Piasta, S.B., Logan, J.A.R., Zettler-Greeley, C.M., Bailet, L.L., Lewis, K., & Thomas, L.J.G. (2022). Small-group, emergent literacy intervention under two implementation models: Intent-to-treat and dosage effects for preschoolers at-risk for reading difficulties. *Journal of Learning Disabilities,* https://doi.org/10.1177%2F00222194221079355.

Small-Group, Emergent Literacy Intervention under Two Implementation Models: Intent-to-Treat and Dosage Effects for Preschoolers At-Risk for Reading Difficulties

Shayne B. Piasta, Ph.D.[1,2]

201 Arps Hall, 1945 North High Street, Columbus OH 43210, piasta.1@osu.edu

Jessica A.R. Logan, Ph.D.[3]

211A Ramseyer Hall, 29 W. Woodruff Avenue, Columbus, OH 43210, logan.251@osu.edu

Cynthia M. Zettler-Greeley, Ph.D.[4]

10140 Centurion Parkway N., Jacksonville, FL 32256, czettler@nemours.org

Laura L. Bailet, Ph.D.[5]

1310 Lewisville Clemmons Rd, Lewisville, NC 27023, laurabailet04@gmail.com

Kandia Lewis[6]

2200 Concord Pike, Wilmington, DE 19803, kandia.lewis@nemours.org

& Leiah J.G. Thomas, M.A.[1,2]

333 Arps Hall, 1945 North High Street, Columbus OH 43210, leiahgroom@gmail.com

[1] Department of Teaching and Learning, The Ohio State University

[2] Crane Center for Early Childhood Research and Policy, The Ohio State University

[3] Department of Educational Studies, The Ohio State University

[4] Center for Health Delivery Innovation, Nemours Children's Health, Jacksonville, Florida

[5] Kaplan Early Learning Company, Lewisville, North Carolina

[6] Center for Health Delivery Innovation, Nemours Children's Health, Wilmington, Delaware

Author Note

Shayne B. Piasta https://orcid.org/0000-0003-3655-4702

Jessica A.R. Logan https://orcid.org/0000-0003-3113-4346

Cynthia M. Zettler-Greeley https://orcid.org/0000-0003-2256-2971

Laura L. Bailet https://orcid.org/0000-0003-2744-9409

Kandia Lewis https://orcid.org/0000-0002-7891-9258

Leiah J.G. Thomas https://orcid.org/0000-0001-9988-3285

Laura L. Bailet is now retired from Kaplan Early Learning Company.

This study was preregistered with the Open Science Framework (Registration doi: 10.17605/OSF.IO/UWNRF).

Nemours receives royalties through the sale of the commercially available curriculum involved in this research. The potential for bias in reporting study results was minimized through the adoption of the following precautions, as outlined in the investigators' signed Memorandum of Understanding: (a) Nemours' institutional responsibilities for this grant were limited to instructor professional development curriculum training and implementation fidelity monitoring, including the development and maintenance of the implementation fidelity database for this study, (b) Nemours investigators Zettler-Greeley, Bailet (affiliated with Nemours until April,

2018; at Kaplan Early Learning Company until June, 2021), and Lewis had no role in participant/site recruitment or data collection for the study, were blind to classroom and participant study assignment, and did not participate in data analyses concerning program impacts, (c) The Ohio State University (OSU) investigators conducted all data analyses concerning impacts of the Nemours BrightStart! program, and (d) OSU investigators retained the final decision as to the findings and interpretation that are reported.

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A160261 awarded to The Ohio State University (Piasta). The opinions expressed are those of the authors and do not represent views of the Institute or U.S. Department of Education. The authors gratefully acknowledge the support of the research project staff, the cooperation of the Ready4Success initiative as led by Shelby Dowdy, Robin Sayers for her thoughtful feedback on manuscript drafts, and the early childhood professionals and children/families without whom this research would not be possible.

Correspondence concerning this article should be addressed to Shayne B. Piasta, Department of Teaching and Learning, The Ohio State University, 201 Arps Hall, 1945 North High Street, Columbus, OH 43210. Email: piasta.1@osu.edu

Supplemental Material are provided with the published version of this article, on the journal website.

Abstract

Preschool-aged children identified as at risk for later reading difficulties can benefit from supplemental, small-group emergent literacy intervention. As such interventions become commercially available and marketed to preschool programs, it is important to understand their impacts when implemented by intended end users under routine conditions. In this study, we examined effects of the Nemours BrightStart! (NBS!) intervention on children's emergent literacy skills, when implemented by teachers and community aides in authentic preschool classrooms. We randomly assigned 98 classrooms to one of three conditions (NBS! teacher-implemented, NBS! community aide-implemented, or control). Children enrolled in these classrooms who met eligibility criteria and were identified as at risk via an early literacy screener (n = 281) completed pretest and posttest emergent literacy assessments; those assigned to NBS! conditions received intervention from their classroom teacher or a community aide affiliated with a local kindergarten-readiness initiative. Intent-to-treat analyses showed no significant impacts of NBS! on any outcome, and an instrumental variables, as-treated approach showed one significant intervention effect on letter writing. Consequently, we did not replicate results of prior, highly controlled efficacy trials. Findings have implications for revising the NBS! theory of change, conducting dosage and as-treated analyses, and moving research-based interventions towards scale up.

Keywords: small-group instruction, emergent literacy intervention, prevention of reading difficulties, early childhood, dosage

Small-Group, Emergent Literacy Intervention under Two Implementation Models: Intent-to-Treat and Dosage Effects for Preschoolers At-Risk for Reading Difficulties

The preschool years mark the development of many foundational emergent literacy skills underpinning a child's later reading success, including print knowledge, phonological awareness, oral language, and emergent writing (National Early Literacy Panel, 2008). Many children attain these skills through systematic exposure to print and high-quality instructional activities. However, for a variety of reasons, nearly one-third of children are one or more standard deviations behind their peers on measures of emergent literacy as they enter kindergarten, placing them at risk for later reading difficulties (Fielding et al., 2007; Greenwood et al., 2013). Early identification of reading difficulties paired with targeted intervention is encouraged as early as preschool, as these children are unlikely to catch up to their peers with general classroom instruction alone (Dickinson et al., 2004; Ferrer et al., 2015).

Fortunately, children identified as needing additional emergent literacy support can benefit from supplemental, small-group literacy instruction (Bailet et al., 2009; Byrne & Fielding-Barnsley, 1991, 1995; Gettinger & Stoiber, 2008; Goldstein et al., 2017; Gonzalez et al., 2011; Kelley et al., 2015; Lonigan & Phillips, 2016; Phillips et al., 2016; Pollard-Durodola et al., 2011; Spencer et al., 2015; Zettler-Greeley et al., 2018). Supplemental, domain-specific instruction that supports "code-focused" skills (e.g., print knowledge, phonological awareness) and "meaning-focused" skills (e.g., language comprehension) promotes gains in print knowledge, phonological awareness, and language outcomes (Lonigan et al., 2013; see also National Early Literacy Panel, 2008). Both code-focused and meaning-focused skills are essential for conventional reading, as espoused in the Simple View of Reading (Gough & Tunmer, 1986), and are supported by decades of empirical studies linking growth in these skills

with later reading outcomes (Hjetland et al., 2020). Effective, small-group interventions incorporate systematic, scaffolded, intentional, and differentiated instructional approaches (Piasta, 2016; Wasik, 2008).

Small-Group Emergent Literacy Interventions: Impacts and Implementation Models

Small-group, supplemental emergent literacy interventions can improve children's literacy skills when delivered by a variety of implementers (Bailet et al., 2009, 2013; Byrne & Fielding-Barnsley, 1991, 1995; Goldstein et al., 2017; Gonzalez et al., 2011; Kelley et al., 2015; Kruse et al., 2015; Phillips et al., 2016; Pollard-Durodola et al., 2011; Spencer et al., 2012; Zettler-Greeley et al., 2018). Many interventions described in the literature have been implemented by research staff or highly trained interventionists, which is necessary to evaluate whether an intervention is effective under ideal conditions (i.e., efficacy trials; Byrne & Fielding-Barnsley, 1991; Lonigan & Phillips, 2016; Phillips et al., 2016; Spencer et al., 2015; Zettler-Greeley et al., 2018). For instance, when researchers delivered the PAth to Literacy intervention, preschoolers demonstrated significant improvements in phonological awareness and print knowledge (Kruse et al., 2015). Likewise, researcher implementation of the Sound Foundations intervention resulted in significant effects on preschoolers' phonemic awareness and significant long-term effects on decoding ability in first and second grade (Byrne & Fielding-Barnsley, 1991, 1995). Language in Motion and Story Champs (Phillips et al., 2016; Spencer et al., 2015) have shown positive impacts on aspects of children's oral language when implemented by research staff, and another early literacy intervention, Nemours BrightStart!, demonstrated positive impacts on preschoolers' print knowledge and phonological awareness when implemented by trained early literacy specialists (Bailet et al., 2009, 2013; Zettler-Greeley et al., 2018). A few small-group, emergent literacy interventions also have evidenced effects when

implemented under more authentic conditions. For example, when implemented by teachers or teacher aides, the Sound Foundations and PAth to Literacy interventions have shown improvements in preschoolers' phonological awareness (Byrne & Fielding-Barnsley, 1995; Goldstein et al., 2017) and the Words of Oral Reading and Language Development intervention improves vocabulary learning (Gonzalez et al., 2011; Pollard-Durodola et al., 2011).

