Integrating Literacy and Science Instruction in Kindergarten: Results From the Efficacy Study of *Zoology One*

Abigail M. Gray, Philip M. Sirinides, Ryan E. Fink & A. Brooks Bowden

Published online: 21 Jul 2021.

Submit your article to this journal 🗗

Article views: 4998

View related articles 🗗

View Crossmark data 🗗

Routledge
Taylor & Francis Group

INTERVENTION, EVALUATION, AND POLICY STUDIES

OPEN ACCESS   Check for updates

Integrating Literacy and Science Instruction in Kindergarten: Results From the Efficacy Study of *Zoology One*

Abigail M. Gray[a], Philip M. Sirinides[b], Ryan E. Fink[a] and A. Brooks Bowden[a] (iD)

[a]Graduate School of Education, University of Pennsylvania, Philadelphia, Pennsylvania, USA; [b]Institute of State and Regional Affairs, Pennsylvania State University, Middletown, Pennsylvania, USA

ABSTRACT

This study examines the efficacy, cost, and implementation of an integrated science and literacy curriculum for kindergarten. The study was conducted in a large urban district and included 1,589 students in 71 classrooms in 21 schools. The research includes a multi-site cluster-randomized controlled trial and mixed-methods cost and implementation studies. Analysis revealed significant impacts on comprehension, letter-naming fluency, and motivation to read. No main impacts were observed on decoding, word identification, or writing; however, exploratory analysis revealed that students whose teachers implemented the treatment with fidelity performed statistically significantly better in writing and decoding. The cost to produce the observed effects was estimated at $480 per student, two-thirds of which was borne by the school. Despite this cost, treatment classrooms achieved savings by using an average of three fewer instructional programs than control classrooms. Teachers reported positive effects from the integrated curriculum on student engagement, learning, and behavior.

Introduction

Decades of research demonstrate that students who struggle with reading in the earliest grades often suffer long-term academic consequences (Cunningham & Stanovich, 1997; Ozernov-Palchik et al., 2017; Stanley et al., 2018). A primary goal of early education must therefore be to effectively and efficiently help young children become readers. This study focuses on an innovative approach to improving literacy learning in kindergarten: the integration of literacy and science instruction. More specifically, the study investigates the hypothesis that the effects of evidence-based literacy instruction on young students' learning are magnified by the infusion of science content.

The study's setting is a large, economically challenged urban district. Following decades of troublingly low literacy levels among its students, the School District of Philadelphia (SDP) has made early literacy a priority. This study offers a comparison of business-as-usual literacy instruction in SDP schools and a curriculum that includes science integration. The research described here—which includes a multi-site, longitudinal randomized controlled trial (RCT) designed to produce impact, cost and implementation findings—supports causal inference for intervention impacts, in a context with urgent need.

Our work is informed by a growing body of consonant research on early science exposure, early literacy instruction, integrated curricula, and the role of motivation in learning to read. The need for early science instruction is increasingly recognized as shortages of entrants to science, technology, engineering, and math (STEM) careers grow pronounced (Stine, 2009). The U.S. lags behind many other nations in numbers of citizens earning degrees in STEM fields (Okahana et al., 2016), and this gap begins early: Large-scale assessments reveal that U.S. students consistently score lower in science than students in other advanced countries beginning in elementary school (National Center for Educational Statistics, 2019). Additionally, despite indications that young children have natural scientific proclivity (Clements & Sarama, 2016), motivation for science learning tends to diminish as students get older (Anderman & Young, 1994; Vedder-Weiss & Fortus, 2011). Shortages in STEM fields are particularly stark among women and minorities, and gender and racial gaps emerge in science achievement as early as fourth grade (National Science Foundation, 2019). Other research indicates that science knowledge increases with the volume of instructional time devoted to the subject (Curran & Kitchin, 2019), and that resource disparities disproportionately affect minority students, further contributing to science achievement gaps (Curran & Kellogg, 2016). Indeed, many female and minority students have already ruled out STEM careers by the end of elementary school (Wendt et al., 2018).

These trends suggest that science exposure is critical for all children, and particularly those unlikely to pursue careers in STEM. Furthermore, exposure must begin early; knowledge gaps evident in kindergarten contribute to science achievement disparities in subsequent grades (Morgan et al., 2016). Early science exposure establishes foundational scientific concepts students can build on later and develops children's motivation for science learning (Henrichs & Leseman, 2014; Sahin et al., 2014). However, the opportunity to provide this early exposure is often missed: One study found that kindergarteners receive an average of only 2.3 min of science instruction per day (Wright & Neuman, 2014), while another found that average instructional time in science for kindergarten through third-grade classrooms was 19 min per day (Banilower et al., 2013).

Integrated Literacy and Science Curricula

The benefits of literacy and science integration are substantiated by two decades of research focused on older students (Cervetti et al., 2012; Goldschmidt & Jung, 2011; Guthrie et al., 1999; Guthrie & Humenick, 2004; Pearson et al., 2010; Romance & Vitale, 2001; Slavin et al., 2014; Shanahan et al., 2010; Wigfield et al., 2008). Curricular

integration is efficacious, in part, because the parallel cognitive skills required by literacy and science allow learning in one domain to support the other. Cervetti et al. (2006) note that "science and literacy are more than supportive and synergistic, they are in fact *isomorphic*" (p. 9), and that the strategies required for reading comprehension parallel the inquiry strategies science demands (Baker, 1991; Cervetti et al., 2006; Cervetti et al., 2012; Padilla et al., 1991).

In a quasi-experiment, Romance and Vitale (2001) examined the impacts of the literacy and science curriculum *Science IDEAS* on reading and science learning of 540 third- through fifth-grade students, and observed significant positive effects on achievement in reading and science (Romance & Vitale, 1992, 2001). Another quasi-experimental study examined the impact of *Concept-Oriented Reading Instruction (CORI)*, an integrated literacy and science program, on fifth-grade students' achievement (Guthrie & Wigfield, 2009). This study revealed statistically significant positive effects on reading comprehension (ES = .59) and science content knowledge (ES = 1.59), and positive though not statistically significant impacts in several other domains. A 2011 RCT study examined the impacts of *Science IDEAS* on first and second graders' achievement in science and reading (Vitale & Romance, 2011), again finding significant positive effects in both domains. Additional advantages may include the development of background knowledge and vocabulary, the activation of prior knowledge, and the cultivation of curiosity and motivation to read (Duke et al., 2011)—all of which predict long-term achievement (French, 2004; Grissmer et al., 2010; Strickland & Riley-Ayers, 2006).

Despite the promise of integrated literacy and science curricula, no prior study has rigorously examined the impacts of this approach in kindergarten. In a small quasi-experimental study, Wright and Gotwals (2017) examined children's (n = 147) oral language outcomes after receiving four weeks of an integrated science and disciplinary language and literacy curriculum called *SOLID Start*. Children who received the curriculum outperformed children in the control group in their use of vocabulary in a science context, knowledge of receptive science vocabulary, and their ability to make claims and give evidence-based supports. Another recent study (n = 120) investigated the impacts of *LINKS*, an integrated science and literacy curriculum implemented in kindergarten for ten weeks (Kurz, 2018). While this study found a significant impact on the treatment group's understanding of science and depth of science knowledge, it is limited by its short implementation period and small sample size and does not examine impacts on literacy.

Effective Literacy Instruction

To fully assess the promise of integrated curricula, it is critical to select an intervention that incorporates evidence-based best practices in literacy instruction. Research has produced clear insights about which instructional practices benefit early readers most, and these insights are reflected in the program model for the *Zoology One* curriculum. Program elements include code-focused instruction emphasizing alphabet knowledge, phonics, and phonological awareness. Code-focused instruction is based in theory that asserts that understanding of the alphabetic principle—the recognition that sounds are represented by letters, which in turn comprise words—is a critical early step in the development of reading fluency and comprehension (Ehri, 1991, 2005; Juel, 1991;

Stanovich, 1986). Significant evidence supports the effectiveness of code-focused reading instruction for beginning readers. The Report of the National Early Literacy Panel (NELP, 2009) presents a meta-analysis of 83 experimental or quasi-experimental studies with treatment-control equivalence at baseline. This analysis found that "code-focused interventions usually had moderate to large effects both on measures of conventional literacy (i.e., reading, spelling) and measures of precursor literacy skills (e.g., phonological awareness, alphabet knowledge)" (Lonigan et al., 2008, p. 109).

Evidence further suggests that high-volume print exposure yields important benefits for beginning readers (Jorm & Share, 1983; Share, 1995). This includes both complex-text exposure via teacher read-alouds or shared reading, and teacher-supported independent reading practice in leveled, high-interest texts (Duke, 2000; Miller & Moss, 2013; Reutzel et al., 2008; Topping et al., 2007). Correlational research spanning decades associates reading practice with long-term reading proficiency (Allington, 1977; Cipielewski & Stanovich, 1992; Donahue et al., 2001; Garan & DeVoogd, 2008; Samuels & Wu, 2003). In their meta-analysis of studies on reading volume, Mol and Bus (2011) found that 12% of language proficiency in preschool and kindergarten was explained by print exposure alone, and that this effect was accretive: The explanatory power of print exposure on reading achievement increases as students age, suggesting that the benefits of early, high-volume reading are exponential.

