M	SD	Min–Max	Skewness	Kurtosis	$ICC_{classroom}$	$ICC_{preschool}$
Time 1 (Baseline)								
Alphabet Knowledge	846	8.41	9.44	0–26	0.79	1.99	0.02	0.19
Vocabulary–Oral Language	846	12.51	8.02	0–32	0.05	2.13	0.02	0.26
Comprehension	846	9.57	6.04	0–23	−0.06	1.87	0.04	0.19
Phonological Awareness	846	3.46	4.94	0–15	1.11	2.67	0.01	0.23
PELI Composite Score	846	124.51	76.88	0–309	0.46	2.32	0.03	0.28
Gender	802	0.52	0.50	0–1	−0.06	1.00	0.01	0.00
Age	924	450.14	7.03	27.73–66.73	−0.25	2.15	0.71	0.00
Assessment Version	1044	0.55	0.50	0–1	−0.22	1.05	0.43	0.08
Family Income	428	39,651.04	50,195.62	0–733,443	7.27	90.05	0.06	0.13
Race/Ethnicity	802							
Hispanic/Latine	428	0.53	0.50	0–1	−0.13	1.02	0.21	0.00
Asian	30	0.03	0.19	0–1	4.88	24.78	0.03	0.00
Black/African American	140	0.17	0.38	0–1	1.71	3.94	0.14	0.00
Multiracial/Other	30	0.04	0.19	0–1	4.88	24.78	0.19	0.02
White	174	0.22	0.41	0–1	1.37	2.89	0.02	0.05
Time 2								
Alphabet Knowledge	905	12.19	10.12	0–26	0.16	1.35	0.04	0.13
Vocabulary/Oral Language	905	15.01	7.80	0–32	−0.22	2.22	0.06	0.19
Comprehension	905	12.15	6.64	0–23	−0.34	1.93	0.06	0.20
Phonological Awareness	905	5.08	5.56	0–15	0.51	1.60	0.00	0.21
PELI Composite Score	905	144.83	77.75	0–288	−0.09	1.92	0.04	0.23
Time 3								
Alphabet Knowledge	949	14.45	9.98	0–26	−0.19	1.38	0.05	0.12
Vocabulary/Oral Language	949	17.68	8.33	0–34	−0.32	2.42	0.05	0.20
Comprehension	949	13.33	6.20	0–23	−0.67	2.50	0.05	0.11
Phonological Awareness	947	6.58	5.99	0–15	0.13	1.33	0.04	0.18
PELI Composite Score	947	163.69	78.25	0–306	−0.34	2.15	0.04	0.21

2.4.1. Missing Data

All preschools and their classrooms remained in the sample for the duration of data collection. Because the effect of access to RULER was at the preschool level, the study is classified by What Works Clearinghouse standards [71] as low-attrition. Nonetheless, there was non-trivial missing data at the child level. Rates of child-level missing data ranged from 0 to 59% (see Table 2). Among the covariates, rates of missing data were significantly correlated, $r_{total}(1042) = 0.66$, $p < 0.001$. Little's Missing Completely at Random (MCAR) test [72] revealed that data were *not* MCAR, χ^2 (84) = 983.87, $p < 0.001$. Follow-up missing-data analyses revealed that early literacy missingness was not systematically dependent on the values of the covariates, but children with no age data were more likely to have higher PELI scores. A follow-up missing-data analysis explored whether early literacy missing data were dependent on the values of the covariates by regressing the missingness of early literacy data over time onto the covariates using multilevel logistic regression, accounting for preschool-, classroom-, and child-level nesting. Results revealed that early literacy missingness was not systematically dependent on the values of the covariates. Early literacy scores were then regressed over time onto covariate missingness to explore whether early literacy scores depended on the missingness of the observed covariates, revealing a statistically significant association with missing data on child age after controlling for missingness on the other covariates; $b = 15.06$, $p = 0.019$. Given these patterns, missing data were handled in two ways. First, doubly robust inverse probability-of-missing weighting was used to adjust the RULER impact estimate, using regression-based covariate adjustment as well as the conditional probability of being in the sample based on the covariates [73]. Propensity scores were estimated using a multi-level probit model. Propensity scores were estimated using a multi-level probit model, which exhibited sufficient balance on covariates (pseudo-R^2 = 0.01, χ^2 (1) = 16.94, p = 0.076, M_{bias} = 6%, Rubin's B = 18%, Rubin's R = 1.89) except for a greater proportion of Black/African American children compared with Hispanic/Latine children with missing data (bias = 10%, t = 2.13, p = 0.033). Overall, propensity scores reduced this bias by between 31–96%. Inverse probability-of-missing weights were calculated based on the propensity scores and subsequently included in the primary analyses at the occasion level. The second way that missing data were handled in the primary analyses was by using the expectation–maximum likelihood algorithm, which accounts for multivariate missing data by producing unbiased parameter estimates.

2.4.2. Baseline Equivalency and Selection Bias

Testing for baseline equivalency is a requirement for quasi-experimental studies, because observed baseline differences, as well as unobserved differences, are threats to causal inference [71,74,75]. According to What Works Clearinghouse standards [71], baseline equivalency was achieved for the early literacy outcome ($M_{\text{No access to RULER at baseline}}$ = 134.47, $M_{\text{Access to RULER at baseline}}$ = 125.75, g = 0.08, given a Hedges' $g \leq$ cutoff of 0.25 for baseline-adjusted analyses). Baseline equivalency was also achieved for all covariates (i.e., gender, family income, assessment version, and race/ethnicity) except for child age (i.e., g = 0.40), Black/African American (g = 0.32), and Hispanic/Latine (g = 0.26). with overall standardized differences ranging between 0.04 and 0.40 standard deviations (see Table 3).

Table 3. Standardized differences between no-access-to-RULER and access-to-RULER groups in preschoolers' early literacy scores and covariates before and after propensity-of-treatment weighting.

		No Access to RULER	Access to RULER	Standardized Difference
PELI Composite Score	Before	132.47	125.75	0.08
	After	118.71	119.38	0.03
Gender	Before	0.56	0.48	0.17
	After	0.59	0.56	0.07
Age	Before	47.15	44.33	0.40
	After	45.77	45.31	0.07
Assessment Version	Before	0.36	0.38	0.04
	After	0.42	0.42	0.00
Family Income	Before	50,800.22	40,485.44	0.19
	After	39,334.20	41,135.68	0.03
Hispanic/Latine	Before	0.66	0.53	0.26
	After	0.57	0.57	0.00
Asian	Before	0.05	0.03	0.07
	After	0.04	0.04	0.02
Black/African American	Before	0.14	0.26	0.32
	After	0.27	0.24	0.08
Multiracial/Other	Before	0.00	0.02	0.11
	After	0.00	0.02	0.02
White	Before	0.15	0.16	0.00
	After	0.13	0.13	0.00

Note. Adjusted means were calculated using primary model specification and were either linear or logit depending on the variable. Standardized differences are in standard deviation units and are equivalent to Hedges' g.

Threats to causal inference based on baseline differences were addressed in two ways. First, doubly robust inverse probability-of-treatment weighting was used in the primary analyses to provide more weight to cases of children in the access-to-RULER group that resembled—based on covariates—children in the no-access-to-RULER group, as recommended by What Works Clearinghouse [71]. The conditional probability of being in the access-to-RULER group was estimated using baseline covariates in a probit model. Inverse probabilities of access-to-RULER weights were then calculated and included at the preschool level. With this technique, the resulting propensity scores were then inversed and used as weights along with covariate regression adjustment. Ergo, children in the no-access-to-RULER group that resembled children in the access-to-RULER group received more weight. Estimation was similar to the inverse probability of the missing model providing propensity scores that exhibited sufficient balance on covariates at baseline (pseudo-R^2 = 0.01, χ^2 (1) = 13.77, p = 0.13, M_{bias} = 5.00%, Rubin's B = 18%, Rubin's R = 1.46). Overall, propensity scores reduced bias between the groups by 56–174%, with standardized differences between 0.00 and 0.19 (see Table 3), suggesting that the estimated potential of early literacy growth will likely be independent of the probability of RULER conditional on the covariates. The second way threats to casual inference were addressed was by better understanding the potential influence of unobserved confounding variables on the impact estimate, using Frank's (2000) approach to estimate how strong an unobserved covariate would have to be to invalidate the results [76]. This approach compliments both the multilevel difference in growth model estimation, which may account for some selection bias due to unobserved confounding variables, and the use of inverse probability-of-treatment weighting that reduced selection bias in the impact estimate due to observed confounding variables.