Investigating the impacts of interventions when implemented by intended end users, such as teachers, is important as we strive to move research-based practices into classrooms (Durlak & DuPre, 2008). Ultimately, promising interventions developed by researchers must be placed in the hands of implementers who deliver instruction in authentic classroom contexts. Whereas these are often effective under more controlled conditions, small-group interventions may be challenging for preschool teachers to implement. Their implementation requires additional time for lesson preparation and planning as well as considerable classroom and behavior management (Wasik, 2008; Zucker et al., 2013); preschool teachers rarely receive formal training on the latter (Wyatt & Chapman-DeSousa, 2017). Time and management constraints may make small-group intervention untenable for many preschool teachers without the aid of another adult in the classroom (Farley et al., 2017), contributing to less-than-ideal implementation and affecting intervention dosage in particular (i.e., amount of the intervention delivered/received; Piasta et al., 2021). In turn, children who do not receive recommended amounts of a high-quality intervention are unlikely to optimize learning gains (Bailet et al., 2009; Durlak & DuPre, 2008; Ehri et al., 2001; cf. Suggate, 2016).

Efforts to facilitate teachers' use of research-based, small-group emergent literacy interventions include providing additional training in small-group instruction and classroom management (Zucker et al., 2013), assigning teacher aides to manage the class while the lead

teacher delivers the small-group instruction (Gonzalez et al., 2011; Pollard-Durodola et al., 2011), or employing automated instructional programs that do not require direct adult facilitation (Kelley et al., 2015). Other approaches enlist members of the community to assist. Such "community aides" may be individuals with or without prior teaching experience who are affiliated with local community organizations and who are trained to implement intervention as tutors or assistants (Al Otaiba et al., 2005; Fitzgerald et al., 2018; Markovitz et al., 2015; Nelson et al., 2009; see also Ritter et al., 2006, 2009). Importantly, literacy interventions that capitalize on trained community volunteers, tutors, and assistants have shown positive effects on children's emergent literacy and conventional reading skills (Ritter et al., 2006, 2009; Savage & Carless, 2005; Savage et al., 2003). For example, in the Minnesota Reading Corps PreK program, trained AmeriCorps members conduct early literacy screening and provide supplemental emergent literacy instruction to children in Head Start programs. Findings from a quasi-experimental study suggested positive effects of this approach on preschoolers' print knowledge, phonological awareness, and vocabulary outcomes (Markovitz et al., 2015). Similarly, Nelson and colleagues (2009) showed positive effects of small-group phonological awareness and alphabet knowledge intervention as delivered by trained community members in Head Start classrooms. Collectively, these studies demonstrate that implementation by community aides may be a viable option for conducting small-group emergent literacy intervention in preschool classrooms.

Employing community aides to administer small-group intervention in preschool classrooms may confer advantages over other implementation models. First, this solution may alleviate time constraints faced by preschool teachers tasked with numerous classroom and administrative responsibilities, enabling increased intervention dosage (Piasta et al., 2021). Second, a community aide can deliver small group instruction to children across multiple centers

and classrooms, efficiently serving larger numbers of children. Third, unlike automated programs, community aides can engage in the types of responsive, supportive feedback identified as essential for literacy development (Piasta, 2016; Wasik, 2008). Consequently, preschool teachers may have greater flexibility in how they facilitate emergent literacy learning if community aides can deliver small-group, emergent literacy interventions effectively.

With more widespread use of these interventions in preschool programs, it is necessary to understand the impacts of small-group emergent literacy interventions when implemented by intended end users under authentic conditions. In recognition of potential challenges for teachers as intended end users, such as time constraints, alternative implementation models such as those in which community aides rather than teachers implement the intervention must be considered to inform scale up. In the present study, we examined the effects of the Nemours BrightStart! intervention on children's emergent literacy skills, when implemented by teachers and community aides in authentic preschool classrooms. In doing so, we expand prior efficacy work in which the intervention was implemented by early literacy specialists hired and trained by the research/development team and, to our knowledge, are the first to systematically compare teacher- versus community aide-implementation models.

Nemours BrightStart! Intervention

The Nemours BrightStart! intervention (NBS!) provides small-group, supplemental emergent literacy instruction to young children who have been identified as at risk for later reading difficulties. Informed by the Simple View of Reading (Gough & Tunmer, 1986) extended to the preschool context (Hjetland et al., 2020), NBS! targets growth in code-focused and meaning-focused skills. The 20 NBS! lessons are contextualized by thematic and dialogic read alouds and emergent writing activities that provide opportunities for rich linguistic

exchanges between children and instructors, while also providing explicit and systematic print and phonological awareness instruction. The intervention is grounded in research concerning early literacy development and best practices for supporting such development (e.g., Anthony & Francis, 2005; Justice et al., 2003; Koutsoftas et al., 2009; Lonigan et al., 2009; National Early Literacy Panel, 2008; Pearson & Gallagher, 1983; Piasta, 2016; Piasta & Wagner, 2010; Schneider et al., 2000) and aligns with recommended practices for small-group instruction (Wasik, 2008) in that lessons are structured, scaffolded, and manualized.

The initial 18-lesson intervention was evaluated in several randomized controlled trials using a delayed control-group design. Early literacy specialists hired and trained by the developers implemented the intervention in over 100 preschool and childcare settings in one large metropolitan community in the southeastern United States (Bailet et al., 2009, 2013; Zettler-Greeley et al., 2018). On average, specialists exhibited 98% adherence to core lesson components and implemented 16 of 18 lessons, reflecting near-optimal implementation (Zettler-Greeley et al., 2018). Moreover, NBS! resulted in statistically significant effects on children's print knowledge (d = 0.20–0.74), rhyming (d = 0.24–0.37), alliteration (d = 0.44), elision (d = 0.46), blending (d = 0.41), and general early literacy skills (d = 0.24–0.39) in these efficacy trials (Bailet et al., 2009, 2013; Zettler-Greeley et al., 2018). Whereas statistically significant treatment effects for oral language were not evidenced, additional analyses indicated that vocabulary gains were positively correlated with intervention dosage, defined as the number of lessons that children completed; similar associations with dosage were noted for other outcomes as well.

Despite showing the promise of NBS!, these efficacy trials had several limitations. First, the research was conducted by the intervention developers and thus lacked independence. Second, child outcome measures did not include emergent writing, and oral language measures

were limited to expressive vocabulary (Bailet et al., 2009). Third, lessons were provided by early literacy specialists, many of whom had considerable instructional and education experience and who were accountable for high implementation fidelity, as they reported directly to the research/development team. Fourth, based on this prior research, NBS! was modified to a 20-lesson version which has now been made commercially available to teachers and preschool programs. The slightly revised, commercially available version has not been tested empirically.

The Current Study

In the current study, we address these limitations and expand the research on NBS! to increase generalizability of past findings and move towards scale up (Gottfredson et al., 2015). We evaluated NBS! as implemented by preschool teachers and community aides, to determine its impacts under these two implementation models when used under routine conditions and in authentic preschool classrooms. To do so, we conducted a randomized controlled trial of the commercially-available intervention and provided teachers and community aides with the materials and professional development that typically accompanies its purchase. The study was led by researchers unaffiliated with NBS! and conducted in a region of the U.S. in which it had not been widely used. We also included a broader array of thirteen distinct outcome measures to more fully represent the constructs targeted in intervention content.

We addressed three study aims. In Aim 1, we examined the impact of NBS!, when implemented by teachers or community aides, on children's emergent literacy skills under a rigorous, intent-to-treat framework (Sagarin et al., 2014). Intent-to-treat frameworks include all participants in analyses, as originally randomized and regardless of how much intervention was received. As such, intent-to-treat results allow for unbiased causal estimates of intervention effects and, critical to the current work, which involved implementing NBS! under authentic

conditions, these estimates reflect how the intervention was actually used and delivered within the context of the study. Based on prior efficacy trials, we hypothesized that NBS! would significantly and positively impact children's print knowledge, phonological awareness, oral language, and emergent writing relative to children assigned to a business-as-usual control condition. In Aim 2, we explored whether NBS! impacts differed based on implementer (i.e., teacher or community aide). We did not have a directional hypothesis for this aim. On one hand, both types of instructors received the same materials and training; on the other, prior evidence suggests potential tradeoffs in implementation between teachers (higher adherence) and community aides (higher dosage; Piasta et al., 2021).

In Aim 3, we conducted as-treated analyses that considered the impacts of NBS! after accounting for variation in dosage. This addressed a limitation of intent-to-treat frameworks, in that participants do not always receive intervention as intended. Specifically, given the potential challenges of implementing a small-group intervention in preschool classrooms, we had a priori expectations that not all children would complete all 20 lessons and, based on prior correlational work showing positive associations between dosage and learning gains (Bailet et al., 2009; Zettler-Greeley et al., 2018), hypothesized that dosage may need to be taken into account when examining intervention impacts. We thus used an instrumental variables approach (Angrist et al., 1996) as a modern means of addressing as-treated impacts (Sagarin et al., 2014). This approach capitalized on our use of random assignment as an instrument for removing selection biases and thereby provided a causally interpretable estimate of intervention effects after accounting for how much of the intervention children received.

Method

Participants

The study involved three sequential cohorts of participating preschool classrooms in one Midwestern state; each classroom participated for one academic year. Classrooms were eligible for participation if (a) the classroom enrolled 3- to 5-year old children, (b) the lead or co-lead teacher voluntarily provided informed consent and agreed to participate in all study activities, (c) the administrator agreed to study participation and signed the project's memorandum of understanding, and (d) at least one child in the classroom was eligible for study participation, as described below. Based on a priori power calculations (with minimally detectable effect sizes of .30 to .36; see the study's Open Science Framework preregistration; doi: 10.17605/OSF.IO/UWNRF), we enrolled a total of 98 classrooms in the project.