While research on early writing instruction is less abundant, meta-analyses highlight impactful and promising practices. These practices include explicit instruction in the process and mechanics of writing and routines and structures that help make writing a pleasant and familiar experience (Graham et al., 2015). Research also supports the close integration of reading and writing instruction, so that skills taught in one domain are reinforced through exposure and practice in the other (Graham & Hebert, 2011). Additionally, frequency of writing is understood to be important for developing writers (Graham et al., 2015).

Motivation to Read

The current study explores the impact of integrated science and literacy instruction on student motivation to read. Increased motivation is identified as an important potential factor because interest-driven motivation to read is "the link between frequent reading and reading achievement" (Guthrie & Wigfield, 2000, p. 405). Motivation is acknowledged as a key factor in the development of the reading habits that support long-term achievement in literacy (McGeown et al., 2015; Wigfield et al., 2016). Furthermore, motivation is a significant predictor of reading comprehension both in young and older students (Jean et al., 2018; Taboada et al., 2009; Wang & Guthrie, 2004). A rigorous quasi-experimental study yielded impacts as high as .71 SD on comprehension from gains in motivation (Guthrie et al., 2006).

Research Questions

Here, we comprehensively examine the impacts, cost, and implementation of an integrated science and literacy curriculum for kindergarten. We also explore heterogeneity

of impacts based on student characteristics and teacher implementation fidelity. The research questions we address are:

Impact

1. Do students in kindergarten classrooms using an integrated science and literacy curriculum outperform students in business-as-usual control classrooms in:
   a. decoding and comprehension, as measured by the Woodcock Reading Mastery Test, 3rd Edition (WRMT)?
   b. reading and letter naming fluency, as measured by the Developmental Reading Assessment (DRA) and AIMSWeb curriculum-based assessment, respectively?
   c. writing, as measured by the Kaufman Test of Educational Achievement (KTEA)?
   d. science, as measured by a researcher-developed science assessment?
   e. motivation to read, as measured by Kindergarten Reading Motivation Scale (KRMS)?

Heterogeneity

2a. Do treatment effects vary among subgroups of students based on gender, language and home language status, IEP status, or eligibility for lunch assistance?

2b. Do outcomes for students in treatment classrooms vary based on teachers' fidelity of implementation?

Cost

3. What is the cost of *Zoology One* relative to business as usual literacy instruction?

Implementation

4. How was *Zoology One* implemented by teachers in the treatment group and what factors contributed to variations in fidelity?

Intervention and Context

The curriculum that is the focus of this evaluation is *Zoology One: Kindergarten Research Labs*, developed by American Reading Company (2019) (the program was later renamed *ARC Core Kindergarten*). *Zoology One* was selected for study because it is a widely used example of an integrated curriculum for young children, because its literacy and science content are both standards-aligned, and because it employs the evidence-based literacy practices discussed earlier.

Zoology One is a full-year curriculum centered around a daily 120-min integrated literacy and science instructional block. The program includes four 9-week units, implemented in succession. The first unit is introductory, designed to orient students to the basics of books and literacy and to build key classroom procedures. Following this introduction, the curriculum proceeds through a Zoology unit, an Ecology unit, and an Entomology unit. Teachers receive a new set of topically aligned instructional materials

and texts to use with each 9-week unit, but the structures and practices that guide instructional delivery are consistent throughout the year.

Zoology One uses a balanced literacy framework that incorporates each of the following in every daily instructional block: direct instruction in reading, writing, and science; complex text exposure delivered via multiple daily, themed teacher read-alouds; high-volume print exposure via supported independent reading in themed, leveled texts; formative assessment and progress monitoring implemented by the teacher during individual conferences or small-group instruction; high-volume writing practice related to the science theme; and science inquiry, including hands-on science activities and drama, music and art activities oriented around the science themes. *Zoology One* also includes a focus on parental involvement; students are expected to build the stamina to read for 30 min in class, and 30 min at home each day.

Teachers implementing *Zoology One* as part of this study received a full day of startup training at the beginning of the school year plus 10 visits throughout the year—approximately one visit per month—from coaches employed by the curriculum developer. While schools implementing this program can purchase varying amounts of coaching, 10 coaching visits are recommended by the developer. Two coaches provided all of the coaching to treatment teachers in this study. During their visits to implementing classrooms, the coaches provided a range of supports, including modeling components of *Zoology* One instruction such as whole-group instruction or skill-based small-group intervention; confirmation of student reading levels as determined by teachers; side-by-side coaching during one-to-one conferencing; and use of the program's formative assessment tools. Along with these supports, which were provided to all teachers, some teachers participated in an optional half-day introduction to the first science-themed unit, and some received additional support via phone or email from the coaches in between classroom visits.

The setting for the study is the School District of Philadelphia (SDP). Philadelphia is a city of approximately 1.5 million people (U.S. Census Bureau, 2019) and the poorest of the nation's large cities; approximately one-quarter of residents live in poverty (Hunger Free America, 2018). Despite a large charter- and private-school sector that serves nearly 40% of school-age children, approximately 134,000 students attend regular District-managed public schools. The demographics of the student body of SDP deviate from those of the city overall: For example, although over one-third of Philadelphia residents identify as non-Hispanic White, only 15% of SDP students do so; and although 44% of Philadelphia residents identify as African-American, more than 80% of SDP students are African-American. Median income of parents of children in public school is $3,000 below average for residents, and only 17% of parents of public school students have a bachelor's degree or higher (NCES, 2019), as compared with 27% of residents (U.S. Census Bureau, 2019). Philadelphia public schools serve large populations of English Language learners (10% of students) and students with Individualized Education Plans (18% of students) (NCES, 2019).

At the time of this study, the School District of Philadelphia had invested significantly in early literacy. Kindergarten teachers across the district were trained and supported in using a balanced literacy approach whose basic components parallel those of *Zoology One*.

Table 1. ITT and TOT student samples and attrition.

	Treatment	Control	Total
ITT Sample			
Cohort 1	239	253	492
Cohort 2	148	152	300
Combined	387	405	792
TOT Sample[a]			
Cohort 1	479	472	951
Cohort 2	340	298	638
Combined	819	770	1,589
Partial/no data[b]			
Cohort 1	123	115	238
Cohort 2	166	130	296
Combined	289	245	534
All known students			
Cohort 1	602	587	1189
Cohort 2	506	428	934
Combined	1108	1015	2,123
TOT attrition			
Cohort 1	20%	20%	20%
Cohort 2	33%	30%	32%
Combined	26%	24%	25%

Notes: [a]TOT sample includes ITT sample and early joiners. [b]Students with partial/no data include those who enrolled after fall assessments (140), students who unenrolled before the spring assessment (141), those who enrolled but never attended (140), those who were absent, ineligible, or did not consent (113). All known students are the sum of TOT sample and students with partial/no data.

Method

Evaluation Design

This study encompasses a multi-site cluster-randomized controlled trial (RCT) with embedded cost and implementation research. The impacts of *Zoology One* relative to business-as-usual instruction were estimated via the RCT, in which we randomly assigned entire kindergarten classrooms—including the teacher and all students— to conditions, within schools. Participating schools ranged in size, with the smallest having two kindergarten classrooms and the largest having nine. Classrooms assigned to treatment were expected to implement *Zoology One* in place of regular literacy instruction for 120 min per day, for the full school year. Classrooms assigned to control were expected to implement SDP's business-as-usual literacy program for 120 min per day. Teachers in the treatment condition were asked not to provide any science instruction over and above that provided via the *Zoology One* curriculum. Teachers in the control condition were asked to provide the same science instruction they normally would.

Within the RCT framework, we assessed the costs of the treatment condition relative to the business-as-usual control condition using the ingredients method (Levin et al., 2018). A mixed-methods approach was used to understand treatment teachers' implementation of the *Zoology One* program and the treatment/control contrast. Cost and implementation activities were coordinated to increase efficiency and provide a holistic understanding of both topics.

Participants

Participants in the evaluation of *Zoology One* included 71 kindergarten teachers in 21 schools and their students ($n = 1,589$). None of the teachers in the study had previously

Table 2. Baseline comparisons by sample and subgroup.

	ITT		TOT	
	Control	Treatment	Control	Treatment
Student demographics				
ELL	9%	8%	8%	7%
Female	49%	52%	50%	51%
FFT	62%	68%	70%	74%
IEP	6%	9%	7%	8%
Non English Home Language	17%	16%	14%	13%
Fall assessment means (and SD)				
WRMT Readiness Cluster	84.27	83.91	81.37	80.68
	(16.60)	(17.20)	(15.90)	(15.90)
DRA Level	0.74	0.81	0.64	0.67
	(0.70)	(0.77)	(0.65)	(0.72)

Note: sample groups were statistically significant on all measured student attributes measured and baseline assessments.

implemented the *Zoology One* program. The RCT was implemented in two cohorts of schools. Cohort 1 included 12 schools (with 40 kindergarten classrooms) during the 2016–2017 school year. Cohort 2 included 9 additional schools (with 31 kindergarten classrooms) during the 2017–2018 school year. At the start of each study year, the research team randomly assigned classrooms to treatment and control, within school. We conducted analyses using both an intent-to-treat (ITT) analytic sample and a treatment-on-treated (TOT) analytic sample. Table 1 summarizes the size of ITT and TOT samples by group and cohort.

The ITT sample included students who were rostered prior to random assignment of classrooms to treatment condition, and who were assessed in the fall and spring of Kindergarten. The TOT sample includes all students in the ITT sample and any other student with complete pre and post assessment data. We identified students who enrolled or were rostered after assignment as *early joiners*. Table 2 presents baseline demographic attributes and fall assessment data, all of which were equivalent in both ITT and TOT samples.