2.4.3. Multi-Level Growth Model Specification

Five models were run using a model-building approach [77]: (1) unconditional, (2) unconditional growth, (3) unconditional growth with random slopes, (4) conditional growth with random slopes, and (5) moderated growth with random slopes. All models included inverse probability-of-missing-data and treatment weights at the occasion level (i.e., for missing data) and school level (i.e., for treatment), respectively. The unconditional model explored variance in early literacy skills between preschools, classrooms, and children, and provides a point of comparison for other models. One assumption in testing whether preschool-level access to RULER is associated with growth in preschooler's early literacy skills is that children exhibit growth regardless of condition. The unconditional growth model tests this assumption of growth by modeling the average rate of change across the three timepoints. To test the assumption that growth in children's early literacy skills is constant between preschools, classrooms, and children or varies between preschools, classrooms, and children, the unconditional growth with random slopes model was run. Growth–slope-intercept covariance was also explored in this model, which tests the assumption that children's growth does not covary with their baseline scores [66]. Random-slope models are often superior to fixed-slope models in that they have the potential to reduce Type I errors and explain additional variance [78].

The conditional growth with random slopes model explored the influence of potential confounding variables and included gender (boys = 1), age in months, family income, race and ethnicity (Hispanic/Latine vs. Asian, Hispanic/Latine vs. Black/African American, Hispanic/Latine vs. White, and Hispanic/Latine vs. multiracial/other), assessment version (4–5-year-old-books = 1), and the main effect of RULER access = 1.

The primary research question was addressed using the moderated growth with random slopes model by testing the degree to which growth in early literacy scores was associated with preschool-level access to RULER. The moderated growth with random slopes model explored cross-level moderation between time and access to RULER. The four levels of equations for the ith observation for the jth child in the kth classroom at the lth school can be represented in a combined equation:

$$Y_{ijkl}(\text{PELI}) = \gamma_{0000} + \gamma_{1000}(\text{Time}) + \gamma_{0n00}(\text{Covariates}) + \gamma_{0001}(\text{Access to RULER}) + \gamma_{1001}(\text{Time X Access to RULER}) + r_{0jkl} + r_{1jkl} + \mu_{00kl} + \mu_{10kl} + s_{000l} + s_{100l} + e_{ijkl}$$

The main effect of access to RULER represents the covariate-adjusted weighted difference in PELI scores at baseline as it accounts for covariance associated with growth. Variance components are reported in standard deviation units to facilitate interpretation. Family income was centered for each child on their school-level group mean, because 13% of the variance in family income was between preschools, 6% was between classrooms, and the remaining 81% was between children. Similarly, age in months was centered on the classroom-level group mean, because 0% of the variance in age was between preschools, 71% between was classrooms, and the remaining 29% was between children.

2.4.4. Effect Size

For a metric of standardized difference, a coefficient equivalent to Cohen's *d* was used to gauge the impact of offering access to RULER on children's early literacy skills. Based on recommendations by Feingold (2013) [79], effect size was calculated using the equation

$$\delta = (\gamma_{1001} \times 2)/\text{sqrt}(\tau_{\text{baseline pooled}})$$

where γ is the unstandardized, covariate-adjusted, weighted fixed-effect difference in the rate of change between the access-to-RULER and no-access groups. The "2" represents the number of timepoints from baseline, ensuring that the effect size metric is comparable with common pre–post experimental designs. Finally, $\tau_{\text{baseline pooled}}$ is the pooled total variance across the access-to-RULER and no-access groups at baseline, which best represents the variance in the population without the influence of RULER [80,81]. Cohen's *U3*—the proportion of overlap in group distributions [82]—was also used to provide an education-relevant interpretation of effect for the access-to-RULER and no-access-to-RULER groups.

Based upon recommendations by LaHuis and colleagues [83] (see also, Snijders and Bosker, 2012 [84]), the proportion of residual variance explained between models without random slopes (i.e., pseudo-R^2_{between}) was reported. Also reported was the within-model rather than the between-model value, pseudo-R^2_{within}, an alternative proposed by Nakagawa and Schielzeth (2013) based on the proportion of variance in the predicted scores divided by the predicted score variance and total residual variance [85].

Finally, regarding internal and external validity, and to further contextualize the estimated impact of RULER on children's early literacy skills, the robustness of the average treatment effect was also estimated by calculating how much bias there must be in the estimate to invalidate the estimate of causal inference as well as the likelihood of omitted-variable bias. An approach developed by Frank and colleagues (2013) [86] was used to estimate how many cases (i.e., children over time) in the access-to-RULER group would have to exhibit the same pattern as the no-access-to-RULER group for an effect size equivalent to zero, effectively invalidating the results. The larger the number of cases needed to invalidate the results, the less presumed bias. How strong a confounding variable would have to be to invalidate the results was also quantified using an approach developed by Frank [76]. Both methods factor in effect size, standard error, observations, and number of covariates.

3. Results

3.1. Assumptions

The unconditional model revealed that children's initial early literacy scores were 140.87 on average (see Table 4) and significantly varied between preschools ($\tau_{\psi 0000} = 37.33$), classrooms ($\tau_{\beta 0000} = 13.44$), and children ($\tau_{\pi 0000} = 56.98$, all *p*s < 0.001) (see Table 5). Thus, most of the variation in scores was between children, followed by between preschools, with the least amount of variance between classrooms. With the addition of time as a variable, the unconditional growth model fit the data statistically significantly better, with 4% variance explained relative to the unconditional model (Table 5), revealing that children's early literacy scores at baseline were 114.93 on average and significantly varied over time at a rate of 25.49 points per timepoint on average, $p < 0.001$ (Table 4).

The unconditional growth with random slopes model fit the data statistically significantly better than the unconditional growth model but did not explain more total variance (Table 3). However, this more-complex model did reveal how growth in early literacy scores significantly varied between preschools ($\tau_{\psi 1000} = 6.20$), classrooms ($\tau_{\beta 1000} = 5.15$), and children ($\tau_{\pi 1000} = 13.06$, all *p*s < 0.01) (Table 5). Preschools with children initially scoring lower in their early literacy scores tended to grow more over the year than preschools with children initially scoring higher ($r\tau_{\psi 0000}$, $\tau_{\psi 1000} = -0.71$, $p < 0.001$). The conditional growth with random slopes model was explored next and subsequently, fit the data statistically significantly better with 27% variance explained and 18% explained by the fixed effects (Table 5), revealing that children who were administered the 4–5-year-old book version scored statistically significantly higher, by 41.26 points, in their average early literacy scores over the year than children administered the 3–4-year-old book version (adjusted $M_{3\text{–}4\ \text{Version}}$ = 109.86, adjusted $M_{4\text{–}5\ \text{Version}}$ = 153.86) (see Table 4).

Table 4. Fixed effects by model.