Classroom teachers completed a questionnaire at the beginning of the academic year to report background characteristics about themselves and their classrooms. Most classrooms were situated in early childhood centers (84%; 9% in public schools; 7% other/unreported); these were located in urban (77%), suburban (21%) and rural (1%) areas. Forty percent were affiliated with Head Start, and 62% accepted public subsidies (6% unreported). Teachers reported using Creative Curriculum in the majority of classrooms (78%); additional reported curricula included Assessment, Evaluation, and Programming System (12%), the Reggio Emilia Approach (11%), Scholastic (10%), Handwriting Without Tears (8%), Building Language for Literacy (7%), and Everyday Mathematics (7%), with an assortment of other curricula reported by less than 5% of teachers. Class sizes ranged from 6 to 24 children (M = 13.96, SD = 4.59). Most classrooms were staffed by lead/co-lead teachers who were female (96%). Fifty-one percent of teachers were Black, 43% were White, 2% were Asian, and 3% were multiracial/other (1% unreported); 3% were Hispanic/Latinx. Highest degrees completed included a high school diploma or equivalent (18%), associate's degree (33%), bachelor's degree (33%), and master's degree (9%); 2% did not

have a high school diploma (5% unreported). Twenty-five percent held teaching licenses. Preschool teaching experience ranged from 0 to 33 years (M = 10.77, SD = 8.25).

Within each classroom, we sampled up to four children who met study eligibility criteria. All children were between 3- to 5-years of age, had caregiver consent to participate, had minimal attendance issues and no severe behavior problems as reported by the teacher, and were identified as at risk for later reading difficulties based on the Get Ready to Read!-Revised (Whitehurst & Lonigan, 2010), as further described in the Assessment Procedures and Measures section. Additionally, parents completed a background questionnaire on which they reported on their child's English proficiency, disability status and medical history, and classroom functioning. For children for whom English was not their primary language, we required that parents reported that their child understood and spoke English with at least basic fluency. For children with disabilities or medical conditions, we required that parents reported these having little impact on children's abilities to learn and participate in classroom activities. These criteria were enacted to ensure that the intervention and study assessments were appropriate for participating children. For classrooms in which more than four children met these criteria, we randomly selected four to participate.

A total of 281 children met these criteria and enrolled in the study. Forty-eight percent were girls, and the average age was 4.28 years (SD = 0.50) at study start. Fifty-four percent were Black, 20% were White, 12% were multiracial, and 9% were of other races (5% unreported); 13% were Hispanic/Latinx. Two percent had an Individualized Education Plan or 504 plan. The highest degrees earned by children's parents included a high school diploma or GED (61%), associate's degree (12%), bachelor's degree (10%), and graduate degree (6%); 10% did not have a high school diploma. Annual family income was $25,000 or less for 56% of children, between

$25,001 and $75,000 for 35% of children, and greater than $75,000 for 6% of children (3% unreported). We provide descriptive information by condition in the Supplemental Material (Table S.1), along with the study CONSORT diagram.

Random Assignment and Intervention Procedures

We randomly assigned eligible classrooms to one of three study conditions: NBS! teacher-implemented, NBS! community aide-implemented, or business-as-usual control. We conducted randomization separately for each cohort, using an Excel random number generator. Children who were enrolled in classrooms assigned to the teacher-implemented or community aide-implemented conditions and selected to participate (maximum of four children per classroom) received the NBS! intervention, delivered by the classroom teacher or a community aide, as assigned, across the academic year. Children in the control condition only experienced their typical classroom instruction.

NBS! Instructors

Instructors for NBS! lessons were children's classroom teachers in the teacher-implemented condition and community aides in the community-aide implemented condition. Full descriptive information for teachers who implemented NBS! is provided in Table S.1 in the Supplemental Material. All were female, 57% identified as Black, 42% identified as White, and none identified as Hispanic/Latina. Education levels ranged from a high school diploma or GED through holding a master's degree.

Community aides were adults employed by a kindergarten readiness initiative that had been partnering with local early childhood programs since 2013 to provide emergent literacy screenings and professional development. Community aides were hired and supervised by the initiative staff to provide NBS! and received hourly wages; they did not report to research staff.

All had an interest in working with young children, successfully completed a background check, and voluntarily agreed to study activities. Across the three cohorts, nine community aides implemented NBS!. Two community aides resigned from their positions after implementing 5 to 6 lessons; their duties were reassigned to other community aides. Community aides completed a questionnaire at the beginning of the academic year to report background characteristics about themselves; full descriptive information is provided in Table S.1 in the Supplemental Material. All were female, 57% identified as Black, 29% identified as White, 14% identified as multiracial, and none identified as Hispanic/Latinx. Education levels ranged from a high school diploma or GED through holding a doctoral degree. All had previous experience working with children, including two as a classroom or substitute teacher. Others served as a reading tutor, Sunday School teacher, classroom volunteer, choir director, and summer camp aide.

NBS! Intervention

NBS! is designed to be implemented with small groups of four or fewer preschool children identified as at risk for later reading difficulties based on early literacy screening scores. The intervention includes systematic, small-group emergent literacy instruction to support the development of print knowledge, phonological awareness, oral language and comprehension, and emergent writing skills. Lesson content, structure, and instructional strategies reflect research and established best instructional practices at the time of program development, including the findings of the National Early Literacy Panel (2008). The intervention consists of 20 scripted lessons, each of which follows a seven-part instructional routine and involves interactive manipulatives to support learning (e.g., magic erase boards, boxes and gel bags for letter writing; mirrors for viewing mouth and tongue movements; magnifying glasses for letter identification; picture cards and magnets for illustrating alliterative concepts). Lessons feature a

gradual release of responsibility model (Pearson & Gallagher, 1983) in which concepts are explicitly taught and modeled, practiced with instructor support, and practiced independently. Instruction is "failure-free" in that corrective feedback is provided but child mastery is not required prior to introduction of subsequent concepts. More information on lesson organization, content, and activities is provided in Supplemental Material (see also Bailet et al., 2009, 2013; Zettler-Greeley et al., 2018). For purposes of the current study, each lesson was subdivided into two parts and implemented in two 20-30-min sessions per week with the small groups of selected children in NBS! teacher-implemented and NBS! community aide- implemented classrooms.

Prior to implementing, instructors received all necessary materials (i.e., instructor guide, lesson plans, books, manipulatives) and attended a 2-day, face-to-face professional development training. The professional development was the same as provided by the publisher to those who purchase the intervention and led by a staff member affiliated with the developers. It included an overview of NBS! and its instructional design, step-by-step instructions for implementing all lesson components, demonstrations and video exemplars, and hands-on practice. Notably, two teachers assigned to the NBS!-teacher implemented condition stopped participating prior to this training and did not attend (or implement any lessons; both resigned from their positions) and a third teacher missed the face-to-face professional development training but completed training with research staff. Given the intent-to-treat design, we retained participating children in these classrooms in analyses.

We report detailed information about implementation fidelity procedures in Piasta et al. (2021), in which we also provide a multiple-method analysis of implementation for the first two cohorts. We summarize key information regarding implementation fidelity for the full sample here. Instructors video recorded and logged each lesson in order to monitor implementation

fidelity. The research team coded the first two lesson videos to provide immediate feedback on implementation plus a randomly selected 50% of subsequent lessons. We measured adherence using the Nemours Fidelity Implementation Record (Nemours BrightStart!, 2016), a checklist on which raters score whether or not key elements of NBS! lessons were implemented as intended. The number of items on the checklist ranged from 55-75 per lesson and was specific to the content contained within each lesson. Per the developers, instructors are expected to implement 75% or more of these key elements. Interrater reliability for a randomly selected 20% of coded lessons was high, with an intraclass correlation (ICC) of .97. On average, teachers implemented 79% of key lesson elements (*SD* = 0.15) and community aides implemented 75% (*SD* = 0.11). Lesson logs indicated that, on average, teachers implemented 8 of the 20 lessons (*SD* = 6.83) and community aides implemented 16 (*SD* = 3.72). We also coded videos for quality of delivery and participant responsiveness/child engagement using the Quality of Intervention Delivery and Receipt observation tool (Harn et al., 2012); the ICC was .84 for the 20% of videos randomly selected for double coding. Quality of delivery, measured on a scale of 0 (*lack of implementation*) to 3 (*expert implementation*), averaged 2.16 (*SD* = 0.37) for teachers and 2.10 (*SD* = 0.30) for community aides. Participant responsiveness/child engagement, measured on a scale of 0 (*no or only one child responding*) to 3 (*all children responding*), averaged 2.57 (*SD* = 0.34) for teachers and 2.38 (*SD* = 0.30) for community aides.

Assessment Procedures and Measures

All children completed an initial early literacy screening, pretest assessments at the beginning of the academic year/prior to intervention and the same assessments as posttests at the end of the academic year. All assessments aligned with the emergent literacy skills targeted by the NBS! intervention and were conducted 1:1 by trained assessors in quiet spaces at children's

preschools. Research staff, supervised by the principal investigator and PhD-level project coordinator with expertise in literacy and language assessment, completed training on each assessment prior to administration. This included reviewing technical manuals and administration protocols, scoring at least 80% correct on administration quizzes, and completing practice administrations while being observed. Additional information for all measures is provided in the Supplemental Material.