We tested for differences between study groups at baseline using a similar model specification to the one used for estimating impacts (see Analysis section). This analysis revealed no significant differences between groups at baseline based on scores on the Reading Readiness cluster of the WRMT, $t(772) = 0.09$, $p = 0.925$; DRA scores, $t(746) = 0.68$, $p = 0.495$; ELL status, $t(666) = -0.50$, $p = 0.620$; Female, $t(730) = 0.82$, $p = 0.414$; Free from Tape (FFT) status (a free/reduced-price lunch indicator) $t(730) = 1.30$, $p = 0.195$; IEP status, $t(666) = 1.73$, $p = 0.085$; and Non-English home language status, $t(730) = -1.28$, $p = 0.202$.

Student Outcome and Teacher Implementation Measures

We administered the Reading Readiness cluster of the WRMT to each cohort in the fall as a baseline assessment, and administered the Passage Comprehension, Word Attack, and Word Identification subtests in the spring as measures of reading comprehension and decoding. The WRMT was individually administered by trained, monitored assessors. To further investigate impacts on reading, we obtained secondary data from SDP's district-wide literacy assessments. SDP classroom teachers in both treatment and control

groups collected data in fall and spring each year using DRA and AIMSWeb, both widely used classroom-administered measures of reading achievement. The AIMSWeb probe assesses letter naming fluency, a known predictor of future reading achievement (Leppänen et al., 2008; Stage et al., 2001). The research team also administered the KTEA-3, an individually administered assessment, for the analysis of impacts on writing.

Although the intervention's theory of change does not posit impacts on math from *Zoology One*, we conducted a math assessment in Cohort 1 to examine the possibility that the expanded focus on literacy and science in the treatment classrooms might negatively impact math achievement. The measure for the Math outcome was the Kaufman Test of Educational Achievement (KTEA-3), which we administered to a random sample of 359 Cohort 1 students.

At the time of this study, there were few appropriate science assessments for kindergarten, and we were not able to identify an existing assessment that was comprehensive, accessible to pre-readers, and feasible to administer. As a result, we assessed science outcomes for this study using an instrument designed by the research team, in collaboration with advisors with expertise in both science and assessment development. In accordance with the Standards for Educational *and* Psychological Testing (2014), we developed, piloted, and selected items for this assessment over multiple rounds. The final instrument included 21 multiple choice items spanning all Next Generation Science Standards (NGSS) in Life Science from kindergarten through fifth grade, with 2–5 items for each standard. Along with items designed to assess life sciences content knowledge, the assessment included items that cover the science and engineering processes outlined in NGSS (such as using diagrams and graphs). Trained assessors administered the assessment individually to kindergarten students via Qualtrics using a touch screen and pictures. In order to eliminate confounds with students' English language and/or literacy proficiency, the assessment did not require students to read or speak.

The theory of change guiding the study posits that *Zoology One*'s engaging, animal-themed texts will result in increased motivation to read. To assess this hypothesis, researchers designed and validated a new measure, the Kindergarten Reading Motivation Scale (KRMS). Trained assessors individually administered the KRMS to 878 treatment and control students in the spring of 2017 as a measure of motivation to read. The measure includes 19 items probing students' feelings about reading (e.g., "Do you like to read?" "Can you learn new things from books?" "Do you like to look at books by yourself?").

We collected data from all classrooms in our sample to examine how treatment classrooms delivered the curriculum in terms of resource use and implementation fidelity and to identify the contrast between treatment instruction and business-as usual-instruction. Data sources included teacher surveys, daily activity logs, interviews, and school district documents outlining the scope and sequence for literacy and required instructional activities. We administered online surveys to teachers in the treatment and control conditions in the spring. Thirty-six of 37 treatment teachers completed the survey and 32 of 34 control teachers. Teachers responded on a range of topics, including their comfort teaching science, the materials and curricula they used for instruction, the quantity and perceived quality of the coaching they received, and any supports or interventions provided to students in addition to the regular curriculum. To understand

differences in science dosage, we also asked control teachers about the quantity of science instruction their students received. Treatment teachers were asked how much, if any, science instruction was provided over and above *Zoology One*. We also collected data about teachers' allocation of planning and instructional time across content areas using a daily activity log. The logs asked teachers to record their activities throughout the day in 30-min increments on three randomly selected school days from late fall to early spring. We used teacher logs to explore implementation fidelity and to examine contrasts between treatment and control in teacher time and resource use. In addition, we invited teachers to participate in interviews in order to expand on the same topics from the survey and logs. Forty-nine teachers participated in the interviews, (28 treatment and 21 control). The interview protocol was based on the program's theory of change, theory regarding instructional program implementation, and theory regarding teachers' implementation decisions.

Analyses

Impact (Research Question 1)

The impact of the Zoology One program after one school year of the intervention is based on an ITT analysis, where treatment status is determined at the time of randomization. Students who enrolled in a study school after random assignment to conditions were excluded from ITT analysis and treated as joiners for exploratory analysis of TOT impacts. A multilevel analysis of covariance was performed to estimate single-year treatment effects for student i, in classroom j, and school k, for this multi-site cluster randomized trial. The modeling approach allows for variation in treatment effects across schools (Raudenbush & Bryk, 2002). We treat classrooms and schools as random effects with school as the site block, producing four sources of variability: within-classroom variance (i.e., Level-1 student residual); between-classroom, within-school variance (i.e., Level-2 classroom intercept); between-school variance (i.e., Level-3 school intercept); and between-school variance in program effects (i.e., Level-3 school treatment effects). To estimate the average treatment effect across sites, a mixed-effect model was used in which the outcome of interest is a function of a fixed student effect associated with fall pretest scores (Level 1), and a fixed classroom effect associated with treatment status (Level 2). The three-level model used for estimating all impacts is presented in Equations (1)–(3).

$$Y_{ijk} = \pi_{0jk} + \beta_{1jk}X_{ijk} + e_{ijk} e_{ijk} \sim N(0, \sigma^2) \tag{1}$$

$$\pi_{0jk} = \beta_{00k} + \beta_{02k}T_{jk} + r_{0jk}r_{0jk} \sim N(0, \tau_\pi) \tag{2}$$

$$\beta_{00k} = \gamma_{000} + u_{00k}\text{var}(u_{00k}) \sim \tau_{\beta_{00}} \tag{3}$$

$$\beta_{02k} = \gamma_{010} + u_{01k}\text{var}(u_{01k}) \sim \tau_{\beta_{01}}$$

Equation (1) is the person-level model, where

π_{0jk} is the mean for classroom j in school k;
β_{1jk} is the pretest effect for student i in classroom j in school k;
X_{ijk} is the individual Reading Readiness cluster score at baseline;
e_{ijk} is the error associated with each student; and
σ^2 is the within-school variance.

Equation (2) is the classroom-level model, where

β_{00k} is the mean for school k;
β_{02k} is the treatment effect at school k;
T_{jk} is a treatment contrast indicator;
r_{0jk} is the random effect associated with each classroom; and
τ_π is the variance between classrooms within school.

Equation (3) is the school-level model, where

γ_{000} is the grand mean;
γ_{010} is the average treatment effect;
u_{00k} is the random effect associated with each school;
u_{01k} is the random effect associated with each school treatment effect;
$\tau_{\beta_{00}}$ is the variance between school means; and
$\tau_{\beta_{01}}$ is the variance between schools on treatment effects.

In this study, γ_{000} indicates the average scores for students in the control group, and γ_{010} is the main effect of treatment, both adjusted for student pretest scores. The error terms u_{00k}, u_{01k} allow these average effects to differ by school.

Heterogeneity and Fidelity (Research Question 2)

To explore heterogeneity of effects, we estimated treatment impacts separately for sub-groups based on gender, English Language Learner (ELL) status, participation in district feeding program (FFT), IEP status, and home language (Research Question 2a). Finally, we computed fidelity scores for each teacher, grouped the teachers into quartiles by fidelity score, and tested for differences between the students of high-fidelity teachers and those of low-fidelity teachers. This test is exploratory in nature (Research Question 2b).

Cost Analysis (Research Question 3)

We applied the ingredients method in a cost-effectiveness framework to estimate the cost of *Zoology One* relative to business-as-usual (Levin et al., 2018). We used the program's design, theory of change, and theoretical treatment contrast to design the cost study so that the cost estimate would capture the achieved relative strength of the program in resource terms (Hulleman & Cordray, 2009; Weiss et al., 2014). We measured the resources (ingredients) allocated for literacy instruction in treatment and control classrooms to estimate the cost of all ingredients used to produce the impacts we observed. Following cost-effectiveness standards, it is important to note that we apply the economic definition of costs and estimate the cost of all resources used, regardless of who financed them. Accordingly, we describe the ingredients used, illustrate how the program changed literacy instruction for the treatment group in practice, estimate the cost per student to achieve any observed change, and describe distribution of costs. We distinguish between total cost to produce effects—which

includes important inputs like parent/caregiver time for home reading and changes in other literacy curricula used in the treatment classrooms—and the purchase price to buy the curriculum.

First, we outlined, described, and quantified the ingredients related to literacy instruction in *Zoology One* and control classrooms, focusing data collection on those resources that were most likely to differ across conditions and drive a change in student learning (Levin, 1975; Levin et al., 2018). We collected data on personnel, training, materials, data management/software, and facilities components of each condition. Because we randomly assigned classrooms within schools, both conditions equally used facilities, transportation, food, and other resources related to schooling, so these inputs are not included in our analyses.