Variable	Unconditional	Unconditional Growth	Unconditional Growth Slopes	Conditional Growth Slopes	Moderated Growth Slopes
Intercept γ_{0000}	140.87 *** (9.28)	114.93 *** (10.33)	114.20 *** (10.65)	82.76 *** (10.37)	85.77 *** (10.61)
Time γ_{1000}		25.49 *** (1.50)	25.85 *** (1.83)	29.63 *** (3.85)	24.55 *** (3.66)
Gender γ_{0200}				6.08 (5.85)	5.95 (5.90)
Age γ_{0300}				−0.35 (1.05)	−0.33 (1.04)
Assessment Version γ_{0400}				41.26 *** (7.30)	41.49 *** (7.24)
Family Income γ_{0500}				0.00 (0.00)	0.00 (0.00)
Hispanic/Latine vs. Asian γ_{0600}				−31.98 * (15.17)	−31.96 * (15.12)
Hispanic/Latine vs. Black/African American γ_{0700}				−8.86 (5.88)	−8.82 (5.83)
Hispanic/Latine vs. Multiracial/Other γ_{0800}				4.04 (10.83)	3.91 (10.77)
Hispanic/Latine vs. White γ_{0900}				−4.92 (9.41)	−4.57 (9.40)
Access to RULER γ_{0001}				−2.76 (7.09)	−8.77 (7.21)
Time X Access to RULER γ_{1001}					9.21 *** (2.71)

Note. Effects were estimated using maximum likelihood and were adjusted using the inverse probability of missing data and treatment assignment; values in parentheses are cluster-robust standard errors; time was coded as 0, 1, or 2, corresponding to T1–T3; gender was coded as girl = 0, boy = 1; age was centered at the child's classroom mean; assessment version was coded as 3–4-year-old = 0, 4–5-year-old = 1; family income was centered at the child's school mean; for race and ethnicity comparisons, Hispanic/Latine = 0; access to RULER = 1; N_{children} = 1051, $N_{\text{classrooms}}$ = 95, $N_{\text{preschools}}$ = 19; significance was evaluated by z-score; * $p < 0.05$, *** $p < 0.001$.

Table 5. Random effects by model.

Variance Components	Unconditional	Unconditional Growth	Unconditional Growth with Random Slopes	Conditional Growth with Random Slopes	Moderated Growth with Random Slopes
Residual σ^2	39.24 *** (1.43)	30.04 *** (0.94)	25.68 *** (1.27)	26.88 *** (4.32)	26.87 *** (4.33)
Between-child Intercept Variance $\tau_{\pi0000}$	56.98 *** (2.16)	59.14 *** (2.11)	60.89 *** (2.98)	53.58 *** (5.66)	53.80 *** (5.65)
Between-child Time-slope $\tau_{\pi1000}$			13.86 *** (1.61)	9.69 (7.53)	10.01 (7.28)
Child Intercept-Time-slope Correlation $\tau_{\pi0000,\ \pi1000}$			−.20 (0.16)	−0.09 (0.25)	−0.10 (0.24)
Between-classroom Variance $\tau_{\beta0000}$	13.44 * (3.51)	14.37 *** (3.51)	13.03 ** (4.70)	11.43 (7.22)	10.50 * (6.08)
Between-classroom Intercept Time-slope $\tau_{\beta1000}$			5.15 *** (1.00)	8.30 *** (1.78)	5.27 ** (2.08)
Classroom Intercept-Time-slope Correlation $\tau_{\beta0000}$, $\tau_{\beta1000}$			0.11 (0.37)	−0.17 (1.01)	0.25 (1.24)
Between-preschool Intercept Variance $\tau_{\psi0000}$	37.33 *** (4.58)	38.38 *** (4.86)	43.29 *** (5.46)	30.27 *** (9.12)	30.96 *** (9.31)
Between-preschool Time-slope $\tau_{\psi1000}$			6.20 ** (2.20)	8.75 ** (2.88)	11.28 *** (3.29)
Preschool Intercept-Time-slope Correlation $\tau_{\psi0000}$, $\tau_{\psi1000}$			−0.71 *** (0.18)	−0.49 * (.29)	−0.51 ** (0.21)
Model Fit					
df	5	6	12	21	22
AIC	29,430.89	28,549.06	28460.47	24,552.32	24,546.54
BIC	29,460.38	28,584.45	28531.26	24,656.43	24,655.61
χ^2		883.83 ***	100.59 ***	3926.15 ***	7.78 **
pseudo-R^2_{within}		0.07	0.07	0.18	0.20
pseudo-R^2_{between}		0.04	0.04	0.27	0.27

Note. Random effects were estimated using restricted maximum likelihood; values in parentheses are standard errors; variance components are in standard deviation units; pseudo-R^2_{within} represents the proportion of variance explained by the fixed effects; pseudo-R^2_{between} represents the proportion of variance explained relative to the unconditional model; AIC = Akaike information criterion; BIC = Bayesian information criterion; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The early literacy growth rate changed very little with the addition of covariates. The covariates and access to RULER also collectively reduced the early-literacy child, classroom, and preschool intercept and growth covariances from −0.20, 0.11, and −0.70 to −0.09, −0.17, and −0.49, respectively, indicating that these factors accounted for the association in covariation between the child, classroom, and preschool intercept variance and between the child, classroom, and preschool growth rate variance.

3.2. Growth in Early Literacy Associated with RULER Access

Finally, the conditional growth moderation with random slopes model was run fitting the data statistically significantly better with 27% total variance explained and 20% explained by the fixed effects themselves (see Table 5). Access to RULER was statistically significantly associated with greater early literacy growth in children compared with children in preschools with no access to RULER, $\gamma_{1001} = 9.21$, $p < 0.001$. Comparison of simple slopes revealed a stronger growth rate in early literacy skills for children in preschools with access to RULER than for children without access to RULER ($b_{\text{RULER Access}} = 33.71$, $b_{\text{No RULER Access}} = 24.52$, $p_{\text{slopes}} < 0.001$) (see also Figure 1). The end-of-year difference was equivalent to 0.25 standard deviations, 95% CI [0.14, 0.40], which is a small effect using Cohen's conventions for tightly controlled laboratory studies [82]. Converted to Cohen's *U3*, the improvement index was +10, meaning that the average preschooler scoring at the 50th percentile in schools without access to RULER would perform at the 60th percentile had they been in schools with access to RULER.

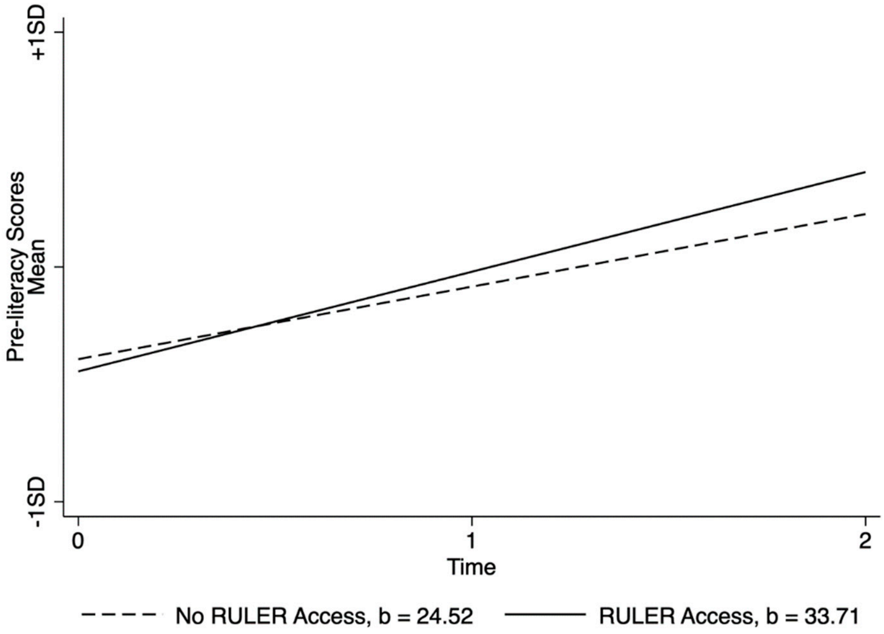

Figure 1. Simple slope analyses comparing children in preschools without access to RULER to those in preschools with access to RULER.

Using the formula by Frank and colleagues [86], 37% of the RULER average treatment effect would have to be due to bias to invalidate the results, equivalent to 391 of the observations or 130 children over the three timepoints in the access-to-RULER group scoring the same as the children in the no-access-to-RULER group for an effect size of = 0.06 standard deviations (i.e., not statistically significant at $\alpha = 0.05$). Using the formula by Frank [76], it was also found that an unobserved confounding variable or combination of variables would have to exhibit a correlation of 0.20 with access to RULER and 0.20 with early literacy skills to invalidate the inference. For comparison, the strongest association with access to RULER was assessment version with $r = 0.20$, with corresponding rs = 0.43–0.46 for associations with early literacy. Altogether, this means that an unobserved confound or combination of confounds would have to be at least as important as the version of the PELI to invalidate the inference.