Emergent Literacy: Get Ready to Read!-Revised

Children completed the Get Ready to Read!-Revised (GRTR-R; Whitehurst & Lonigan, 2010) at screening and posttest. The GRTR-R focuses on print knowledge and phonological awareness but is concurrently and longitudinally correlated with measures of oral language and reading comprehension (Phillips et al., 2009), supporting its validity as a measure of emergent literacy. We used the initial administration of the GRTR-R to identify children eligible to participate in the study. We re-administered the GRTR-R at posttest given that prior studies of NBS! have included it as an outcome measure (Bailet et al., 2009, 2013; Zettler-Greeley et al., 2018). Although all other assessments were conducted by research staff, GRTR-R was primarily administered by community aides as part of the local kindergarten readiness initiative's existing emergent literacy screening service; they completed the kindergarten readiness initiative's GRTR-R training. The number of correct responses is summed (max = 25; $\alpha = .80$). This sum is compared to age-based performance levels (i.e., benchmarks) to identify children at risk for later reading difficulties for screening purposes or used as raw score when considered as an outcome.

Print Knowledge

Research staff administered three measures of print knowledge. We assessed children's understanding of print concepts and forms (e.g., concept of a word, letter recognition) using the

Print Knowledge subtest of the Test of Preschool Early Literacy (TOPEL; Lonigan et al., 2007). Correct items are summed to provide the total score (max = 36; α = .97). We assessed children's letter-name knowledge using the Quick Letter Name Knowledge Assessment (Q-LNK; Tortorelli et al., 2017) and letter-sound knowledge using the Letter Sound Short Forms (Piasta et al., 2016). These measures provide item-response theory (IRT) scores on a scale of 0–26; IRT-derived reliabilities range from .89–.92 for the Q-LNK and .90–.93 for the Letter Sound Short Forms.

Phonological Awareness

Research staff administered three measures of phonological awareness. We assessed children's abilities to delete and blend sound units (words/syllables, phonemes) using the Phonological Awareness subtest of the TOPEL (Lonigan et al., 2007; max score = 27). We assessed children's rhyme awareness and initial sound awareness using the Rhyming and Alliteration subtests of the Individual Growth and Development Indicators of Early Literacy, 2nd edition (IGDIs; McConnell et al., 2012; max score on each = 15). For each of these measures, correct items are summed to provide total scores (α = .67, .89, and .64, respectively).

Language and Comprehension

Research staff administered three measures of language and language comprehension. We measured children's narrative language skills using the Narrative Assessment Protocol, 2nd edition (Bowles et al., 2020). The measure provides IRT-based scores with $M = 20$ ($SD = 2$); IRT-derived reliability is .81. We measured children's vocabulary using the Picture Naming subtest of the IGDIs (McConnell et al., 2012). Correct items are summed to provide the total score (max score = 15; α = .73). We measured children's listening comprehension using the Oral Comprehension subtest of the Woodcock-Johnson Tests of Achievement III (Woodcock et al., 2001, 2007), which provides IRT-based W scores (test-retest $r = .82$ per the manual).

Emergent Writing

We assessed three aspects of children's emergent writing using the tasks developed by Gerde and colleagues (see Gerde et al., 2015; Thomas et al., 2020). Children are provided with paper and writing utensils and asked to write their name, ten dictated letters (T, B, H, M, S, A, D, J, C, P), and five dictated consonant-vowel-consonant words (sad, hug, lip, net, job). *Name writing* is scored on an ordinal scale of 0 = refusal/no response, 1 = scribbling/drawing, 2 = use of letter-like shapes, 3 = use of at least one letter, or 4 = name spelled entirely correctly. *Letter writing* is scored similarly, with 0 = refusal/no response, 1 = scribbling/drawing, 2 = letter-like shape, 3 = a letter other than the letter dictated to the child, or 4 = correct letter; these scores are averaged across all ten letters ($\alpha = .92$). Invented spelling of the words is scored on an ordinal scale of 0 = refusal/no response, 1 = scribbling/drawing, 2 = use of letters or letter-like shapes, 3 = use of a letter to represent an initial or salient sound, 4 = use of letters to represent multiple sounds in the word, or 5 = word spelled correctly; these are averaged across all five words ($\alpha = .93$). Double scoring of a randomly selected 20% of writing samples indicated high interrater reliability, with ICCs ranging from .79 to .92 for the three emergent writing measures.

Intervention Dosage

Instructors reported intervention dosage for individual children on lesson logs. Given that each of the 20 lessons was split into two sessions per week, instructors recorded whether each child was present or absent for each of these sessions. In cases of child absence, instructors provided makeup lessons whenever possible, and these makeup lessons were also recorded. The research team checked all lesson logs and verified against submitted lesson videos. We measured intervention dosage by summing the total number of lessons that each child experienced, including makeup sessions, with each session counting as .5 of a lesson.

Results

We preregistered all analyses pursuant to the primary aims of this study on the Open Science Framework (Registration doi: 10.17605/OSF.IO/UWNRF). Prior to examining our three research aims, we conducted several preliminary analyses. First, we examined whether children assigned to the NBS! intervention and control conditions were initially equivalent on key child characteristics (i.e., gender, race and ethnicity, English as the primary language, whether the child had an IEP or 504 plan, and parents' highest degree) and all pretest measures. We conducted these analyses using SAS Proc Mixed and accounted for the nested structure of the data. Results, provided in the Supplemental Material (Table S.2), showed no significant differences between conditions on any of the examined child characteristics or pretest measures. Next, we examined the ratio of within- to between-classroom variance for each outcome of interest. We estimated ICCs based on unconditional models (i.e., models without predictors), fitting separate models for each outcome. ICCs ranged from .05–.28. Based on these findings, we used hierarchical linear modelling in all analyses to account for the nested structure of the data (i.e., children in small groups). We also estimated ICCs accounting for additional nesting within community aide (given that some community aides taught multiple small groups) but found that the variance components were too small to be estimable for all variables. Finally, we examined missing data. Missing data ranged from 0% to 16% across assessment points and measures. Through a series of hierarchical linear modeling analyses, we determined that missingness was unrelated to sample characteristics and was equivalent across conditions. We used SAS Proc MI to impute 30 datasets and conducted all subsequent analyses via SAS Proc Mixed using all 30 imputed datasets, with results combined using SAS Proc MIAnalyze.

Aims 1 and 2: Intent-to-Treat Effects of the NBS! Intervention

In our first two aims, we considered whether NBS!, as implemented by teachers or community aides, impacted children's print knowledge, phonological awareness, language and comprehension, and emergent writing skills under an intent-to-treat framework. As described in the preregistration, we tested this using a multilevel approach, with posttest scores as the dependent variables, pretest scores on the same measure as covariates, and variables representing teacher-implemented or community aide-implemented conditions as the independent variables of interest. We estimated separate models for each of the 13 outcomes, initially contrasting each NBS! condition with the control condition (Aim 1) and then re-estimating with the community aide-implemented condition as the reference group to contrast the two NBS! conditions (Aim 2).

Descriptive statistics for pretest and posttest scores are reported by condition in Table 1, and full results from the multilevel inferential analyses are available in the Supplemental Material (Table S.3). For Aim 1, there were no significant differences between the two NBS! conditions and the control condition for any child outcomes (all p-values > .05). Cohen's d effect sizes were generally small and ranged in absolute value from $d = 0.01$ to 0.25 with the exception of letter writing ($d = 0.37$ for community aide versus control comparison). For Aim 2, we found a significant difference between the teacher- and community aide-implemented conditions only on name writing ($d = .38$; favoring community aides). All other comparisons were non-significant, and most effect sizes were small ($d < 0.28$) with the exceptions of GRTR-R, name writing, and letter writing ($d = 0.38$ to 0.47; favoring community aides).

Aim 3: Impacts Accounting for Variation in Dosage

For Aim 3, we considered the impact of NBS! on outcomes after accounting for differences in intervention dosage, as not all children received all 20 lessons. The number of lessons received by children ranged from 0 to 20, with a mean of 5.52 ($SD = 6.10$) in the teacher-

implemented condition and mean of 12 (*SD* = 6.04) in the community aide-implemented condition. As described in the preregistration, we employed an instrumental variables approach to address this aim, following work by Angrist et al. (1996).

We conducted the two-stage instrumental variables analysis using SAS Proc Syslin. In the instrumental variables (first-stage) model, the number of lessons received by each child was predicted by the instrument, which was the randomly assigned study condition. In the second-stage model, children's posttest scores were predicted by the estimated number of lessons from the first-stage model, which represent the number of lessons received by children after selection biases were removed; pretest scores were also controlled. The key results are presented in Table 2. The instrumental variables analyses accounting for dosage did not show impacts of NBS! on any outcome except one: Children who received NBS! had higher scores on letter writing compared to those in the control condition (b = .02, p = .035).

Post-Hoc Exploration of Associations between Dosage and Outcomes

We hypothesized that NBS! would positively impact children's emergent literacy skills, particularly after accounting for dosage given correlational evidence suggesting that receiving more NBS! lessons was associated with better outcomes (Bailet et al., 2009; Zettler-Greeley et al., 2018). We thus conducted additional, post-hoc analyses not included in our preregistration to determine whether we could replicate previously reported correlations between dosage and outcomes with data from the current study.

These analyses mirrored those conducted for Aim 1, except that number of lessons replaced the condition variables as the independent variable of interest; pretest scores were again controlled. Note that, in these analyses, the number of lessons is simply based on the raw data; unlike the instrumental variables analyses, these do not account for potential selection biases that

may lead some children to receive different numbers of lessons. Results are reported in Table 2. The number of lessons received was positively associated with children's posttest scores on the GRTR-R (b = .08, p = .027), print knowledge (b = .22, p = .006), letter-name knowledge (b = .37, p = .003), and letter writing (b= .01, p = .050) but not any of the other outcomes.