Second, we matched ingredients with standard average national prices to reflect market rates relevant for an efficacy RCT designed to inform the field. We obtained price data from the Department of Labor, ARC, and publicly available market prices for other literacy programs used in classrooms. Prices were adjusted for inflation to reflect 2018 US Dollars and amortized, when appropriate, to reflect the portion of an ingredient used during the year. For example, each classroom received a kit containing over 400 books. Most of the books will last longer than one year. We used teacher-reported data on the proportion of books that were lost or destroyed coupled with data from the Zoology One program records to estimate the frequency of book replacement. For the supplemental curricula, we assumed a life of 5 years and a classroom size of 22.5 students. We tested the sensitivity of our amortization assumptions by varying the years of available life of each resource to ensure that our findings are robust to these decisions.

Third, we calculated the total cost of Zoology One above and beyond business-as-usual. We examined the cost distribution focusing on costs borne by schools and parents/caregivers. Below, we present costs per student to correspond to the effectiveness estimates.

Implementation (Research Question 4)

Research Question 4 was addressed through a combination of mixed-methods analysis of implementation fidelity and qualitative analysis of teachers' explanations for how and why their implementation of *Zoology One* varied. We used data from the teacher surveys and daily activity logs to assess fidelity to the Zoology One program. The fidelity framework was developed at the outset of the study in collaboration with *Zoology One's* developer. In the analysis, we measured the extent to which teachers used 13 core components of the *Zoology* One curriculum. These components are identified in the program's logic model and pertain to the domains of resources and materials, training and coaching, and instruction.

For five of the 13 core components, we measured teachers' implementation fidelity via items on the teacher survey that asked about the consistency with which they implemented activities (e.g., "Students select and exchange books daily," "Students engage with science content daily via reading, writing, and hands-on science-themed activities"). We assigned one point if the teacher met the specified fidelity metric for the component (e.g., reported using the component "always or almost always"), and 0 points if not. For nine components, we used data from daily activity logs to assess implementation fidelity.

For each time period when they were implementing *Zoology One*, teachers were asked to indicate the components that they used. For each day, one point was assigned when the teacher indicated implementing the component at any time throughout the day. We averaged points for each component across the days for which logs were completed. Finally, the total fidelity score was computed by averaging the points assigned for each of the 13 components, with total possible scores ranging from 0 to 1.

To understand the reasons for variation in teachers' implementation of *Zoology One,* we analyzed data from interviews with twelve treatment teachers in Cohort 1 and 16 treatment teachers in Cohort 2, representing a total of 21 schools. We developed and used codes that emerged both inductively and deductively from the *Zoology One* logic model or interviews. We applied the codes to randomly selected transcripts, then discussed discrepancies to arrive at common understandings until 80% reliability was reached. To enhance the validity of our findings regarding factors contributing to variation in teachers' implementation of *Zoology One*, the research team applied three analytic strategies to the data coded with the "implementation factors" main code. First, we counted how many transcripts included a particular sub-code of the "implementation factors" main code (e.g., "coach") at least one time. This was a way to determine the prevalence of that sub-code within the overall sample. Next, we counted sub-codes within individual transcripts, and tallied the number of transcripts within which a given sub-code appeared most frequently. Last, we looked at which sub-codes emerged most often across all transcripts. After applying all three analytic methods to the data coded "implementation factors," we ranked the sub-codes by frequency across all three methods to derive our key themes.

Findings

Investigation of all differences between treatment and control conditions is important for interpreting the results of experimental research. Both treatment and control classrooms used a balanced literacy instructional approach designed to engage students with the Kindergarten Common Core Standards for Reading, Writing, and Speaking and Listening across a range of instructional settings. In both conditions, students experienced a combination of direct instruction, teacher read-alouds, shared reading, small-group instruction, and independent practice in both reading and writing.

In terms of science, our expectation at the outset of the study was that students in the treatment group would receive only the science instruction embedded in *Zoology One,* and that students in the control group would receive little science instruction at all. We observed, however, that students in both groups received more direct science instruction than we anticipated, and that total minutes of daily, direct instruction in science were similar across the two groups. In the control group, 41% of teachers reported that their classes received science as an enrichment class, taught by a specialized teacher. Fifty-three percent of control-group teachers reported that they taught science to their own students, either in addition to or instead of the enrichment class. In all, control teachers reported that their students received, on average, 28 min of science instruction per day.

Treatment teachers were asked not to teach science outside the *Zoology One* block, and most (89%) reported that they complied with this request. However, treatment teachers reported that their classes received enrichment science at similar rate to control classes (36%). Combined with the explicit science instruction provided as part of *Zoology One*, we estimate that this resulted in comparable total minutes per day of explicit science instruction for both groups.

Furthermore, we found that the teachers themselves were similar, on average. We observed no differences between treatment and control teachers based on years teaching, years teaching kindergarten, and years teaching at current school. There were no significant differences between groups in the number of teachers who identified as certified reading specialists.

In terms of differences, we found that control teachers used more packaged curricular programs and interventions than *Zoology One* teachers. Most control teachers delivered direct instruction in phonics using a commercially available whole-class program, Saxon Phonics, while most treatment teachers delivered phonics instruction within *Zoology One's* instructional components, and often in small-group or individual settings. Following the *Zoology One* curriculum, treatment teachers used a formative assessment framework to guide and target instruction during small-group work and individual conferences, and they used formative assessment data to select texts each day for read-alouds and shared reading activities. Teachers in the control group more often used basal readers and/or textbooks; instruction was paced in accordance with an established scope and sequence rather than students' progress data. Two other key differences were noted: First, *Zoology One* emphasizes home reading and teachers are instructed to send books home with students each day to support this component. This was not a widely observed practice in the control classrooms. Second, while *Zoology One* embeds science instruction within the instructional components of the literacy block, the business-as-usual literacy program did not. Thus, in addition to explicit science instruction, *Zoology One* students also had up to 120 min per day of sustained immersion in science content through the program's science-themed teacher read-alouds, student texts, and writing activities.

Research Question 1: Do Students in Kindergarten Classrooms Using an Integrated Science and Literacy Curriculum Outperform Students in Business-as-Usual Control Classrooms?

Table 3 presents results of baseline equivalency tests on the WRMT Reading Readiness Cluster (used as pretest measure in impact models). The following four columns present impacts on decoding and comprehension via the Word Identification, Word Attack, Word Comprehension, and Passage Comprehension subtests of the WRMT. Students in the treatment group scored significantly higher on one outcome, Passage Comprehension ($b = 1.90$, $t(771) = 2.11$, $p = 0.035$). Passage Comprehension had an SD of 11.86 in the control group, producing an ITT Glass's Delta effect size of 0.16. We find no differences between treatment groups on the Word Attack, Word Identification, or Word Comprehension WRMT subtests.

We also observed significant differences on two other outcomes, letter naming fluency and motivation to read. In the ITT sample, letter naming fluency ($b = 8.35$ $t(459) =$

Table 3. ITT and TOT impact analysis results with baseline equivalence findings.

	Baseline	WRMT				District tests		Other assessments		
	RRC	WI	WA	WC	PC	LNF	DRA	KRMS	KTEA-3	Science
Intent to treat										
Sample size	792	792	792	792	792	471	783	737	518	295
Random effects										
School intercept	52.87*	0.93	0.54	0.27	1.74	0.00	0.04	0.00	2.52	20.79
School	–	6.19	2.42	4.41	1.33	101.26	0.00	0.00	4.08	0.52
treatment effect										
Classroom intercept	0.00	11.49*	5.10*	0.90	4.13	14.44	0.25**	0.00	8.84*	14.07
Residual (student)	237.1**	88.72**	65.3**	45.1**	79.80**	499.38**	0.64**	0.11**	68.81*	225.24**
Fixed effects										
Constant	83.70	40.76	55.15	59.06	51.22	–37.03	–1.08	2.35	58.69*	24.07*
Treatment	0.11	1.37	–0.27	1.07	1.90*	8.35*	0.07	0.11*	–0.10	–0.48
Fall RC	–	0.58**	0.42*	0.38**	0.47**	1.01**	0.05**	0.00	0.45**	0.45**
Treatment on the treated										
Sample size	1,589	1,589	1,589	1,589	1,589	916	1,562	1,387	933	626
Random effects										
School intercept	38.79*	1.64	0.00	0.00	0.82	0.00	0.05	0.00	2.38	12.53
School	–	0.51	2.68	4.24*	0.00	58.20	0.00	0.00	0.78	0
treatment effect										
Classroom intercept	2.28	14.57**	5.51**	0.80	3.10*	39.62**	0.23**	0.00	9.37*	0
Residual (student)	216.38**	74.54**	57.90**	34.50*	70.90**	507.51**	0.62**	0.11**	58.96**	218.58**
Fixed effects										
Constant	81.48	40.45	57.40	64.30	52.74	–44.24	–1.30	2.26	58.39	19.24
Treatment	–0.27	0.66	–0.55	0.83	1.59*	5.00	0.13	0.10**	0.82	1.67
Fall RC	–	0.58**	0.40*	0.34**	0.46**	1.11**	0.05**	0.00**	0.44**	0.48**

Notes: *$p < 0.05$; ** $p < 0.01$. RRC: WRMT Reading Readiness Cluster; WA: WRMT Word Attack subtest; WC: WRMT Word Comprehension subtest; PC: WRMT Passage Comprehension subtest; KRMS: Kindergarten Reading Motivation Scale; LNF: AimsWeb Letter Naming Fluency; WI: WRMT Word Identification subtest; DRA: Developmental Reading Assessment; KTEA: Kauffman Test of Educational Achievement.