Finally, post-hoc analyses of each PELI subscale were run to unpack the omnibus effect of RULER. Results revealed that the omnibus effect of RULER access on children's early literacy was mostly due to growth in Vocabulary/Oral Language and Phonological Awareness, equivalent to end-of-the year differences in standard deviation units of 0.20, 95% CI [0.07, 0.33] and 0.24, 95% CI [0.00, 0.48], respectively. Equivalent effects for Alphabet Knowledge, 0.03, 95% CI [−0.25, 0.28]) and Comprehension, 0.17, 95% CI [−0.11, 0.44]) were not statistically different from zero.

4. Discussion

The current study tested whether preschool-level access to RULER was associated with young children's growth in early literacy skills using a quasi-experimental, community-level design. We hypothesized that access to RULER's high-quality SEL programming would be associated with children's development of early literacy skills based on theory and meta-analyses across pre-K–12 settings as well as RULER's specific approach to SEL [1,3,8].

Multilevel growth modeling accounting for variance associated with gender, age, race and ethnicity, family income, classroom-level variance, and preschool-level variance revealed that children in preschools with access to RULER grew significantly more in their early literacy skills than children without access to RULER resulting in an end-of-year difference equivalent to about one-fourth of a standard deviation. The findings were robust across gender, age, family income, race, ethnicity, and omitted variables and were consistent with Lonigan and Shanahan's [60] analysis of literacy-focused interventions, wherein age, SES, and race did not relate to intervention effectiveness. Further, children at preschools with access to RULER ended up better over the school year across three timepoints, a pattern suggesting a possible causal association given our multi-timepoint design [75].

The size of the effect was slightly smaller than Durlak and colleagues' 2011 meta-analytic estimate of 0.27 for the effect of access to SEL programming on K–12 academic outcomes, which included but was not limited to literacy [3], but larger than the estimate of 0.18 reported by Blewitt and colleagues in their 2018 meta-analysis of SEL programming in preschool settings [1]. For a community- or district-level educational effect of a universal program, a difference of 0.25 standard deviations is quite meaningful [87]. Lipsey and colleagues (2012) found that, among 259 studies, whole-school, universal curriculum approaches like RULER typically yield effect sizes between 0.11 and 0.13 [88]. For comparison, in 105 randomized controlled trials on reading, the average effect size was 0.07 standard deviations [89]. Effect sizes are typically larger in studies using researcher-made measures aligned with the intervention and larger in studies that only test high compliance [90,91]. Our results are therefore robust given our use of independent assessments and our intent-to-treat analytical approach. Effect sizes are also typically larger for younger students compared with older students [92]. It is worth noting that we found evidence for RULER's impact on children's growth in early literacy despite imperfect data collection in an arguably more ecologically valid study. Unlike the many documented replication challenges plaguing education and psychological research, there is reason to believe that our results may be replicable in other communities [93,94].

When post-hoc analyses were run for each early literacy subscale, results revealed that the omnibus effect of RULER on children's growth in early literacy was mostly due to growth in Vocabulary/Oral Language and Phonological Awareness but not Alphabet Knowledge or Comprehension in the access-to-RULER group. Lonigan and colleagues (2000) found significant variation in young children's phonological sensitivity (i.e., sensitivity and manipulation of oral language) despite it being highly stable in early elementary school and predictive of other early literacy components [28]. The authors suggested that preschool-aged children may especially benefit from practices that formally and informally enhance phonological sensitivity, which is supported by meta-analyses showing that phonemic awareness and phonics interventions work well for young children [95,96]. Given the variability in typically developing children, RULER practices may support teachers in meeting children where they are as emergent readers in their phonological and phonemic awareness, in addition to their vocabulary and oral language skills [26].

These findings are consistent with how RULER works and likely influences children's early literacy skills through (1) enhancing educator skills and teacher–child interactions and (2) providing classroom curricula focused on expanding language, integrating and improving literacy instruction, and helping children better understand personal and social narratives. These mechanisms include an emphasis on building and enhancing children's emotion vocabulary using teacher–child interactions and social narratives as well as storybook read-alouds. Given these mechanisms and what has been reported in the literature [50,57], we expected RULER-associated growth in children's Vocabulary/Oral Language as the subscale taps into naming objects from a story and explaining as much as they can about those objects. This finding, in particular, is consistent with existing frameworks describing the relationship between language acquisition and early literacy [25]. Likewise, we expected to see little contribution to the overall impact estimate in Alphabet Knowledge. Thus, enhancing teacher–child interactions and dialogic reading practices via RULER seems to do little for children's letter or word recognition.

However, we did not expect to find RULER-associated growth due to changes in Phonological Awareness, or children's ability to sound out words they hear. There could be a connection between RULER's alignment to NAEYC best practices for embedding language learning into stories that children can relate to and the specific nature of the PELI assessment of Phonological Awareness (e.g., looking at pictures depicting a scene from a story and identifying the sounds associated with the beginning of a target word). However, Medford and McGeown (2016) found that prosocial behavior was a significant predictor of letter-sound knowledge [39]. Perhaps children who developed emotion and social skills by way of RULER were better able to attend and engage in instructional time focused on sounding out words. More research would be needed to explore this mechanism.

Another unexpected result was the low contribution to overall growth from Comprehension (i.e., story prediction and inference and story recall). In their study of RULER's curriculum on upper elementary students' literacy scores, Brackett and colleagues [16] (see also Rivers et al., 2013 [54]) hypothesized that RULER may enhance children's overall understanding of a story by examining the feelings of different characters, including the causes and consequences of their feelings and how they regulated them. Given our contradictory findings, this may not be the case with young children or with early literacy, or at least, the mechanism of RULER may drive understanding of vocabulary (i.e., Vocabulary–Oral Language) more than story details (i.e., Comprehension).

Our findings are further contextualized by other research exploring children's growth in early literacy using the PELI as an indicator of early literacy skills. Overall, our sample of 1051 two-to-five-year-old children started lower than the 1764 three-to-five-year-old children in Kaminski and colleagues' (2014) sample [24]. Kaminski and colleagues (2014) found a raw rate of change of 22.45 in PELI total score [24] compared with the raw rate of change of 13.56 in our sample) and grew less ($M_{\text{Kaminski et al. 2014 T1}}$ = 132.45, $SD_{\text{Kaminski et al. 2014 T1}}$ = 61.15; $M_{\text{current study T1}}$ = 123.00, $SD_{\text{current study T1}}$ = 76.88; $M_{\text{Kaminski et al. 2014 T3}}$ = 199.58, $SD_{\text{Kaminski et al. 2014 T3}}$ = 62.53; $M_{\text{current study T3}}$ =163.69, $SD_{\text{current study T3}}$ = 78.15 However, growth rates in PELI subscales were similar between studies. The differences in baseline may be due to differences in sample characteristics [26,97]. In addition, the lower rates of growth in our sample partially parallel results from data collected between 2010 and 2017 from over two million children showing declines in kindergarten reading skills by about 0.15 standard deviations [26,97]. Kaminski and colleagues' (2014) sample was 77% White and had 23% students of color [24], compared with our much more diverse sample, containing 78% students of color and 22% White students. As well, our sample had a lower average income, with 40% in poverty, compared with 1% in the Kaminski study [24]. Extant research shows that family income and parental education—used to operationalize factors associated with children from less-advantaged backgrounds—often correlate with early literacy skills, like how cultural and environmental factors are associated with home language, race, ethnicity, and even quality of instruction [98–100].