Discussion

NBS! is a small-group, emergent literacy intervention that has shown promise for improving early literacy skills when implemented under highly controlled conditions. In the current study, we examined the impacts of the commercially available, 20-lesson version of the intervention when used by teachers and community aides in authentic preschool classrooms. Our results speak to the ongoing challenge of moving research-based interventions towards scale up and have implications for analysis of dosage effects and revising the NBS! theory of change.

Our major finding was that we did not replicate results of prior efficacy trials indicating positive impacts of NBS! relative to control on children's print knowledge, phonological awareness, or GRTR-R outcomes (Bailet et al., 2009, 2013; Zettler-Greeley et al., 2018), nor did we expand this work to show impacts on language and comprehension skills or emergent writing skills, with the one exception of letter writing. We did not find any significant effects in our intent-to-treat analyses, and the effect on letter writing was statistically significant only in our as-treated analyses. Most effect sizes relative to control were less than 0.30, which we were not statistically powered to detect but were also below our threshold for practically meaningful effects based on the literature (Chatterji, 2006; Kuhfeld et al., 2020; Lipsey et al., 2012). Effect sizes were also much lower in the current study (e.g., d = -0.22–0.24 for print knowledge) for constructs also assessed in prior efficacy trials (e.g., d = 0.20–0.74 for print knowledge). Although the effect on letter writing may be notable, given the need to better support emergent

writing during preschool (Gerde et al., 2015; Zhang et al., 2015), we caution that this was the sole effect out of many outcomes and only statistically significant after accounting for variation in dosage. With respect to language and comprehension, our failure to find effects echoes recent work highlighting the difficulty of improving these outcomes via classroom-based instruction (Dickinson, 2011; Haley et al., 2017; Rogde et al., 2019; cf. Spencer et al., 2012; Spencer et al., 2015) especially when using standardized rather than intervention-aligned measures (Herrera et al., 2021). Importantly, differences in children's outcomes relative to control were minimal, indicating that the null results were neither statistically significant nor practically meaningful, and also not due to lack of statistical power. Rather, it appears that NBS!, as implemented in the current study, did not improve most emergent skills relative to typical instruction.

By design, the current study differed from prior efficacy trials of NBS! in several ways. These differences supported our goal of understanding intervention impacts when implemented by intended end users under more authentic conditions. While such testing for generalizability is critical when working towards scale up (Gottfredson et al., 2015), these differences need to be considered in light of the largely null findings. One difference was that the study was led by an independent research team rather than the developers; evaluations conducted by intervention developers often show more positive effects than those conducted independently (Wolf et al., 2020). A second difference was that prior efficacy trials tested the penultimate, 18-lesson version of NBS!; however, the revisions to create the commercially available, 20-lesson version tested in this study were minor and not anticipated to decrease effects. The early childhood policy and funding context was yet another difference between past and current studies. The state in which initial efficacy trials were conducted has an established voluntary prekindergarten program, in which language and literacy learning is emphasized and centrally tracked, reported, and

supported by the state. The current study, in contrast, was conducted in a state in which early childhood programming is less centralized with greater local control and minimal accountability for outcomes. This larger ecology may have influenced results in terms of the perceived need for and commitment to emergent literacy intervention (Durlak & DuPre, 2008).

Another difference was in the study sample. The current study involved a greater proportion of children from lower socioeconomic backgrounds than prior NBS! efficacy trials, and this included children enrolled in Head Start classrooms. More than 50% of participating children had family incomes that fell at or below the federal poverty level, which is more than double the child poverty rate in the U.S. (Jiang et al., 2017) and in prior NBS! studies (Zettler-Greeley, 2018). The preponderance of study participants from low-income families is not surprising, given documented opportunity gaps, and thus differences in school-entry literacy skills, for children from lower socioeconomic backgrounds (Chatterji, 2006; Kuhfeld et al., 2020). Of note, children from lower socioeconomic backgrounds may have lower language capabilities than their more affluent peers, which may, in turn, further impact literacy skill development (Lonigan et al., 2009). Also, whereas past studies involved only those children matriculating to kindergarten the following year, the current study was open to all children in participating preschool classrooms who qualified. It may be that some of the younger children in the sample would no longer be identified as needing extra literacy support after receiving language and literacy learning opportunities during preschool. These child and family factors and others (e.g., oral language ability; Kruse et al., 2015) may moderate effects and will be examined in future work.

The major difference from past NBS! efficacy trials was that the intervention was implemented by classroom teachers, as the intended end users, and by community aides, as an

important step in moving towards effectiveness or scale-up studies (Gottfredson et al., 2015). Collectively, these instructors may have been less prepared to use NBS! as compared to the early literacy specialists involved in past efficacy trials. Most early literacy specialists had masters degrees in education, early childhood, or a related field, along with prior experience as licensed teachers. In contrast, the vast majority of classroom teachers implementing in the current study did not have advanced degrees, and very few teachers or community aides held teaching licenses. This is not atypical in the preschool sector (Rhodes & Huston, 2012). Although education level, per se, is not consistently related to instructional quality or early literacy outcomes (e.g., Lin & Magnuson, 2018), many preschool teachers and community aides in the current study may not have had opportunities to learn about small-group instruction or build the requisite classroom and behavior management skills for implementing small groups (Wyatt & Chapman-DeSousa, 2017; see also Zucker et al., 2013).

The teachers and community aides also had less familiarity with NBS! compared to the early literacy specialists. The early literacy specialists practiced their implementation with numerous small groups each year, often across multiple years, which may refine implementation (Clements et al., 2015). They also received ongoing support and training from the NBS! developers, to whom they directly reported. The teachers and community aides were in their initial years of using the intervention and received the 2-day professional development training that accompanies purchase of NBS!. Although a 2-day training may not be optimal (Hamre et al., 2017), similar levels of professional development have been effective in prior studies involving classroom teachers or community aides (Al Otaiba et al., 2005; Bleses et al., 2018; Gonzalez et al., 2011; Nelson et al., 2009), and this sort of training realistically aligns with what preschool teachers often receive, given cost and time constraints. In the current study, this professional

development appeared to be sufficient to enable most teachers to implement NBS! reasonably well. Only six teachers scored below the 75% adherence benchmark set by the developers, and both teachers and community aides exhibited relatively high quality of delivery and participant responsiveness/child engagement. Five community aides, serving 15 small groups, did not meet the adherence benchmark, however. On one hand, it is likely that lower adherence attenuated effects. On the other, we saw no advantage for implementation by teachers, who exhibited stronger adherence over community aides (see also Piasta et al., 2021), signaling that adherence did not seem to determine effects.

Dosage was considerably lower than in previous NBS! efficacy trials. In some cases, limited dosage was due to reasons typical of the preschool context (e.g., children moving classrooms). As we projected, however, small-group intervention proved challenging for many preschool teachers, who cited time, resource, and classroom/behavior management constraints as barriers to implementation when asked to comment on what made completing lessons difficult (see Piasta et al., 2021). This resulted in less-than-recommended dosage. In anticipation of this, we included the community aide-implemented condition and made the a priori decision to conduct as-treated analyses that accounted for dosage variation via instrumental variables analyses. Yet, although community aides achieved higher dosage and, in a few cases, better child outcomes relative to teachers, there were no impacts compared to the control condition. Moreover, the instrumental variables analyses failed to demonstrate effects with the single exception of letter writing. Based on these findings, we cannot conclude that the lack of effects was due to low dosage. Prior work did not identify differences between teachers and community aides in their quality of delivery or participant responsiveness during lessons, although teachers had significantly higher adherence (Piasta et al., 2021). Yet, in the current study, we did not find

better outcomes for the teacher-implemented condition. Thus, it is not readily apparent that any of the commonly used implementation metrics (dosage, adherence, quality of delivery, participant responsiveness; Durlak & DuPre, 2008), as applied in this study, were solely responsible for our inability to replicate prior effects. Rather, we may need to consider more complex patterns of associations between NBS! implementation and learning gains (Suggate, 2016), along with more nuanced aspects of implementation (e.g., how well instructors paced or differentiated instruction, or managed group dynamics), and refine the theory of change accordingly.

Our correlational dosage analysis partially replicated previous findings. Specifically, we also found positive associations between number of lessons and children's outcomes on the GRTR-R, two measures of print knowledge, and letter writing. We did not replicate associations previously found for phonological awareness or vocabulary, despite use of the same or highly similar measures (Bailet et al., 2009; Zettler-Greeley et al., 2018), nor did we identify any associations with language and comprehension or other emergent writing skills. At first glance, the replication of some associations seems at odds with findings from the instrumental variables analyses, in which the number of lessons was not significantly associated with outcomes beyond letter writing. Our interpretation is that correlations between dosage and outcomes are due to confounding variables; children who receive more lessons are systematically different from those who receive fewer lessons. Such selection biases are accounted for in the instrumental variables analyses. These findings highlight the need for using this type of rigorous approach to as-treated analyses and caution in interpreting correlational dosage analyses (Sagarin et al., 2014) within the education literature. Children who differ in dosage may also differ in a variety of other factors, such as attendance, initial skill levels, motivation, self-regulation/behavioral skills, and

child-teacher and peer relationships, which may affect intervention delivery, or may attend classrooms that differ in the extent to which intervention is likely to be provided. Identifying factors that predict dosage is an important future direction not only to inform methodology and statistical analysis, but also to better support preschool programs and teachers as they work to provide intervention to all children who might benefit, rather than a select subgroup.