Table 4. Standardized treatment effects for student samples and subsamples.

	Baseline	WRMT				District tests			Other assessments	
	RRC	WI	WA	WC	PC	LNF	DRA	KRMS	KTEA-3	Science
Intent to treat										
All	0.01	0.10	−0.03	0.13	0.17*	0.27*	0.05	0.32*	−0.01	−0.03
Female	−0.03	0.09	−0.03	0.08	0.08	0.25	0.01	0.28*	0.07	−0.14
Male	0.04	0.14	0.03	0.21	0.30*	0.31*	0.14	0.33*	−0.05	−0.06
FFT	0.04	0.04	−0.12	0.06	0.12	0.18	0.02	0.20*	−0.04	0.06
Non-FFT	−0.03	0.17	0.22	0.29	0.22	0.33*	0.06	0.44*	0.03	−0.39
ELL	−0.47	−0.01	−0.12	−0.11	−0.01	0.39	0.05	−0.04	0.24	−0.35
Non-ELL	0.03	0.08	−0.03	0.16	0.14	0.21	0.01	0.33*	−0.07	0.01
IEP	−0.09	0.20	−0.13	0.30	0.14	0.51	0.49	0.14	0.48	0.57
Non-IEP	0.00	0.08	−0.02	0.14	0.13	0.20	0.00	0.31*	−0.05	−0.03
English home	0.04	0.11	−0.05	0.14	0.17	0.26	0.03	0.32*	−0.02	−0.07
Non-Engl home	−0.18	0.05	0.08	0.07	0.12	0.38	0.29	0.33	0.08	0.38
Treatment on the treated										
All	−0.02	0.05	−0.05	0.10	0.14*	0.16	0.10	0.29*	0.07	0.10
Female	−0.03	0.02	−0.08	0.06	0.09	0.15	0.03	0.21*	0.14	0.09
Male	−0.03	0.09	−0.02	0.14	0.20*	0.19	0.19	0.35*	0.03	0.07
FFT	−0.04	0.03	−0.09	0.04	0.12*	0.16	0.14	0.25*	0.06	0.10
Non-FFT	0.03	0.05	0.05	0.24	0.17	0.17	0.00	0.34*	0.04	0.05
ELL	−0.09	0.05	−0.05	0.05	0.14	0.50*	0.14	0.22	0.16	−0.59
Non-ELL	0.00	0.04	−0.05	0.12	0.14*	0.12	0.09	0.28*	0.05	0.11
IEP	0.00	0.11	−0.11	0.16	0.12	0.32	0.29*	0.16	0.36	0.44
Non-IEP	−0.02	0.05	−0.03	0.11	0.14*	0.14	0.09	0.30*	0.05	0.09
English home	−0.01	0.07	−0.05	0.12	0.15*	0.15	0.10	0.28*	0.10	0.07
Non-Engl home	−0.04	0.00	−0.09	0.02	0.06	0.32	0.13	0.36	0.06	0.14

Notes: *$p < 0.05$. FFT: Free From TAPE (free & reduced lunch indicator); RRC: WRMT Reading Readiness Cluster; WA: WRMT Word Attack subtest; WC: WRMT Word Comprehension subtest; PC: WRMT Passage Comprehension subtest; KRMS: Kindergarten Reading Motivation Scale; LNF: AimsWeb Letter Naming Fluency; WI: WRMT Word Identification subtest; DRA: Developmental Reading Assessment; KTEA: Kauffman Test of Educational Achievement.

2.02, $p = 0.044$) had an SD of 29.37 in the control group, producing an ITT Glass's Delta effect size of 0.28. Analysis of data from the KRMS revealed that *Zoology One* students scored statistically significantly higher than control students in reading motivation ($b = 0.11$, $t(716) = 4.58$, $p < 0.0001$), producing an ITT Glass's Delta effect size of 0.32 SD. We find no differences between the treatment and control groups on science, as measured by our researcher-developed assessment, on the general reading outcome measured by the DRA, and no differences on the writing outcome.

Research Question 2a: Do Literacy Treatment Effects Persist within Subgroups Based on Gender, Language and Home Language Status, IEP Status, or Lunch Assistance?

Exploratory analyses of treatment effects within salient student groups are presented as fully standardized effect sizes for ITT and TOT in Table 4. Results demonstrate that the overall treatment effects, both significant and non-significant, bore out in most subgroups.

Standardized effects across all student subgroups suggest that the reported average treatment effect is robust and generalizable. Notably, estimated impacts for boys' reading comprehension were somewhat larger than those for girls'; impacts for native English speakers were larger than those for English Language Learners; and impacts for students who do not qualify for free or reduced-price lunch were larger than those for students who do qualify.

Table 5. Effect of teacher fidelity on student mediators and outcomes.

	Baseline	WRMT				District tests		Other assessments		
	RRC	WI	WA	WC	PC	LNF	DRA	KRMS	KTEA-3	Science
Intent to treat										
Sample size	210	210	210	210	210	129	207	197	121	61
Random effects										
School intercept	14.82∼	0.00	2.09	0.62	0.00	158.28	0.00	0.00	7.15	2.56
Classroom intercept	0.00	6.99∼	0.00	0.33	2.03	0.00	0.17	0.00	0.00	0.00
Residual (student)	226.19**	89.63**	51.11**	29.17**	74.39**	565.82**	0.58**	0.09**	54.03**	191.20**
Fixed effects										
Constant	79.35**	44.17**	61.84**	67.29**	55.40**	−34.35*	−0.78*	2.36**	65.12**	13.93
Teacher fidelity	5.70*	4.40*	2.73*	1.88*	2.75∼	0.15	0.26	0.04	3.09	−11.23
Fall RC	–	0.51**	0.31**	0.27**	0.41**	1.04**	0.04**	0.00	0.33**	0.56**
Treatment on the treated										
Sample size	423	423	423	423	423	232	414	395	243	156
Random effects										
School intercept	15.21*	0.00	0.93	0.27	0.00	92.62∼	0.00	0.00	4.02	0.37
Classroom intercept	0.00	7.80*	3.25	0.14	1.53	0.00	0.20	0.00	0.00	0.00
Residual (student)	199.71**	71.11**	53.12**	25.37**	68.05**	537.85**	0.63**	0.09**	50.81**	197.74**
Fixed effects										
constant	77.74**	43.94**	61.74**	69.49**	56.98**	−41.25**	−1.22**	2.35**	60.85**	10.91
Teacher fidelity	3.98∼	4.11*	2.40*	1.79*	1.40	−2.38	0.12	0.03	4.57*	−4.90
Fall RC	–	0.51**	0.31**	0.24**	0.41**	1.13**	0.05**	0.00	0.38**	0.59**

Notes: $**p < 0.001, *p < 0.05, \sim p < 0.10$, Treatment group only comparing students of teachers with highest and lowest fidelity (quartiles). RRC: WRMT Reading Readiness Cluster; WA: WRMT Word Attack subtest; WC: WRMT Word Comprehension subtest; PC: WRMT Passage Comprehension subtest; KRMS: Kindergarten Reading Motivation Scale; LNF: AimsWeb Letter Naming Fluency; WI: WRMT Word Identification subtest; DRA: Developmental Reading Assessment; KTEA3: Kauffman Test of Educational Achievement.

Research Question 2b: Do Literacy Scores in Treatment Classrooms Vary Based on Teachers' Fidelity of Implementation?

An exploratory analysis compared literacy impacts for students of high-fidelity implementers with those of low-fidelity implementers. This contrast revealed statistically significant differences between top-quartile and bottom-quartile implementers on WRMT Word Attack and Word Identification subtests and on the KTEA assessment of writing (Table 5).

These findings suggest that the intervention is effective on some outcomes even with lower fidelity, while other outcomes require faithful implementation. Because teachers were not randomly assigned to fidelity condition it is understood that teachers with high fidelity may be different from teachers with low fidelity in ways that would relate to student achievement.

Research Question 3: What Are the Relative Costs Associated with the Curriculum?

We observed that the *Zoology One* was largely implemented as designed, in terms of resources. We found that home reading was not fully achieved, with students averaging 75 min per week versus the recommended 150. *Zoology One* teachers, on average, reported using the curriculum for more minutes per day than required (170 vs. 120 min).

Zoology One is designed to be a comprehensive, balanced literacy curriculum, reducing the need for additional curricula to teach literacy. Thus, to estimate costs that

Table 6. Differences in instructional program between treatment and control.

Program	Treatment	Control
District endorsed programs		
Saxon	8%	91%
Lexia	17%	59%
Ready gen	3%	75%
Self-created programs	75%	97%
Teacher reports using "other" programs	50%	94%
Mean number of programs reported	3	6

Note: Quantities rounded to the nearest whole number.

Table 7. Average incremental cost per student of *Zoology One*.