Limitations and Future Directions

Interpretation of the estimated impact of RULER on children's early literacy skills in this quasi-experimental study is not without several limitations. First, as a secondary data analysis within a research-to-practice partnership, data were collected by school district staff, and decisions about data collection were likely not in the interest of controlled research, as is often the case in field studies and studies involving data collected by districts [101]. Myriad unobserved variables provide irrefutable alternative hypotheses in this study. Although we found that the access-to-RULER and no-access-to-RULER groups were equivalent at baseline according to What Works Clearinghouse standards [71], threats to causal inference may be due to observed and unobserved confounding variables (e.g., home language, multilingual exposure and proficiency, and immigration status). Complex associations were not possible to explore in this study—it could be that children in RULER preschools are differentiated in their initial emotional intelligence, which drove the differential growth we attributed to RULER [38,102]. With inverse probability weighting, we are confident in the observed covariate-adjusted balance between the groups of preschools at baseline. However, despite our analysis that unobserved covariates (e.g., home language, multilingual exposure and proficiency, and immigration status) must be at least as strong in association with the PELI as the association with the version of the PELI used (i.e., $>r = 0.43$), it is still possible that our results were influenced by unobserved confounding variables resulting in unobserved selection bias.

Relatedly, no information in this study was provided by the district about the teaching staff. It could be that preschools or classrooms were initially different with respect to teacher characteristics and classroom climate, which then led to the differential growth due to RULER. Some teacher characteristics, such as teacher experience, have been associated with student learning in preschool settings [103–106]. Other teacher-level factors that we were unable to obtain may be important as well, including teachers' existing knowledge of SEL [103]. Still, director characteristics, class size, or even children's executive function also might have contributed to our findings [107–109]. One child-level variable not collected was home language, which is especially important given that the PELI was administered in English [110,111]. Given the large percentage of Hispanic/Latine children in the study, it may be that scores for these children were not an accurate representation of their native- or home-language early literacy skills, even if many children with these backgrounds are in fact proficient in English or both English and Spanish.

Third, future studies, especially secondary data analyses with research-to-practice partnerships, should explore additional available outcomes of interest, especially those encompassing preschoolers' school readiness (i.e., preschool adjustment, approaches to learning, social, emotional, and cognitive skills, and other pre-academic skills). Social and emotional skills, in particular, are proximal to SEL interventions, with impact explored into kindergarten and beyond, especially in relation to other important long-term outcomes, such as well-being [1,112].

Fourth, outcomes often vary by fidelity of implementation [3,113], and high-quality implementation data were not collected in this study beyond RULER implementation team members' beliefs in their confidence using RULER and in the value of RULER. Based on this limitation, this study used an intent-to-treat design where all preschools and all classrooms were included despite potential variation in compliance. The analyses in this study thus focused on RULER access and not RULER use. Certainly, RULER's impact on young children's early literacy skills could have been due to several mechanisms that make up RULER's theory of change [8], which is likely given meta-analyses that point to variation in SEL impact due to how programs are implemented [3]. It is also possible there was variation in literacy-related activities in the preschools that did and did not implement RULER. Exploring this variation and mechanisms will be important in future studies.

5. Conclusions

We estimated an effect size of 0.25 standard deviations' difference at the end of the year in early literacy scores due to preschool-level access to RULER. In a community-level context for a universal SEL program, this effect size is quite meaningful, especially considering that language and literacy instruction in general tends to be of lower quality in preschool classrooms [48]. Although there were several limitations, the results of this study bolster the evidence base for RULER as a universal approach to SEL with elements that support the whole child through integrative pedagogical approaches (see Brackett et al., 2019 [8] for a review).

Author Contributions: Conceptualization, C.S.B.; methodology, C.S.B.; software, C.S.B.; validation, C.S.B., O.M. and E.D.; formal analysis, C.S.B.; investigation, C.S.B.; resources, C.S.B.; data curation, E.D.; writing—original draft preparation, C.S.B.; writing—review and editing, O.M. and E.D.; visualization, C.S.B.; supervision, C.S.B.; project administration, C.S.B. and O.M.; funding acquisition, C.S.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Grossman Family Foundation.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Yale University Institutional Review Board (protocol #2000023481 and date of approval: 12 July 2018).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Participant confidentiality and our agreement with the school district prohibits publicizing the raw data for this study.

Acknowledgments: The authors are forever indebted to the tireless early childhood educators in Norwalk who bring learning alive and are essential to young children's learning. The authors also want to express gratitude to Paula Palermo, Data Director of Norwalk ACTS and Jamie Zyla, the Norwalk School Readiness Coordinator, and Norwalk Public Schools, for their support on this project and commitment to children and families in the community. We also thank Marc Brackett for his steadfast support of this work.

Conflicts of Interest: Craig Bailey is a co-developer of RULER and works full time at the Yale Child Study Center. Olivia Martinez works full-time at the Yale Child Study Center. Elizabeth DiDomizio has no conflicts of interest to disclose.

References

1. Blewitt, C.; Fuller-Tyszkiewicz, M.; Nolan, A.; Bergmeier, H.; Vicary, D.; Huang, T.; McCabe, P.; McKay, T.; Skouteris, H. Social and Emo-tional Learning Associated with Universal Curriculum-Based Interventions in Early Childhood Education and Care Centers. *JAMA Netw. Open* **2018**, *1*, e185727. [CrossRef]
2. Corcoran, R.P.; Cheung, A.C.; Kim, E.; Xie, C. Effective universal school-based social and emotional learning programs for improving academic achievement: A systematic review and meta-analysis of 50 years of research. *Educ. Res. Rev.* **2018**, *25*, 56–72. [CrossRef]
3. Durlak, J.A.; Weissberg, R.P.; Dymnicki, A.B.; Taylor, R.D.; Schellinger, K.B. The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions. *Child. Dev.* **2011**, *82*, 405–432. [CrossRef] [PubMed]
4. Murano, D.; Sawyer, J.; Lipnevich, A. A Meta-Analytic Review of Preschool Social and Emotional Learning Interventions. *Rev. Educ. Res.* **2020**, *90*, 227–263. [CrossRef]
5. Payton, J.; Weissberg, R.; Durlak, J.; Dymnicki, A.; Taylor, R.; Schellinger, K.; Pachan, M. The positive impact of social and emotional learning for kindergarten to eighth-grade students: Findings from three scientific reviews. *Child. Care Pract.* **2008**, *15*, 261–278.
6. Doyle, B.G.; Bramwell, W. Promoting Emergent Literacy and Social-Emotional Learning through Dialogic Reading. *Read. Teach.* **2011**, *59*, 554–564. [CrossRef]
7. Kozak, S.; Recchia, H. Reading and the Development of Social Understanding: Implications for the Literacy Classroom. *Read. Teach.* **2018**, *72*, 569–577. [CrossRef]