Although disappointing, our findings are aligned with other work showing smaller—or null—effects as interventions are implemented by intended end users or at larger scale (Bleses et al., 2018; de Boer et al., 2014; Kim, 2019). The modest effects in the current and other studies require continued research examining the reasons behind these attenuated effects to understand aspects of the intervention or its implementation that can be strengthened. Acknowledging that NBS! was not effective as implemented in this study is important for the sake of scientific transparency (Polanin et al., 2016) and to guide continued work on NBS! and other small-group, emergent literacy interventions such that these ultimately can be brought to scale. The current findings compel further attention to the NBS! theory of change and elements necessary to realize intended effects. As alluded to above, the theory of change may need to acknowledge (1) key structural and contextual factors such as the perceived need for supplemental emergent literacy intervention and availability of staff, space, time, and other resources (Durlak & DuPre, 2008); (2) the critical role of classroom routines and management affording small-group instruction; (3) required expertise of instructors; (4) additional, more nuanced aspects of implementation. The NBS! professional development may also require changes, as grounded in the extant literature (Hamre et al., 2017; Wyatt & Chapman-DeSousa, 2017; Zucker et al., 2013), informed by lessons learned over the course of this study (see also Piasta et al., 2021), and aligned with a revised theory of change (e.g., incorporation of strategies for small-group and behavior

management, more opportunities to practice lessons and ongoing feedback, sharing implementation challenges and brainstorming solutions). As acknowledged in other work (e.g., Hamre et al., 2017), ongoing support and troubleshooting through coaching or communities of practice may be valuable. In these ways, and coupled with insights into dosage analysis, our null results contribute to improving the "rigor, relevance, and reproducibility" (Kim, 2019, p. 599) of research on small-group emergent literacy interventions. These indispensable insights impel hypothesis generation for future studies as we work to achieve intended effects of NBS! and similar interventions when used under more routine, authentic conditions.

References

Al Otaiba, S., Schatschneider, C., & Silverman, E. (2005). Tutor-assisted intensive learning strategies in kindergarten: How much is enough? *Exceptionality, 13*(4), 195-208. https://doi.org/10.1207/s15327035ex1304_2

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association, 91*(434), 444-455. http://www.jstor.org/stable/2291629

Anthony, J. L., & Francis, D. J. (2005). Development of phonological awareness. *Current Directions in Psychological Science, 14*(5), 255-259. https://doi.org/10.1111/j.0963-7214.2005.00376.x

Bailet, L. L., Repper, K., Murphy, S., Piasta, S., & Zettler-Greeley, C. (2013). Emergent literacy intervention for prekindergarteners at risk for reading failure: Years 2 and 3 of a multiyear study. *Journal of Learning Disabilities, 46*(2), 133-153. https://doi.org/10.1177/0022219411407925

Bailet, L. L., Repper, K. K., Piasta, S. B., & Murphy, S. P. (2009). Emergent literacy intervention for prekindergarteners at risk for reading failure. *Journal of Learning Disabilities, 42*(4), 336-355. https://doi.org/10.1177/0022219409335218

Bleses, D., Højen, A., Justice, L. M., Dale, P. S., Dybdal, L., Piasta, S. B., Markussen-Brown, J., Clausen, M., & Haghish, E. F. (2018). The effectiveness of a large-scale language and preliteracy intervention: The SPELL randomized controlled trial in Denmark. *Child Development, 89*, 342-363. https://doi.org/10.1111/cdev.12859

Bowles, R. P., Justice, L. M., Khan, K. S., Piasta, S. B., Skibbe, L. E., & Foster, T. D. (2020). Development of the Narrative Assessment Protocol-2: A tool for examining young

children's narrative skill. *Language, Speech, and Hearing Services in Schools, 51*(2), 390-404. https://doi.org/doi:10.1044/2019_LSHSS-19-00038

Byrne, B., & Fielding-Barnsley, R. (1991). Evaluation of a program to teach phonemic awareness to young children. *Journal of Educational Psychology, 83*(4), 451-455. https://doi.org/10.1037/0022-0663.83.4.451

Byrne, B., & Fielding-Barnsley, R. (1995). Evaluation of a program to teach phonemic awareness to young children: A 2- and 3-year follow-up and a new preschool trial. *Journal of Educational Psychology, 87*(3), 488-503. https://doi.org/10.1037/0022-0663.87.3.488

Chatterji, M. (2006). Reading achievement gaps, correlates, and moderators of early reading achievement: Evidence from the Early Childhood Longitudinal Study (ECLS) kindergarten to first grade sample. *Journal of Educational Psychology, 98*(3), 489-507. https://doi.org/10.1037/0022-0663.98.3.489

Clements, D. H., Sarama, J., Wolfe, C. B., & Spitler, M. E. (2015). Sustainability of a scale-up intervention in early mathematics: A longitudinal evaluation of implementation fidelity. *Early Education and Development, 26*(3), 427-449. https://doi.org/10.1080/10409289.2015.968242

de Boer, H., Donker, A. S., & van der Werf, M. P. C. (2014). Effects of the attributes of educational interventions on students' academic performance: A meta-analysis. *Review of Educational Research, 84*(4), 509-545. https://doi.org/10.3102/0034654314540006

Dickinson, D. K. (2011). Teachers' language practices and academic outcomes of preschool children. *Science, 333*(6045), 964-967. https://doi.org/10.1126/science.1204526

Dickinson, D. K., McCabe, A., & Clark-Chiarelli, N. (2004). Preschool-based prevention of reading disability. In C. A. Stone, E. R. Sillliman, B. J. Ehren, & K. Apel (Eds.), *Handbook of language and literacy* (pp. 209-227). Guilford Press.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*(3-4), 327-350. https://doi.org/10.1007/s10464-008-9165-0

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly, 36*(3), 250-287. https://doi.org/10.1598/RRQ.36.3.2

Farley, K. S., Piasta, S., Dogucu, M., & O'Connell, A. (2017). Assessing and predicting small-group literacy instruction in early childhood classrooms. *Early Education and Development, 28*(4), 488-505. https://doi.org/10.1080/10409289.2016.1250549

Ferrer, E., Shaywitz, B. A., Holahan, J. M., Marchione, K. E., Michaels, R., & Shaywitz, S. E. (2015). Achievement gap in reading is present as early as first grade and persists through adolescence. *The Journal of Pediatrics, 167*(5), 1121-1125.e1122. https://doi.org/https://doi.org/10.1016/j.jpeds.2015.07.045

Fielding, L., Kerr, N., & Rosier, P. (2007). *Annual growth for all students: Catch up growth for those who are behind*. The New Foundation Press.

Fitzgerald, T., Robillard, L., & O'Grady, A. (2018). Exploring the impact of a volunteer shared reading programme on preschool-aged children. *Early Child Development and Care, 188*(6), 851-861. https://doi.org/10.1080/03004430.2016.1240679

Gerde, H. K., Bingham, G. E., & Pendergast, M. L. (2015). Reliability and validity of the Writing Resources and Interactions in Teaching Environments (WRITE) for preschool classrooms. *Early Childhood Research Quarterly, 31*, 34-46. https://doi.org/10.1016/j.ecresq.2014.12.008

Gettinger, M., & Stoiber, K. (2008). Applying a response-to-intervention model for early literacy development in low-income children. *Topics in Early Childhood Special Education, 27*(4), 198-213. https://doi.org/10.1177/0271121407311238

Goldstein, H., Olszewski, A., Haring, C., Greenwood, C. R., McCune, L., Carta, J., Atwater, J., Guerrero, G., Schneider, N., McCarthy, T., & Kelley, E. S. (2017). Efficacy of a supplemental phonemic awareness curriculum to instruct preschoolers with delays in early literacy development. *Journal of Speech, Language, and Hearing Research, 60*(1), 89-103. https://doi.org/10.1044/2016_JSLHR-L-15-0451

Gonzalez, J. E., Pollard-Durodola, S., Simmons, D. C., Taylor, A. B., Davis, M. J., Kim, M., & Simmons, L. (2011). Developing low-income preschoolers' social studies and science vocabulary knowledge through content-focused shared book reading. *Journal of Research on Educational Effectiveness, 4*(1), 25-52. https://doi.org/10.1080/19345747.2010.487927

Gottfredson, D., Cook, T., Gardner, F. M., Gorman-Smith, D., Howe, G., Sandler, I., & Zafft, K. (2015). Standards of evidence for efficacy, effectiveness, and scale-up research in prevention science: Next generation. *Prevention Science, 16*(7), 898-926. https://doi.org/10.1007/s11121-015-0555-x

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*, 6-10. https://doi.org/10.1177/074193258600700104

Greenwood, C. R., Carta, J. J., Atwater, J., Goldstein, H., Kaminski, R., & McConnell, S. (2013). Is a response to intervention (RTI) approach to preschool language and early literacy instruction needed? *Topics in Early Childhood Special Education, 33*(1), 48-64. https://doi.org/10.1177/0271121412455438

Haley, A., Hulme, C., Bowyer-Crane, C., Snowling, M. J., & Fricke, S. (2017). Oral language skills intervention in pre-school – a cautionary tale. *International Journal of Language and Communication Disorders, 52*(1), 71-79. https://doi.org/10.1111/1460-6984.12257

Hamre, B. K., Partee, A., & Mulcahy, C. (2017). Enhancing the impact of professional development in the context of preschool expansion. *AERA Open, 3*(4), 2332858417733686. https://doi.org/10.1177/2332858417733686

Harn, B. A., Forbes-Spear, C., Fritz, R., & Berg, T. (2012). *Quality of intervention delivery and receipt observation tool*. University of Oregon.