Ingredient	Treatment quantity	Control quantity	Unit	Total cost	Distribution	
					School	Home
Personnel						
Teaching time	170	170	Min/day			
School principal	1	1	Principal			
Teacher's assistant	1	1	Assistant			
Home reading	40	20	Hours	$160		$160
Training						
Coaching sessions	10	0	Sessions	$150	$150	
Materials						
ARC core materials	1	0	Kit	$160	$160	
Additional programs	2	6	Program	−$40	−$40	
Data and technology						
Computer	1	1	Item			
SchoolPace	1	0	Year	$40	$40	
Facilities						
Reading nook in classroom	1	1	Nook			
Total cost:				$480	$320	$160
					67%	33%

Notes: Total costs reflect the incremental resources to produce effects (treatment—control). We use constant 2018 U.S. dollar, rounded to the nearest ten. Averages are weighted to account for the size of program groups. We amortized costs of ingredients with useful life beyond one year using a 3.5% discount rate.

correspond to effects, we also considered changes in other curricula that contributed to literacy development (Bowden et al., 2017). As noted earlier, we observed that teachers implementing Zoology One used fewer curricular programs in their classrooms overall. On average, the treatment classrooms used *Zoology One* and three additional curricular supports. Control classrooms used six curricula on average. Three programs prominently used in control classrooms but not in treatment classrooms were Lexia, Saxon Phonics, and ReadyGen. While estimating the total costs of these programs was outside the scope of this work, we use purchase prices as minimum values in our analysis to adjust for this reduction in other curricula. We find that the elimination of three programs in *Zoology One* classrooms creates important cost savings for schools, averaging around $40 in savings per student or $900 saved per classroom per year relative to the business as usual programming in control classrooms (Table 6).

As shown in Table 7, the incremental cost of *Zoology One* is about $480 per student on average. This cost reflects the differences in resources received by students between experimental conditions, and thus the cost to produce effects. This means that the cost estimate can be combined with the effectiveness estimate in a cost-effectiveness ratio to compare the efficiency of this curriculum to other whole-class kindergarten literacy

curricula. At the time of this study, there were no comparable curricula listed in the What Works Clearinghouse.

As stated above, the cost to produce effects is not equivalent to the purchase price. For example, when a school purchases coaching support, the school pays a fee for coaches from ARC to serve all teachers delivering the program in the school. In our study, the program was delivered to roughly half of the kindergarten students in each school in our sample. This means that the cost of coaching in our study is being divided by fewer students than a typical implementation. Also, we adjust the price of the coaching to reflect the investment in teaching capital, which lasts for longer than one year. For the purposes of our evaluation, the cost estimate must correspond to the effectiveness estimate to reflect the cost to produce the impact rather than an idiosyncratic purchase price, which does not reflect the true value of this investment.

When we examine the distribution of costs, we find that one-third of the total cost is driven by parent/caregiver time allocated through the home reading component. To estimate the portion of costs borne by the school, we sum the components of the curriculum that are typically purchased by schools (in the evaluation, these were funded by a research grant) and any resources purchased or reallocated by the school to deliver the curriculum. Because the treatment replaced existing curricula, there were very few additional or reallocated costs. We found that the cost to the school for implementing *Zoology One* is valued at approximately $320 per student. Based on this estimate, the school or district bears 67% of the total cost to deliver the program.

Research Question 4: How Was the Integrated Curriculum Implemented by Teachers in the Treatment Group and Why Did Teachers' Implementation Vary?

Our analysis of fidelity of implementation revealed high levels of fidelity overall, with substantial variation. The total possible fidelity score ranged from 0 to 1, where 1 indicated that a teacher used all 13 components of Zoology One and 0 indicated that a teacher used 0 components. The mean teacher fidelity score was .74. There was substantial between-teacher variability in the total fidelity score ($SD = .13$; range [.42, 1]). The sample had a 25th percentile fidelity score of .62 and a 75th percentile fidelity score of .85.

There was also considerable variability in the mean fidelity score among the specific *Zoology One* components. Mean component scores ranged from .47 (students assigned independent reading at home; indicating that approximately half of the teachers in the sample implemented this component) to 1 (implementing the 4 units in succession; indicating that all teachers implemented this component).

Reasons Teachers Varied Implementation
Our analysis of qualitative data from interviews with 28 teachers in 21 schools revealed two primary factors impacting teachers' implementation of Zoology One. These are (1) teachers' assessments of their students' needs and interests, and (2) constrained instructional time. Findings from our three parallel analyses indicate that teachers' perceptions of their students' interests and needs were the most influential factor driving variations

in implementation. More specifically, twenty-one of the 28 teachers interviewed (75%) cited students' needs and interests as a factor that influenced their implementation of Zoology One. Of these 21 teachers, 11 teachers (52%) also identified students' needs and interests more frequently than any other factor.

When teachers talked about how their students' needs and interests influenced their use of Zoology One, it was generally in reference to decisions to a) vary the time allotted to particular activities, or b) supplement the program with other materials or instruction. For example, we found that teachers frequently varied the time allocated for independent reading depending on student engagement, or eliminated components on particular days in order to extend Zoology One activities in which the students were engaged. Teachers also talked about supplementing Zoology One based on their students' needs. Most often, these teachers reported incorporating additional phonics instruction, perhaps instead of other program elements.

"Time" was the second most frequently applied code in the data set; 21 of the 28 teachers referenced time constraints or scheduling issues as a primary reason why they deviated from the Zoology One curriculum. Teachers who identified time as a factor spoke about the unpredictability of their daily schedules and changes to their schedules based on factors outside their control. These interruptions and schedule changes might limit their ability to do all they had planned for a particular day—for example, they might have planned for two read-alouds, in accordance with the Zoology One lesson plan, but ended up having time for only one.

Another factor we identified in the interview data was the availability of support staff to help with various aspects of Zoology One (e.g., guided reading, independent reading, writing). Without extra staff to manage centers, supervise activities, or meet with children individually or in small groups, some teachers reported being unable to implement the whole curriculum on a given day. This theme appeared in 20 of 28 interviews, and seven teachers (35%) cited it more often than any other factor. Finally, of the 28 teachers we interviewed, 15 (54%) mentioned their personal beliefs when discussing decisions about program implementation and five (33%) mentioned beliefs more than any other factor. Teacher beliefs influenced implementation in that teachers place varying levels of priority on different literacy components—independent reading or direct phonics instruction, for example—and therefore implemented those components more or less consistently, or deviated from the Zoology One schedule in order to accommodate them.

Study Limitations

Several limitations to this research should be noted. First, we report impacts at the end of kindergarten only. Work is underway to examine participants' literacy achievement in the years following this portion of the study. The findings of this longitudinal research will be critical to the interpretation and practical applicability of our findings. Similarly, our research into the impacts of Zoology One is limited by the fact that all teachers were in their first year of implementation of the program. Research indicates that teachers' use of curricula improves over time (Ladd & Sorensen, 2017), suggesting

that we might observe different results from teachers beyond their initial year using *Zoology One*.

Additionally, we were unable to identify a validated science assessment that could adequately capture Zoology One's hypothesized impacts on science learning. As a result, our finding of no impacts on science comes with the caveat that our researcher-developed measure was not validated for purposes of measuring science achievement in kindergarten. The research team's future work will assess impacts on this study's participants' 4th-grade state assessments in science.

We assessed writing using a standardized assessment, the KTEA, which we selected in large part for the feasibility of administration and scoring with our large sample of young children. However, this assessment focused on writing mechanics, and scores from our study indicate that kindergarten students in SDP—both treatment and control—significantly underperform the KTEA norming sample in this area, introducing the threat of floor effects. In our teacher interviews, many Zoology One teachers reported observing an increase in students' ability to compose complete and focused sentences and paragraphs as compared with their prior classes of kindergarteners, as well as increased stamina for writing. The KTEA instrument was not designed to detect these changes. As a result, some impacts on student writing may have gone undetected.

Discussion

Despite these limitations, the findings presented here provide important insights and highlight promising directions for future research. Our causal impact findings in reading are mixed; we observed meaningful impacts from *Zoology One* on reading comprehension and letter-naming fluency, but treatment students performed no better or worse than control students in two other key areas of interest, writing and decoding. When juxtaposed with the resource and cost findings, however, the nuance of this statement is meaningful: Students in the treatment group performed no better and no worse in decoding than those in the control group, despite the latter group having participated in a daily, intensive program that teaches phonics in isolation in addition to their other literacy curriculum. Indeed, treatment students, whose teachers overall reported using 3 fewer programs and interventions on a daily basis, performed as well as or better in all areas of literacy achievement than peers in the control group.

Furthermore, our exploratory analysis revealed that students whose teachers implemented *Zoology One* with high fidelity performed statistically significantly better in writing and decoding, two areas where main impacts were not observed. Thus, we found either a significant treatment effect or significant group mean difference based on fidelity for every literacy outcome of interest. While our fidelity finding is not causal, it suggests that even a relatively low-fidelity implementation of *Zoology One* yields impacts in comprehension and letter-naming fluency and that high fidelity can yield additional impacts in decoding and writing.

Our exploratory analysis of impacts on science yielded null findings. This result surprised us given the immersive science focus of *Zoology One.* However, our conclusion that both groups received similar quantities of direct instruction in science most likely explains this finding. A question lingers as to why the daily immersion in science-

themed reading and writing activities throughout the *Zoology One* literacy block did not appear to accelerate treatment students' acquisition of science knowledge. This question warrants further exploration with a validated science assessment. Additionally, our ongoing investigation of longitudinal impacts on the science achievement of students in our study may yield further insights. Based on our current findings alone, however, it appears that *Zoology One* is best viewed as a literacy program, rather than a science intervention.