8. Brackett, M.A.; Bailey, C.S.; Hoffmann, J.D.; Simmons, D.N. RULER: A Theory-Driven, Systemic Approach to Social, Emotional, and Academic Learning. *Educ. Psychol.* **2019**, *54*, 144–161. [CrossRef]
9. Hennessey, A.; Humphrey, N. Can social and emotional learning improve children's academic progress? Findings from a randomised controlled trial of the Promoting Alternative Thinking Strategies (PATHS) curriculum. *Eur. J. Psychol. Educ.* **2019**, *35*, 751–774. [CrossRef]
10. Jennings, P.A.; Greenberg, M.T. The Prosocial Classroom: Teacher Social and Emotional Competence in Relation to Student and Classroom Outcomes. *Rev. Educ. Res.* **2009**, *79*, 491–525. [CrossRef]
11. Wenz-Gross, M.; Yoo, Y.; Upshur, C.C.; Gambino, A.J. Pathways to Kindergarten Readiness: The Roles of Second Step Early Learning Curriculum and Social Emotional, Executive Functioning, Preschool Academic and Task Behavior Skills. *Front. Psychol.* **2018**, *9*, 1886. [CrossRef] [PubMed]
12. CASEL. Criteria for Inclusion. 2021. Available online: https://pg.casel.org/for-providers/#criteria_inclusion (accessed on 15 June 2021).
13. Mayer, J.; Salovey, P. What is emotional intelligence? In *Emotional Development and Emotional Intelligence: Educational Implications*; Basic Books: New York, NY, USA, 1997; pp. 3–34.
14. Mayer, J.; Salovey, P.; Caruso, D. Emotional intelligence as zeitgeist, as personality, and as a mental ability. In *The Handbook of Emotional Intelligence: Theory, Development, Assessment, and Application at Home, School, and in the Workplace*; Bar-On, R., Parker, J., Eds.; Jossey-Bass: San Francisco, CA, USA, 2000; pp. 92–117.
15. Zeidner, M.; Matthews, G. Emotional Intelligence: A Brief Historical Introduction. In *An Introduction to Emotional Intelligence*; Dacre Pool, L., Qualter, P., Eds.; John Wiley & Sons Ltd.: Hoboken, NJ, USA, 2018; pp. 1–17.
16. Brackett, M.A.; Rivers, S.E.; Reyes, M.R.; Salovey, P. Enhancing academic performance and social and emotional competence with the RULER feeling words curriculum. *Learn. Individ. Differ.* **2012**, *22*, 218–224. [CrossRef]
17. Castillo, R.; Fernández-Berrocal, P.; Brackett, M.A. Enhancing teacher effectiveness in Spain: A pilot study of the RULER approach to social and emotional learning. *J. Educ. Train. Stud.* **2013**, *2*, 263–272. [CrossRef]
18. Hagelskamp, C.; Brackett, M.A.; Rivers, S.E.; Salovey, P. Improving Classroom Quality with The RULER Approach to Social and Emotional Learning: Proximal and Distal Outcomes. *Am. J. Community Psychol.* **2013**, *51*, 530–543. [CrossRef]
19. Reyes, M.R.; Brackett, M.A.; Rivers, S.E.; White, M.; Salovey, P. Classroom emotional climate, student engagement, and academic achievement. *J. Educ. Psychol.* **2012**, *104*, 700. [CrossRef]
20. Rivers, S.; Brackett, M.; Reyes, M.; Elbertson, N.; Salovey, P. Improving the social and emotional climate of classrooms: A clustered randomized controlled trial testing The RULER Approach. *Prev. Sci.* **2013**, *14*, 77–87. [CrossRef]
21. Rivers, S.E.; Brackett, M.A. Achieving Standards in the English Language Arts (and More) Using the RULER Approach to Social and Emotional Learning. *Read. Writ. Q.* **2011**, *27*, 75–100. [CrossRef]
22. NoVo Foundation & Education First. SEL in Action: The 2017–2018 Social and Emotional Learning Innovation Awards for Teachers (Electronic Book). 2018. Available online: https://selforteachers.org/wp-content/uploads/2018/10/Social-and-Emotional-Learning-Teacher-Awards-2018-Digital.pdf (accessed on 15 January 2023).
23. Duncan, G.J.; Dowsett, C.J.; Claessens, A.; Magnuson, K.; Huston, A.C.; Klebanov, P.; Pagani, L.S.; Feinstein, L.; Engel, M.; Brooks-Gunn, J.; et al. School readiness and later achievement. *Dev. Psychol.* **2007**, *43*, 1428–1446. [CrossRef]
24. Kaminski, R.; Abbott, M.; Aguayo, K.; Latimer, R.; Good, R. The Preschool Early Literacy Indicators: Validity and Benchmark Goals. *Top. Early Child. Spec. Educ.* **2014**, *34*, 71–82. [CrossRef]
25. Rhyner, P.; Haebig, E.; West, K. Understanding frameworks for the emergent literacy stage. In *Emergent Literacy and Language Development: Promoting Learning in Early Childhood*; Rhyner, P., Ed.; Guilford Press: New York, NY, USA, 2009; pp. 5–35.
26. Lonigan, C.; Schatschneider, C.; Westberg, L. Identification of children's skills and abilities linked to later outcomes in reading, writing, and spelling. In *Developing Early Literacy: Report of the National Early Literacy Panel*; National Institute for Literacy: Washington, DC, USA, 2008; pp. 55–106.
27. Whitehurst, G.; Lonigan, C. Child development and emergent literacy. *Child. Dev.* **1998**, *69*, 848. [CrossRef]
28. Lonigan, C.J.; Burgess, S.R.; Anthony, J.L. Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. *Dev. Psychol.* **2000**, *36*, 596–613. [CrossRef] [PubMed]
29. Wagner, R.; Torgesen, J.; Rashotte, C. Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Dev. Psychol.* **1994**, *30*, 73–87. [CrossRef]
30. Wagner, R.K.; Torgesen, J.K.; Rashotte, C.A.; Hecht, S.A.; Barker, T.A.; Burgess, S.R.; Donahue, J.; Garon, T. Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. *Dev. Psychol.* **1997**, *33*, 468–479. [CrossRef] [PubMed]
31. Foster, W.A.; Miller, M. Development of the Literacy Achievement Gap: A Longitudinal Study of Kindergarten Through Third Grade. *Lang. Speech Hear. Serv. Sch.* **2007**, *38*, 173. [CrossRef] [PubMed]
32. Markussen-Brown, J.; Juhl, C.B.; Piasta, S.B.; Bleses, D.; Højen, A.; Justice, L.M. The effects of language- and literacy-focused professional development on early educators and children: A best-evidence meta-analysis. *Early Child. Res. Q.* **2017**, *38*, 97–115. [CrossRef]
33. Phillips, L.; Norris, S.; Mason, J. Longitudinal effects of early literacy concepts on reading achievement: A kindergarten intervention and five-year follow-up. *J. Lit. Res.* **1996**, *28*, 173–195. [CrossRef]