Herrera, S., Phillips, B. M., Newton, Y., Dombek, J. L., & Hernandez, J. A. (2021). *Effectiveness of early literacy instruction: Summary of 20 years of research* (REL 2021–084). https://ies.ed.gov/ncee/edlabs/projects/project.asp?projectID=4564

Hjetland, H. N., Brinchmann, E. I., Scherer, R., Hulme, C., & Melby-Lervåg, M. (2020). Preschool pathways to reading comprehension: A systematic meta-analytic review. *Educational Research Review, 30*, 100323. https://doi.org/10.1016/j.edurev.2020.100323

Jiang, Y., Granja, M. R., & Koball, H. (2017). *Basic facts about low-income children: Children under 6 years, 2015*. National Center for Children in Poverty.

Justice, L. M., Chow, S.-M., Capellini, C., Flanigan, K., & Colton, S. (2003). Emergent literacy intervention for vulnerable preschoolers: Relative effects of two approaches. *American*

Journal of Speech-Language Pathology, 12(3), 320-332. https://doi.org/10.1044/1058-0360(2003/078)

Kelley, E. S., Goldstein, H., Spencer, T. D., & Sherman, A. (2015). Effects of automated tier 2 storybook intervention on vocabulary and comprehension learning in preschool children with limited oral language skills. *Early Childhood Research Quarterly, 31*, 47-61. https://doi.org/10.1016/j.ecresq.2014.12.004

Kim, J. S. (2019). Making every study count: Learning from replication failure to improve intervention research. *Educational Researcher, 48*(9), 599-607. https://doi.org/10.3102/0013189x19891428

Koutsoftas, A. D., Harmon, M. T., & Gray, S. (2009). The effect of tier 2 intervention for phonemic awareness in a response-to-intervention model in low-income preschool classrooms. *Language, Speech, and Hearing Services in Schools, 40*(2), 116-130. https://doi.org/10.1044/0161-1461(2008/07-0101)

Kruse, L. G., Spencer, T. D., Olszewski, A., & Goldstein, H. (2015). Small groups, big gains: Efficacy of a tier 2 phonological awareness intervention with preschoolers with early literacy deficits. *American Journal of Speech-Language Pathology, 24*(2), 189-205. https://doi.org/10.1044/2015_ajslp-14-0035

Kuhfeld, M., Soland, J., Pitts, C., & Burchinal, M. (2020). Trends in children's academic skills at school entry: 2010 to 2017. *Educational Researcher, 49*(6), 403-414. https://doi.org/10.3102/0013189x20931078

Lin, Y.-C., & Magnuson, K. A. (2018). Classroom quality and children's academic skills in child care centers: Understanding the role of teacher qualifications. *Early Childhood Research Quarterly, 42*, 215-227. https://doi.org/10.1016/j.ecresq.2017.10.003

Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., Roberts, M., Anthony, K. S., & Busick, M. D. (2012, November). *Translating the statistical representation of the effects of education interventions into more readily interpretable forms*. https://ies.ed.gov/ncser/pubs/20133000/

Lonigan, C. J., Anthony, J. L., Phillips, B. M., Purpura, D. J., Wilson, S. B., & McQueen, J. D. (2009). The nature of preschool phonological processing abilities and their relations to vocabulary, general cognitive abilities, and print knowledge. *Journal of Educational Psychology, 101*(2), 345-358. https://doi.org/10.1037/a0013837

Lonigan, C. J., & Phillips, B. M. (2016). Response to instruction in preschool: Results of two randomized studies with children at significant risk of reading difficulties. *Journal of Educational Psychology, 108*(1), 114-129. https://doi.org/10.1037/edu0000054

Lonigan, C. J., Purpura, D. J., Wilson, S. B., Walker, P. M., & Clancy-Menchetti, J. (2013). Evaluating the components of an emergent literacy intervention for preschool children at risk for reading difficulties. *Journal of Experimental Child Psychology, 114*(1), 111-130. https://doi.org/10.1016/j.jecp.2012.08.010

Lonigan, C. J., Wagner, R. K., & Torgesen, J. K. (2007). *Test of Preschool Early Literacy*. Pro-Ed.

Markovitz, C. E., Hernandez, M. W., Hedberg, E., & Silbergelitt, B. (2015). *Outcome evaluation of the Minnesota Reading Corps PreK program*. NORC at the University of Chicago.

McConnell, S. R., Bradfield, T., Wackerle-Hollman, A. K., & Rodriguez, M. (2012). *Individual growth and development indicators of early literacy* (2nd ed.). Regents of the University of Minnesota.

National Early Literacy Panel. (2008). *Developing early literacy*. National Institute for Literacy. www.nifl.gov

Nelson, J. R., Sanders, E. A., & Gonzalez, J. (2009). The efficacy of supplemental early literacy instruction by community-based tutors for preschoolers enrolled in Head Start. *Journal of Research on Educational Effectiveness, 3*(1), 1-25. https://doi.org/10.1080/19345740903381031

Nemours BrightStart! (2016). *Nemours fidelity implementation record*. Nemours Children's Health System.

Pearson, P. D., & Gallagher, M. C. (1983). The instruction of reading comprehension. *Contemporary Educational Psychology, 8*(3), 317-344. https://doi.org/10.1016/0361-476X(83)90019-X

Phillips, B. M., Lonigan, C. J., & Wyatt, M. A. (2009). Predictive validity of the Get Ready to Read! Screener:Concurrent and long-term relations with reading-related skills. *Journal of Learning Disabilities, 42*(2), 133-147. https://doi.org/10.1177/0022219408326209

Phillips, B. M., Tabulda, G., Ingrole, S. A., Burris, P. W., Sedgwick, T. K., & Chen, S. (2016). Literate language intervention with high-need prekindergarten children: A randomized trial. *Journal of Speech, Language, and Hearing Research, 59*(6), 1409-1420. https://doi.org/10.1044/2016_JSLHR-L-15-0155

Piasta, S. B. (2016). Current understandings of what works to support the development of emergent literacy in early childhood classrooms. *Child Development Perspectives, 10*(4), 234-239. https://doi.org/10.1111/cdep.12188

Piasta, S. B., Logan, J. A. R., Thomas, L. J. G., Zettler-Greeley, C. M., Bailet, L. L., & Lewis, K. (2021). Implementation of a small-group emergent literacy intervention by preschool

teachers and community aides. *Early Childhood Research Quarterly, 54*, 31-43. https://doi.org/10.1016/j.ecresq.2020.08.002

Piasta, S. B., Phillips, B. M., Williams, J. M., Bowles, R. P., & Anthony, J. L. (2016). Measuring young children's alphabet knowledge: Development and validation of brief letter-sound knowledge assessments. *The Elementary School Journal, 116*(4), 523-548. https://doi.org/10.1086/686222

Piasta, S. B., & Wagner, R. K. (2010). Developing emergent literacy skills: A meta-analysis of alphabet learning and instruction. *Reading Research Quarterly, 45*(1), 8-38. https://doi.org/10.1598/RRQ.45.1.2

Polanin, J. R., Tanner-Smith, E. E., & Hennessy, E. A. (2016). Estimating the difference between published and unpublished effect sizes: A meta-review. *Review of Educational Research, 86*(1), 207-236. https://doi.org/10.3102/0034654315582067

Pollard-Durodola, S. D., Gonzalez, J. E., Simmons, D. C., Kwok, O., Taylor, A. B., Davis, M. J., Minjung, K. I. M., & Simmons, L. (2011). The effects of an intensive shared book-reading intervention for preschool children at risk for vocabulary delay. *Exceptional Children, 77*(2), 161-183. https://doi.org/10.1177/001440291107700202

Rhodes, H., & Huston, A. (2012). *Building the workforce our youngest children deserve.* (Vol. 26). Society for Research in Child Development.

Ritter, G., Denny, G., Albin, G., Barnett, J., & Blankenship, V. (2006). The effectiveness of volunteer tutoring programs: A systematic review. *Campbell Systematic Reviews, 2*(1), 1-63. https://doi.org/10.4073/csr.2006.7

Ritter, G. W., Barnett, J. H., Denny, G. S., & Albin, G. R. (2009). The effectiveness of volunteer tutoring programs for elementary and middle school students: A meta-analysis. *Review of Educational Research, 79*(1), 3-38. https://doi.org/10.3102/0034654308325690

Rogde, K., Hagen, Å. M., Melby-Lervåg, M., & Lervåg, A. (2019). The effect of linguistic comprehension instruction on generalized language and reading comprehension skills: A systematic review. *Campbell Systematic Reviews, 15*(4), e1059. https://doi.org/10.1002/cl2.1059

Sagarin, B. J., West, S. G., Ratnikov, A., Homan, W. K., Ritchie, T. D., & Hansen, E. J. (2014). Treatment noncompliance in randomized experiments: Statistical approaches and design issues. *Psychological Methods, 19*(3), 317-333. https://doi.org/10.1037/met0000013

Savage, R., & Carless, S. (2005). Learning support assistants can deliver effective reading interventions for 'at-risk' children. *Educational Research, 47*(1), 45-61. https://doi.org/10.1080/0013188042000337550

Savage, R., Carless, S., & Stuart, M. (2003). The effects of rime- and phoneme-based teaching delivered by learning support assistants. *Journal of Research in Reading, 26*(3), 211-233. https://doi.org/10.1111/1467-9817.00199