This study detected notable effects on students' motivation to read. These effects speak to the overarching question of this study: What benefits are realized when early literacy instruction is combined with science? Here, our findings tell a story that assessment scores alone cannot. Particularly in a context like SDP's, where many children enter kindergarten with limited pre-reading experience (Hindman et al., 2016), motivation is a promising lever for accelerating reading experience and proficiency. We observed that *Zoology One* improved students' motivation to read, with an educationally meaningful effect size of .32 SD. This finding suggests that *Zoology One* may be an effective way to activate this known pathway.

Further, our analysis revealed no gender differences in motivation to read. This finding is interesting in multiple regards. First, research points to higher motivation to read among girls generally and suggests that boys' motivation is more fragile, and more intertwined with reading skill, even in the early grades (Logan & Medford, 2011). The equally high levels of motivation to read among *Zoology One* girls and boys may indicate that the curriculum could bolster motivation among boys. This hypothesis is particularly intriguing given that we observed stronger effects from *Zoology One* for boys than for girls on passage comprehension. Viewed together, these two findings suggest that *Zoology One* bolsters both motivation and reading skill in an urban, largely minority population of boys. Given that this group has traditionally lagged in literacy, this is a potentially powerful finding. Our motivation-to-read findings are also interesting given that girls typically exhibit lower interest in science relative to boys; might early, immersive exposure that connects science with literacy counteract this phenomenon? Each of these avenues warrants further research, and additional work from this study on the impacts of *Zoology One* on girls' reading preferences is forthcoming.

ORCID

A. Brooks Bowden http://orcid.org/0000-0001-6079-5456

References

Allington, R. L. (1977). If they don't read much, how they ever gonna get good? *Journal of Reading, 21*(1), 57–61. http://www.jstor.org/stable/40009862

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing (U.S.). (2014). *Standards for educational and psychological testing*. APA.

American Reading Company (2019, December 19). *American Reading Company: Reorder, replenish, renew.* https://www.americanreading.com/solutions/reorder

Anderman, E. M., & Young, A. J. (1994). Motivation and strategy use in science: Individual differences and classroom effects. *Journal of Research in Science Teaching, 31*(8), 811–831. https://doi.org/10.1002/tea.3660310805

Baker, L. (1991). Metacognition, reading and science education. In C. M. Santa & D. E. Alvermann (Eds.), *Science learning: Processes and applications* (pp. 2–13). International Reading Association.

Banilower, E. R., Smith, S., Weiss, I. R., Malzahn, K. J., Campbell, K. H., & Weis, A. H. (2013). *Report of the 2012 national survey of science and mathematics education*. Horizon Research.

Bowden, A. B., Shand, R., Belfield, C. R., Wang, A., & Levin, H. M. (2017). Evaluating educational interventions that induce service receipt: A case study application of City Connects. *American Journal of Evaluation, 38*(3), 405–419. https://doi.org/10.1177/1098214016664983

Cervetti, G. N., Barber, J., Dorph, R., Pearson, P. D., & Goldschmidt, P. G. (2012). The impact of an integrated approach to science and literacy in elementary school classrooms. *Journal of Research in Science Teaching, 49*(5), 631–658. https://doi.org/10.1002/tea.21015

Cervetti, G. N., Pearson, P. D., Bravo, M. A., & Barber, J. (2006). Reading and writing in the service of inquiry-based science. In R. Douglas & M. Klentschy (Eds.), *Linking science and literacy in the K–8 classroom*. National Science Teachers Association.

Cipielewski, J., & Stanovich, K. (1992). Predicting growth in reading ability from children's exposure to print. *Journal of Experimental Child Psychology, 54*(1), 74–89. https://doi.org/10.1016/0022-0965(92)90018-2

Clements, D. H., & Sarama, J. (2016). Math, science, and technology in the early grades. *The Future of Children, 26*(2), 75–94. https://doi.org/10.1353/foc.2016.0013

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33*(6), 934–945. https://doi.org/10.1037/0012-1649.33.6.934

Curran, F. C., & Kellogg, A. T. (2016). Understanding science achievement gaps by race/ethnicity and gender in kindergarten and first grade. *Educational Researcher, 45*(5), 273–282. https://doi.org/10.3102/0013189X16656611

Curran, F. C., & Kitchin, J. (2019). Early elementary science instruction: Does more time on science or science topics/skills predict science achievement in the early grades? *AERA Open, 5*(3), 233285841986108. https://doi.org/10.1177/2332858419861081

Donahue, P. L., Finnegan, R. J., Lutkus, A. D., Allen, N. L., & Campbell, J. R. (2001). *The nation's report card: Fourth-grade reading 2000 (NCES 2001-499)*. U.S. Department of Education, Office of Educational Research and Improvement. National Center for Education Statistics.

Duke, N. K. (2000). 3.6 Minutes per day: The scarcity of informational texts in first grade. *Reading Research Quarterly, 35*(2), 202–224. https://doi.org/10.1598/RRQ.35.2.1

Duke, N. K., Pearson, P. D., Strachan, S. L., & Billman, A. K. (2011). Essential elements of fostering and teaching reading comprehension. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about reading instruction* (4th ed., pp. 51–93). International Reading Association.

Ehri, L. C. (1991). Development of the ability to read words. In R. Barr, M. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 383–417). Routledge.

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading*, 9(2), 167–188. https://doi.org/10.1207/s1532799xssr0902_4

French, L. (2004). Science as the center of a coherent, integrated early childhood curriculum. *Early Childhood Research Quarterly*, 19(1), 138–149. https://doi.org/10.1016/j.ecresq.2004.01.004

Garan, E. M., & DeVoogd, G. (2008). The benefits of sustained silent reading: Scientific research and common sense converge. *The Reading Teacher*, 62(4), 334–336. https://doi.org/10.1598/RT.62.4.6

Goldschmidt, P., & Jung, H. (2011). *Evaluation of seeds of science/roots of reading (CRESST report)*. National Center for Research on Evaluation, Standards, and Student Testing.

Graham, S., Harris, K. R., & Santangelo, T. (2015). Research-based writing practices and the common core: Meta-analysis and meta-synthesis. *The Elementary School Journal*, 115(4), 498–522. https://doi.org/10.1086/681964

Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review*, 81(4), 710–744. 744. 784-785. https://doi.org/10.17763/haer.81.4.t2k0m13756113566

Grissmer, D., Grimm, K. J., Aiyer, S. M., Murrah, W. M., & Steele, J. S. (2010). Fine motor skills and early comprehension of the world: Two new school readiness indicators. *Developmental Psychology*, 46(5), 1008–1017. https://doi.org/10.1037/a0020104

Guthrie, J. T., Anderson, E., Alao, S., & Rinehart, J. (1999). Influences of concept-oriented reading instruction on strategy use and conceptual learning from text. *The Elementary School Journal*, 99(4), 343–366. https://doi.org/10.1086/461929

Guthrie, J. T., & Humenick, N. M. (2004). Motivating students to read: Evidence for classroom practices that increase reading motivation and achievement. In P. McCardle & V. Chhabra (Eds.), *The voice of evidence in reading research* (pp. 329–354). P.H. Brookes Pub.

Guthrie, J. T., & Wigfield, A. (2000). Engagement and motivation in reading. In M. L. Kamil, P.B. Mosenthal, P.D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. III, pp. 403–422). Routledge.

Guthrie, J. T., & Wigfield, A. (2006). The impact of concept-oriented reading instruction on students' reading motivation, reading engagement, and reading comprehension. In J. L. Meece & J. S. Eccles (Eds.), *Handbook of research on schools, schooling, and human development* (pp. 463–475). Routledge.

Guthrie, J. T., Wigfield, A., Humenick, N. M., Perencevich, K. C., Taboada, A., & Barbosa, P. (2006). Influences of stimulating tasks on reading motivation and comprehension. *The Journal of Educational Research*, 99(4), 232–246. https://doi.org/10.3200/JOER.99.4.232-246

Guthrie, J. T., Wigfield, A., Metsala, J. L., & Cox, K. E. (1999). Motivational and cognitive predictors of text comprehension and reading amount. *Scientific Studies of Reading*, 3(3), 231–256. https://doi.org/10.1207/s1532799xssr0303_3

Henrichs, L. F., & Leseman, P. P. (2014). Early science instruction and academic language development can go hand in hand. The promising effects of a low-intensity teacher-focused intervention. *International Journal of Science Education*, 36(17), 2978–2995. https://doi.org/10.1080/09500693.2014.948944

Hindman, A. H., Wasik, B. A., & Snell, E. K. (2016). Closing the 30 million word gap: Next steps in designing research to inform practice. *Child Development Perspectives*, 10 (2), 134–139. https://doi.org/10.1111/cdep.12177

Hulleman, C. S., & Cordray, D. S. (2009). Moving from the lab to the field: The role of fidelity and achieved relative intervention strength. *Journal of Research on Educational Effectiveness*, 2(1), 88–110. https://doi.org/10.1080/19345740802539325

Hunger Free America. (2018). *Hunger Free America: Philadelphia falling behind in freedom to eat* [Greater Philadelphia Hunger Report, 2018]. Hunger Free America.

Jean, E., Morgan, P. L., & Fuchs, D. (2018). Understanding and promoting the motivation to read in young children. In N. D. Young, C. N. Michael, & T. Citro (Eds.), *From floundering to fluent: Reaching and teaching struggling readers* (pp. 41–51). Rowman & Littlefield.