34. Piasta, S.B.; Wagner, R.K. Developing Early Literacy Skills: A Meta-Analysis of Alphabet Learning and Instruction. *Read. Res. Q.* **2010**, *45*, 8–38. [CrossRef]
35. Rimm-Kaufman, S.E.; Larsen, R.A.; Baroody, A.E.; Curby, T.W.; Ko, M.; Thomas, J.B.; Merritt, E.G.; Abry, T.; DeCoster, J. Efficacy of the Responsive Classroom Approach: Results From a 3-Year, Longitudinal Randomized Controlled Trial. *Am. Educ. Res. J.* **2014**, *51*, 567–603. [CrossRef]
36. Zins, J.; Bloodworth, M.; Weissberg, R.; Walberg, H. The scientific base linking social and emotional learning to school success. In *Building Academic Success on Social and Emotional Learning: What Does the Research Say?* Teachers College Press: New York, NY, USA, 2004; pp. 3–22.
37. Zigler, E. The cognitive child versus the whole child: Lessons from forty years of head start. In *Play = Learning: How Play Motivates and Enhances Cognitive and Social-Emotional Growth*; Singer, D., Golinkoff, R., Hirsch-Pasek, K., Eds.; Oxford University Press: New York, NY, USA, 2006.
38. Curby, T.W.; Brown, C.A.; Bassett, H.H.; Denham, S.A. Associations Between Preschoolers' Social-Emotional Competence and Preliteracy Skills. *Infant. Child. Dev.* **2015**, *24*, 549–570. [CrossRef]
39. Medford, E.; Mcgeown, S. Social, emotional and behavioural influences on young children's pre-reading and word reading development. *J. Appl. Dev. Psychol.* **2016**, *43*, 54–61. [CrossRef]
40. Grimm, K.J.; Steele, J.S.; Mashburn, A.J.; Burchinal, M.; Pianta, R.C. Early behavioral associations of achievement trajectories. *Dev. Psychol.* **2010**, *46*, 976–983. [CrossRef] [PubMed]
41. Rabiner, D.; Coiek, J. Early Attention Problems and Children's Reading Achievement: A Longitudinal Investigation. *Child. Adolesc. Psychiatry* **2000**, *39*, 859–867. [CrossRef]
42. Romano, E.; Babchishin, L.; Pagani, L.S.; Kohen, D. School readiness and later achievement: Replication and extension using a nationwide Canadian survey. *Dev. Psychol.* **2010**, *46*, 995–1007. [CrossRef] [PubMed]
43. Rosenblatt, J.L.; Elias, M.J. Dosage Effects of a Preventive Social-Emotional Learning Intervention on Achievement Loss Associated with Middle School Transition. *J. Prim. Prev.* **2008**, *29*, 535–555. [CrossRef] [PubMed]
44. Trentacosta, C.; Izard, C. Kindergarten children's emotion competence as a predictor of their academic competence in first grade. *Emotion* **2007**, *7*, 77–88. [CrossRef] [PubMed]
45. Dunlap, G.; Strain, P.S.; Fox, L.; Carta, J.J.; Conroy, M.; Smith, B.J.; Kern, L.; Hemmeter, M.L.; Timm, M.A.; Mccart, A.; et al. Prevention and Intervention with Young Children's Challenging Behavior: Perspectives regarding Current Knowledge. *Behav. Disord.* **2006**, *32*, 29–45. [CrossRef]
46. Jones, D.; Greenberg, M.; Crowley, M. Early social-emotional functioning and public health: The relationship between kindergarten social competence and future wellness. *Am. J. Public. Health* **2015**, *105*, 2283–2290. [CrossRef]
47. Fantuzzo, J.; Bulotsky, R.; McDermott, P.; Mosca, S.; Lutz, M.N. A Multivariate Analysis of Emotional and Behavioral Adjustment and Preschool Educational Outcomes. *Sch. Psychol. Rev.* **2003**, *32*, 185–203. [CrossRef]
48. Thompson, R.; Raikes, H. The social and emotional foundations of school readiness. In *Social and Emotional Health in Early Childhood: Building Bridges Between Services and Systems*; Knitzer, J., Kaufmann, R., Perry, D., Eds.; Paul H. Brookes Publishing Co.: Baltimore, MD, USA, 2008; pp. 13–35.
49. Graziano, P.; Reavis, R.; Keane, S.; Calkins, S. The role of emotion regulation in children's early academic success. *J. Sch. Psychol.* **2007**, *45*, 3–19. [CrossRef]
50. Howes, C.; Burchinal, M.; Pianta, R.; Bryant, D.; Early, D.; Clifford, R.; Barbarin, O. Ready to learn? Children's pre-academic achievement in pre-Kindergarten programs. *Early Child. Res. Q.* **2008**, *23*, 27–50. [CrossRef]
51. Justice, L.M.; Mashburn, A.J.; Hamre, B.K.; Pianta, R.C. Quality of language and literacy instruction in preschool classrooms serving at-risk pupils. *Early Child. Res. Q.* **2008**, *23*, 51–68. [CrossRef] [PubMed]
52. Bierman, K.L.; Domitrovich, C.E.; Nix, R.L.; Gest, S.D.; Welsh, J.A.; Greenberg, M.T.; Blair, C.; Nelson, K.E.; Gill, S. Promoting Academic and So-cial-Emotional School Readiness: The Head Start REDI Program. *Child. Dev.* **2008**, *79*, 1802–1817. [CrossRef] [PubMed]
53. Deacon, E.; van Rensburg, E. Enhancing emotional and social competence in a group of South-African school be-ginners: A preliminary study. *J. Psychol. Afr.* **2012**, *22*, 677–680. [CrossRef]
54. Rivers, S.E.; Tominey, S.L.; O'Bryon, E.C.; Brackett, M.A. Special Edition article: Developing emotional skills in early childhood settings using Preschool RULER. *Psychol. Educ. Rev.* **2013**, *37*, 19–25. [CrossRef]
55. Chapman, J.; Tunmer, W.; Prochnow, J. Early reading-related skills and performance, reading self-concept, and the development of academic self-concept: A longitudinal study. *J. Ed. Psychol.* **2000**, *92*, 703–708. [CrossRef]
56. Vygotsky, L. Interaction Between Learning and Development. In *Readings on the Development of Children*; Gauvain, M., Cole, M., Eds.; Scientific American Books: New York, NY, USA, 1978; pp. 34–40.
57. Wasik, B.; Hindman, A. Improving vocabulary and pre-literacy skills of at-risk preschoolers through teacher professional development. *J. Educ. Psychol.* **2011**, *103*, 455–469. [CrossRef]
58. McCormick, M.; Cappella, E.; O'Connor, E.; McClowry, S. Social-emotional learning and academic achievement: Using causal methods to explore classroom-level mechanisms. *AERA Open* **2015**, *1*, 2332858415603959. [CrossRef]
59. National Council of Teachers of English. NCTE/IRA Standards for the English Language Arts. 2022. Available online: https://ncte.org/resources/standards/ncte-ira-standards-for-the-english-language-arts/ (accessed on 15 October 2022).

60. Lonigan, C.; Shanahan, T. *Developing Early Literacy: Report of the National Early Literacy Panel*; National Institute for Literacy: Washington, DC, USA, 2008.
61. Noble, C.; Sala, G.; Peter, M.; Lingwood, J.; Rowland, C.; Gobet, F.; Pine, J. The impact of shared book reading on children's language skills: A meta-analysis. *Educ. Res. Rev.* **2019**, *28*, 100290. [CrossRef]
62. Schonfeld, D.J.; Adams, R.E.; Fredstrom, B.K.; Weissberg, R.P.; Gilman, R.; Voyce, C.; Tomlin, R.; Speese-Linehan, D. Cluster-randomized trial demonstrating impact on academic achievement of elementary social-emotional learning. *Sch. Psychol. Rev.* **2015**, *30*, 406–420. [CrossRef]
63. Swanson, E.A.; Vaughn, S.; Wanzek, J.; Petscher, Y.; Heckert, J.; Cavanaugh, C.; Kraft, G.; Tackett, K.K. A Synthesis of Read-Aloud Interventions on Early Reading Outcomes among Preschool through Third Graders at Risk for Reading Difficulties. *J. Learn. Disabil.* **2011**, *44*, 258–275. [CrossRef]
64. United States Census Bureau. Population Estimates. 2017. Available online: https://www.census.gov/quickfacts/fact/table/norwalkcityconnecticut,norwalkcityohio# (accessed on 1 June 2021).
65. Lechner, M. The Estimation of Causal Effects by Difference-in-Difference Methods. *Found. Trends Econom.* **2010**, *4*, 165–224. [CrossRef]
66. Hesser, H. Modeling individual differences in randomized experiments using growth models: Recommendations for design, statistical analysis and reporting of results of internet interventions. *Internet Interv.* **2015**, *2*, 110–120. [CrossRef]
67. Senn, S. Change from baseline and analysis of covariance revisited. *Stat. Med.* **2006**, *25*, 4334–4344. [CrossRef] [PubMed]
68. Ashbeck, E.L.; Bell, M.L. Single time point comparisons in longitudinal randomized controlled trials: Power and bias in the presence of missing data. *BMC Med. Res. Methodol.* **2016**, *16*, 43. [CrossRef] [PubMed]
69. Castro-Schilo, L.; Grimm, K.J. Using residualized change versus difference scores for longitudinal research. *J. Soc. Pers. Relatsh.* **2018**, *35*, 32–58. [CrossRef]
70. Van Breukelen, G. ANCOVA Versus CHANGE From Baseline in Nonrandomized Studies: The Difference. *Multivar. Behav. Res.* **2013**, *48*, 895–922. [CrossRef]
71. What Works Clearinghouse. *What Works Clearinghouse Standards Handbook*; Version 4.1; U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance: Washington, DC, USA, 2020.
72. Li, C. Little's Test of Missing Completely at Random. *Stata J. Promot. Commun. Stat. Stata* **2013**, *13*, 795–809. [CrossRef]
73. Robins, J.; Rotnitzky, A. Semiparametric efficiency in multivariate regression models with missing data. *J. Am. Stat. Assoc.* **1995**, *90*, 122–129. [CrossRef]
74. Austin, P.C.; Stuart, E.A. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat. Med.* **2015**, *34*, 3661–3679. [CrossRef]
75. Shadish, W.; Cook, T.; Campbell, D. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*; Houghton, Mifflin and Company: Boston, MA, USA, 2002.
76. Frank, K. Impact of a confounding variable on the inference of a regression coefficient. *Sociol. Methods Res.* **2000**, *29*, 147–194. [CrossRef]
77. Raudenbush, S.; Bryk, A. *Hierarchical Linear Models: Applications and Data Analysis Methods*; Sage Publications, Inc.: Thousand Oaks, CA, USA, 2002.
78. Schielzeth, H.; Forstmeier, W. Conclusions beyond support: Overconfident estimates in mixed models. *Behav. Ecol.* **2008**, *20*, 416–420. [CrossRef]
79. Feingold, A. A Regression Framework for Effect Size Assessments in Longitudinal Modeling of Group Differences. *Rev. Gen. Psychol.* **2013**, *17*, 111–121. [CrossRef] [PubMed]
80. Becker, B. Synthesizing standardized mean-change measures. *Br. J. Math. Stat. Psychol.* **1988**, *41*, 257–278. [CrossRef]
81. Morris, S.; DeShon, R. Combining effect size estimates in meta-analysis with repeated measures and independent groups designs. *Psychol. Methods* **2002**, *7*, 105–125. [CrossRef]
82. Cohen, J. *Statistical Power Analysis*, 2nd ed.; Lawrence Erlbaum Associates, Inc.: Hillsdale, MI, USA, 1988.
83. Lahuis, D.M.; Hartman, M.J.; Hakoyama, S.; Clark, P.C. Explained Variance Measures for Multilevel Models. *Organ. Res. Methods* **2014**, *17*, 433–451. [CrossRef]
84. Snijders, T.; Bosker, R. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*, 2nd ed.; Sage Publications: Thousand Oaks, CA, USA, 2012.
85. Nakagawa, S.; Schielzeth, H. A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods Ecol. Evol.* **2013**, *4*, 133–142. [CrossRef]
86. Frank, K.A.; Maroulis, S.J.; Duong, M.Q.; Kelcey, B.M. What Would It Take to Change an Inference? Using Rubin's Causal Model to Interpret the Robustness of Causal Inferences. *Educ. Eval. Policy Anal.* **2013**, *35*, 437–460. [CrossRef]
87. Bakker, A.; Cai, J.; English, L.; Kaiser, G.; Mesa, V.; Van Dooren, W. Beyond small, medium, or large: Points of consideration when interpreting effect sizes. *Educ. Stud. Math.* **2019**, *102*, 1–8. [CrossRef]
88. Lipsey, M.W.; Puzio, K.; Yun, C.; Hebert, M.A.; Steinka-Fry, K.; Cole, M.W.; Roberts, M.; Anthony, K.S.; Busick, M.D. *Translating the Statistical Representation of the Effects of Education Interventions into More Readily Interpretable Forms*; National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education: Washington, DC, USA, 2012.
89. Fryer, R.G. Chapter 2 – The production of human capital in developed countries: Evidence from 196 randomized field experiments. *Handb. Econ. Field Exp.* **2017**, *2*, 95–322.