Schneider, W., Roth, E., & Ennemoser, M. (2000). Training phonological skills and letter knowledge in children at risk for dyslexia: A comparison of three kindergarten intervention programs. *Journal of Educational Psychology, 92*(2), 284-295. https://doi.org/10.1037/0022-0663.92.2.284

Spencer, E. J., Goldstein, H., Sherman, A., Noe, S., Tabbah, R., Ziolkowski, R., & Schneider, N. (2012). Effects of an automated vocabulary and comprehension intervention: An early

efficacy study. *Journal of Early Intervention, 34*(4), 195-221. https://doi.org/10.1177/1053815112471990

Spencer, T. D., Petersen, D. B., & Adams, J. L. (2015). Tier 2 language intervention for diverse preschoolers: An early-stage randomized control group study following an analysis of response to intervention. *American Journal of Speech-Language Pathology, 24*(4), 619-636. https://doi.org/10.1044/2015_AJSLP-14-0101

Suggate, S. P. (2016). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *Journal of Learning Disabilities, 49*(1), 77-96. https://doi.org/10.1177/0022219414528540

Thomas, L. J. G., Gerde, H. K., Piasta, S. B., Logan, J. A. R., Bailet, L. L., & Zettler-Greeley, C. M. (2020). The early writing skills of children identified as at-risk for literacy difficulties. *Early Childhood Research Quarterly, 51*, 392-402. https://doi.org/10.1016/j.ecresq.2020.01.003

Tortorelli, L. S., Bowles, R. P., & Skibbe, L. E. (2017). Easy as AcHGzrjq: The quick letter name knowledge assessment. *The Reading Teacher, 71*(2), 145-156. https://doi.org/10.1002/trtr.1608

Wasik, B. A. (2008). When fewer is more: Small groups in early childhood classrooms. *Early Childhood Education Journal, 35*(6), 515-521. https://doi.org/10.1007/s10643-008-0245-4

Whitehurst, G. J., & Lonigan, C. J. (2010). *Get read to read!—revised edition*. Pearson Assessments.

Wolf, R., Morrison, J., Inns, A., Slavin, R., & Risman, K. (2020). Average effect sizes in developer-commissioned and independent evaluations. *Journal of Research on*

Educational Effectiveness, 13(2), 428-447. https://doi.org/10.1080/19345747.2020.1726537

Woodcock, R., McGrew, K. S., & Mather, N. (2001, 2007). *Woodcock-Johnson Tests of Achievement (3rd ed., normative update)*. Riverside.

Wyatt, T., & Chapman-DeSousa, B. (2017). Teaching as interaction: Challenges in transitioning teachers' instruction to small groups. *Early Childhood Education Journal, 45*(1), 61-70. https://doi.org/10.1007/s10643-015-0758-6

Zettler-Greeley, C. M., Bailet, L. L., Murphy, S., DeLucca, T., & Branum-Martin, L. (2018). Efficacy of the Nemours BrightStart! Early literacy program: Treatment outcomes from a randomized trial with at-risk prekindergartners. *Early Education and Development, 29*, 873-892. https://doi.org/10.1080/10409289.2018.1475202

Zhang, C., Hur, J., Diamond, K., & Powell, D. (2015). Classroom writing environments and children's early writing skills: An observational study in Head Start classrooms. *Early Childhood Education Journal, 43*(4), 307-315. https://doi.org/10.1007/s10643-014-0655-4

Zucker, T. A., Solari, E. J., Landry, S. H., & Swank, P. R. (2013). Effects of a brief tiered language intervention for prekindergartners at risk. *Early Education and Development, 24*(3), 366-392. https://doi.org/10.1080/10409289.2012.664763

Table 1

Descriptive Statistics for Child Outcomes (Raw Data) Overall and by NBS! Condition

Child outcome	Pretest					Posttest				
	n	*M*	*SD*	Min	Max	*n*	*M*	*SD*	Min	Max
Control condition (31 classrooms)										
Get Ready to Read- Revised	86	7.95	2.75	2	13	77	13.68	4.87	3	25
Print knowledge	85	6.51	5.93	0	29	72	16.33	9.8	2	35
Letter-name knowledge	86	7.9	11.01	1	48	77	20.62	15.26	1	48
Letter-sound knowledge	86	2.04	2.37	1	16	77	6.33	6.68	1	24
Phonological awareness	86	8.91	4.65	0	22	77	11.7	4.87	1	24
Rhyme awareness	86	2.01	2.88	0	10	77	4.58	4.37	0	14
Initial sound awareness	86	6.16	3.02	0	15	75	7.48	3.49	0	15
Narrative language	86	18.52	2.05	14	22	77	19.7	1.8	14	23
Vocabulary	86	4.03	2.7	0	11	77	6.14	2.91	1	13
Listening comprehension	86	436.69	14.7	418	475	77	444.7	14.43	418	475
Name writing	86	1.77	0.89	1	4	77	2.86	0.96	1	4
Letter writing	86	1.36	0.56	1	4	77	2.09	0.88	1	4
Invented spelling	86	1.07	0.28	0	2	77	1.41	0.58	1	4
Teacher-implemented condition (33 classrooms)										
Get Ready to Read- Revised	101	7.42	2.89	2	13	92	13.02	4.58	4	23
Print knowledge	99	5.78	4.9	0	28	91	13.4	9.63	0	34
Letter-name knowledge	101	7.58	10.34	1	42	93	18.06	15.15	1	48
Letter-sound knowledge	101	2.28	3.2	1	18	93	6.28	6.24	1	24
Phonological awareness	101	8.77	4.44	0	22	93	11.62	4.66	0	22
Rhyme awareness	101	1.39	1.87	0	8	93	3.42	4.51	0	15
Initial sound awareness	101	5.63	2.89	0	12	89	6.94	2.77	0	15

Child outcome	Pretest					Posttest				
	n	*M*	*SD*	Min	Max	*n*	*M*	*SD*	Min	Max
Narrative language	101	18.3	1.96	14	22	93	19.31	1.87	14	22
Vocabulary	99	3.83	2.35	0	11	93	5.99	2.88	0	12
Listening comprehension	100	434.68	13.72	418	471	93	442.23	14.72	418	485
Name writing	101	1.74	0.97	0	4	91	2.57	1.01	1	4
Letter writing	101	1.36	0.59	0	3	93	2.05	0.86	1	4
Invented spelling	101	1.08	0.29	0	2	93	1.31	0.46	0	3
Community aide-implemented condition (34 classrooms)										
Get Ready to Read- Revised	94	8.13	2.66	2	16	76	14.33	4.75	5	24
Print knowledge	94	5.77	5.09	0	32	74	16.38	9.46	1	35
Letter-name knowledge	94	7.14	9.33	1	48	76	21.14	16.02	1	48
Letter-sound knowledge	94	2.93	4.43	1	24	76	6.87	6.38	1	24
Phonological awareness	94	8.96	4.5	0	19	76	12.2	4.77	3	23
Rhyme awareness	94	1.66	2.78	0	15	76	4.16	4.7	0	15
Initial sound awareness	94	5.67	3.16	0	13	73	7.34	3.5	0	15
Narrative language	94	18.49	1.99	14	22	76	19.73	1.76	14	23
Vocabulary	93	3.87	2.54	0	11	75	6.32	2.86	1	11
Listening comprehension	94	434.55	12.19	418	468	76	443.14	13.77	418	471
Name writing	93	1.95	1.04	0	4	74	3.08	0.68	1	4
Letter writing	93	1.35	0.56	0	3	75	2.31	0.87	1	4
Invented spelling	93	1.17	0.41	0	3	75	1.41	0.6	0	4

Note. We used multilevel models, controlling for pretest, to compare conditions at posttest; these analyses included all 281 participating children via multiple imputation. We found no significant differences relative to control and one significant difference between NBS! conditions; see Supplemental Material for full results).

Table 2

Results of Second-Stage Instrumental Variables Analyses and Replication Analyses: Coefficients Associating Number of Lessons with Child Outcomes

	Second Stage Least Squares				Replication analyses			
Outcome of interest	*b*	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*
Get Ready to Read-Revised	0.03	0.05	0.65	.514	0.08	0.04	2.22	.027
Print knowledge	0.12	0.10	1.21	.227	0.22	0.08	2.76	.006
Letter-name knowledge	0.13	0.15	0.88	.377	0.37	0.12	2.98	.003
Letter-sound knowledge	0.03	0.07	0.40	.688	0.09	0.06	1.53	.126
Phonological awareness	0.03	0.05	0.71	.479	0.01	0.03	0.35	.725
Rhyme awareness	-0.01	0.05	-0.16	.876	0.05	0.04	1.33	.183
Initial sound awareness	0.00	0.04	0.06	.950	0.02	0.03	0.76	.448
Narrative language	0.01	0.02	0.37	.714	0.01	0.01	0.58	.563
Vocabulary	0.03	0.02	1.30	.194	0.03	0.02	1.37	.170
Listening comprehension	0.04	0.12	0.35	.730	0.01	0.10	0.11	.911
Name writing	0.01	0.01	1.05	.295	0.00	0.01	0.21	.833
Letter writing	0.02	0.01	2.12	.035	0.01	0.01	1.96	.050[a]
Invented spelling	0.00	0.01	-0.37	.711	0.00	0.00	0.16	.870

Note. Second Stage Least Squares estimates are results of the instrumental variables analyses. Replication analyses reflect correlational analyses between the number of NBS! lessons received and outcomes.

[a] $p = .0497$.