Jorm, A. F., & Share, D. L. (1983). Phonological recoding and reading acquisition. *Applied Psycholinguistics*, *4*(2), 103–147. https://doi.org/10.1017/S0142716400004380

Juel, C. (1991). Beginning reading. In R. Barr, M. L. Kamil, P. B. Mosenthal & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 759–788). Routledge.

Kurz, L. A. (2018). *Literacy integrated with kindergarten science: An investigation of impacts of kindergarten students including those at-risk for learning disabilities* (Publication No. 10809792) [Doctoral dissertation, George Mason University]. ProQuest Dissertations Publishing.

Ladd, H. F., & Sorensen, L. C. (2017). Returns to teacher experience: Student achievement and motivation in middle school. *Education Finance and Policy*, *12*(2), 241–279. https://doi.org/10.1162/EDFP_a_00194

Leppänen, U., Aunola, K., Niemi, P., & Nurmi, J. E. (2008). Letter knowledge predicts grade 4 reading fluency and reading comprehension. *Learning and Instruction*, *18*(6), 548–564. https://doi.org/10.1016/j.learninstruc.2007.11.004

Levin, H. M. (1975). Cost-effectiveness analysis in evaluation research. In M. Guttentag & E. L. Struening (Eds.), *Handbook of evaluation research* (Vol. 2, p. 89). Sage.

Levin, H. M., McEwan, P., Belfield, C., Bowden, A. B., & Shand, R. (2018). *Economic evaluation in education: Cost-effectiveness and benefit cost analysis*. Sage.

Logan, S., & Medford, E. (2011). Gender differences in the strength of association between motivation, competency beliefs and reading skill. *Educational Research*, *53*(1), 85–94. https://doi.org/10.1080/00131881.2011.552242

Lonigan, C. J., Schatschnieder, C., & Westberg, L. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. National Institute for Literacy.

McGeown, S. P., Johnston, R. S., Walker, J., Howatson, K., Stockburn, A., & Dufton, P. (2015). The relationship between young children's enjoyment of learning to read, reading attitudes, confidence and attainment. *Educational Research*, *57*(4), 389–402. https://doi.org/10.1080/00131881.2015.1091234

Miller, M., & Moss, B. (2013). *No more independent reading without support*. Heinemann.

Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin*, *137*(2), 267–296. https://doi.org/10.1037/a0021890

Morgan, P. L., Farkas, G., Hillemeier, M. M., & Maczuga, S. (2016). Science achievement gaps begin very early, persist, and are largely explained by modifiable factors. *Educational Researcher*, *45*(1), 18–35. https://doi.org/10.3102/0013189X16633182

National Center for Educational Statistics. (2019). www.nces.ed.gov

National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. National Institute for Literacy.

National Science Foundation. (2019). *Women, minorities, and persons with disabilities in science and engineering: 2019*. National Center for Science and Engineering Statistics. https://ncses.nsf.gov/pubs/nsf19304/

Okahana, H., Feaster, K., & Allum, J. (2016). *Graduate enrollment and degrees: 2005 to 2015*. Council of Graduate Schools.

Ozernov-Palchik, O., Norton, E. S., Sideridis, G., Beach, S. D., Wolf, M., Gabrieli, J. D., & Gaab, N. (2017). Longitudinal stability of pre-reading skill profiles of kindergarten children: Implications for early screening and theories of reading. *Developmental Science*, *20*(5), e12471. https://doi.org/10.1111/desc.12471

Padilla, M. J., Muth, K. D., & Lund Padilla, R. K. (1991). Science and reading: Many process skills in common. In S. M. Santa & D. E. Alvermann (Eds.), *Science learning—Processes and applications* (pp. 14–19). International Reading Association.

Pearson, P. D., Moje, E., & Greenleaf, C. (2010). Literacy and science: Each in the service of the other. *Science*, *328*(5977), 459–463. https://doi.org/10.1126/science.1182595

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Sage.

Reutzel, D. R., Fawson, P. C., & Smith, J. A. (2008). Reconsidering silent sustained reading: An exploratory study of scaffolded silent reading. *The Journal of Educational Research*, *102*(1), 37–50. https://doi.org/10.3200/JOER.102.1.37-50

Romance, N. R., & Vitale, M. R. (1992). A curriculum strategy that expands time for in-depth elementary science instruction by using science-based reading strategies: Effects of a year-long study in grade four. *Journal of Research in Science Teaching*, *29*(6), 545–554. https://doi.org/10.1002/tea.3660290604

Romance, N. R., & Vitale, M. R. (2001). Implementing an in-depth expanded science model in elementary schools: Multi-year findings, research issues, and policy implications. *International Journal of Science Education*, *23*(4), 373–404. https://doi.org/10.1080/09500690116738

Sahin, A., Ayar, M. C., & Adiguzel, T. (2014). STEM related after-school program activities and outcomes on student learning. *Educational Sciences: Theory & Practice*, *14*(1), 309–322.

Samuels, S. J., & Wu, Y. (2003). *How the amount of time spent on independent reading affects reading achievement: A response to the National Reading Panel*. Univesity of Minnesota.

Shanahan, T., Callison, K., Carriere, C., Duke, N. K., Pearson, P. D., Schatschneider, C., & Torgesen, J. (2010). *Improving reading comprehension in kindergarten through 3rd grade: A practice guide* (NCEE 2010-4038). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, *55*(2), 151–218. https://doi.org/10.1016/0010-0277(94)00645-2

Slavin, R. E., Lake, C., Hanley, P., & Thurston, A. (2014). Experimental evaluations of elementary science programs: A best-evidence synthesis. *Journal of Research in Science Teaching*, *51*(7), 870–901. https://doi.org/10.1002/tea.21139

Stage, S. A., Sheppard, J., Davidson, M. M., & Browning, M. M. (2001). Prediction of first-graders' growth in oral reading fluency using kindergarten letter fluency. *Journal of School Psychology*, *39*(3), 225–237. https://doi.org/10.1016/S0022-4405(01)00065-6

Stanley, C. T., Petscher, Y., & Catts, H. (2018). A longitudinal investigation of direct and indirect links between reading skills in kindergarten and reading comprehension in tenth grade. *Reading and Writing*, *31*(1), 133–153.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*(4), 360–407. https://doi.org/10.1598/RRQ.21.4.1

Stine, D. (2009). *America COMPETES Act: Programs, funding, and selected issues*. Library of Congress, Congressional Research Service.

Strickland, D., & Riley-Ayers, S. (2006). *Early literacy: Policy and practice in the preschool years*. National Institute for Early Education Research (NIEER).

Taboada, A., Tonks, S. M., Wigfield, A., & Guthrie, J. T. (2009). Effects of motivational and cognitive variables on reading comprehension. *Reading and Writing*, *22*(1), 85–106. https://doi.org/10.1007/s11145-008-9133-y

Topping, K. J., Samuels, J., & Paul, T. (2007). Does practice make perfect? Independent reading quantity, quality and student achievement. *Learning and Instruction*, *17*(3), 253–264. https://doi.org/10.1016/j.learninstruc.2007.02.002

U.S. Census Bureau. (2019). *American Community Survey*. https://www.census.gov/

Vedder-Weiss, D., & Fortus, D. (2011). Adolescents' declining motivation to learn science: Inevitable or not? *Journal of Research in Science Teaching*, *48*(2), 199–216. https://doi.org/10.1002/tea.20398

Vitale, M. R., & Romance, N. R. (2011). Adaptation of a knowledge-based instructional intervention to accelerate student learning in science and early literacy in grades 1 and 2. *Journal of Curriculum and Instruction*, *5*(2), 79–93.

Wang, J. H. Y., & Guthrie, J. T. (2004). Modeling the effects of intrinsic motivation, extrinsic motivation, amount of reading, and past reading achievement on text comprehension between U.S. and Chinese students. *Reading Research Quarterly*, *39*(2), 162–186. https://doi.org/10.1598/RRQ.39.2.2

Weiss, M. J., Bloom, H. S., & Brock, T. (2014). A conceptual framework for studying the sources of variation in program effects. *Journal of Policy Analysis and Management*, *33*(3), 778–808. https://doi.org/10.1002/pam.21760

Wendt, J. L., Rockinson-Szapkiw, A., & Cordes, M. (2018). Examining the influence of a STEM certification model on female, minority science outcomes. *The Journal of Educational Research*, *111*(6), 666–677. https://doi.org/10.1080/00220671.2017.1396437

Wigfield, A., Gladstone, J. R., & Turci, L. (2016). Beyond cognition: Reading motivation and reading comprehension. *Child Development Perspectives*, *10*(3), 190–195. https://doi.org/10.1111/cdep.12184

Wigfield, A., Guthrie, J. T., Perencevich, K. C., Taboada, A., Klauda, S. L., McRae, A., & Barbosa, P. (2008). Role of reading engagement in mediating effects of reading comprehension on reading outcomes. *Psychology in the Schools*, *45*(5), 432–445. https://doi.org/10.1002/pits.20307

Wright, T. S., & Gotwals, A. W. (2017). Supporting kindergartners' science talk in the context of an integrated science and disciplinary literacy curriculum. *The Elementary School Journal*, *117*(3), 513–537. https://doi.org/10.1086/690273

Wright, T. S., & Neuman, S. B. (2014). Paucity and disparity in kindergarten oral vocabulary instruction. *Journal of Literacy Research*, *46*(3), 330–357. https://doi.org/10.1177/1086296X14551474