90. Cheung, A.; Slavin, R.E. How methodological features affect effect sizes in education. *Educ. Res.* **2016**, *45*, 283–292. [CrossRef]
91. Kraft, M.A. Interpreting effect sizes of education interventions. *Educ. Res.* **2020**, *49*, 241–253. [CrossRef]
92. Hill, C.J.; Bloom, H.S.; Black, A.R.; Lipsey, M.W. Empirical benchmarks for interpreting effect sizes in research. *Soc. Res. Child. Dev.* **2008**, *2*, 172–177. [CrossRef]
93. Makel, M.; Plucker, J. Facts Are More Important Than Novelty: Replication in the Education Sciences. *Educ. Res.* **2014**, *43*, 304–316. [CrossRef]
94. Pashler, H.; Wagenmakers, E. Editors' Introduction to the Special Section on Replicability in Psychological Science: A Crisis of Confidence? *Perspect. Psychol. Sci.* **2012**, *7*, 528–530. [CrossRef] [PubMed]
95. Bus, A.G.; van Ijzendoorn, M.H. Phonological awareness and early reading: A meta-analysis of experimental training studies. *J. Educ. Psychol.* **1999**, *91*, 403–414. [CrossRef]
96. Ehri, L.C.; Nunes, S.R.; Stahl, S.A.; Willows, D.M. Systematic Phonics Instruction Helps Students Learn to Read: Evidence from the National Reading Panel's Meta-Analysis. *Rev. Educ. Res.* **2001**, *71*, 393–447. [CrossRef]
97. Kuhfeld, M.; Soland, J.; Pitts, C.; Burchinal, M. Trends in Children's Academic Skills at School Entry: 2010 to 2017. *Educ. Res.* **2020**, *49*, 403–414. [CrossRef]
98. Byrne, B.; Coventry, W.L.; Olson, R.K.; Samuelsson, S.; Corley, R.; Willcutt, E.G.; Wadsworth, S.; DeFries, J.C. Genetic and environmental influences on aspects of literacy and language in early childhood: Continuity and change from preschool to Grade 2. *J. Neurolinguistics* **2009**, *22*, 219–236. [CrossRef]
99. Heath, S.M.; Bishop, D.V.M.; Bloor, K.E.; Boyle, G.L.; Fletcher, J.; Hogben, J.H.; Wigley, C.A.; Yeong, S.H.M. A Spotlight on Preschool: The Influence of Family Factors on Children's Early Literacy Skills. *PLoS ONE* **2014**, *9*, e95255. [CrossRef]
100. Waldfogel, J. The Role of Out-of-School Factors in the Literacy Problem. *Futur. Child.* **2012**, *22*, 39–54. [CrossRef]
101. He, Y.; Faircloth, B.S.; Hewitt, K.K.; Rock, M.L.; Rodriguez, S.; Gonzalez, L.M.; Vetter, A. Data management and use through research practice partnerships: A literature review. *Educ. Res. Rev.* **2020**, *31*, 100360. [CrossRef]
102. Rhoades, B.L.; Warren, H.K.; Domitrovich, C.E.; Greenberg, M.T. Examining the link between preschool social–emotional competence and first grade academic achievement: The role of attention skills. *Early Child. Res. Q.* **2011**, *26*, 182–191. [CrossRef]
103. Burroughs, N.; Gardner, J.; Lee, Y.; Guo, S.; Touitou, I.; Jansen, K.; Schmidt, W. A Review of the Literature on Teacher Effectiveness and Student Outcomes. In *Teaching for Excellence and Equity*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 7–17.
104. Danisman, S.; Guler, M.; Karadag, E. The effect of teacher characteristics on student achievement: A meta-analysis study. *Croat. J. Educ.* **2018**, *21*, 1367–1398.
105. Graham, L.; White, S.; Cologon, K.; Pianta, R. Do teachers' years of experience make a difference in the quality of teaching? *Teach. Teach. Educ.* **2020**, *96*, 103190. [CrossRef]
106. Stuhlman, M.W.; Pianta, R.C. Profiles of Educational Quality in First Grade. *Elem. Sch. J.* **2009**, *109*, 323–342. [CrossRef]
107. Francis, J.; Barnett, W.S. Relating preschool class size to classroom quality and student achievement. *Early Child. Res. Q.* **2019**, *49*, 49–58. [CrossRef]
108. Lonigan, C.J.; Allan, D.M.; Phillips, B.M. Examining the predictive relations between two aspects of self-regulation and growth in preschool children's early literacy skills. *Dev. Psychol.* **2017**, *53*, 63–76. [CrossRef]
109. Zembat, R.; Kocyigit, S.; Tugluk, M.; Dogan, H. The relationship between the effectiveness of preschools and leadership styles of school managers. *Procedia Soc. Behav. Sci.* **2010**, *2*, 2269–2276. [CrossRef]
110. Barnett, W.S.; Yarosz, D.J.; Thomas, J.; Jung, K.; Blanco, D. Two-way and monolingual English immersion in preschool education: An experimental comparison. *Early Child. Res. Q.* **2007**, *22*, 277–293. [CrossRef]
111. Páez, M.M.; Tabors, P.O.; López, L.M. Dual language and literacy development of Spanish-speaking preschool children. *J. Appl. Dev. Psychol.* **2007**, *28*, 85–102. [CrossRef]
112. Taylor, J.J.; Buckley, K.; Hamilton, L.S.; Stecher, B.M.; Read, L.; Schweig, J. Choosing and using SEL competency assessments. 2018. Available online: https://measuringsel.casel.org/pdf/practitioner-guidance.pdf (accessed on 1 June 2021).
113. Durlak, J.A.; DuPre, E.P. Implementation Matters: A Review of Research on the Influence of Implementation on Program Outcomes and the Factors Affecting Implementation. *Am. J. Community Psychol.* **2008**, *41*, 327–350. [CrossRef] [PubMed]