

# Scaffolding Fidelity and Adaptation in Educational Program Implementation: Experimental Evidence From a Literacy Intervention

David M. Quinn

*University of Southern California*

James S. Kim

*Harvard University*

*In a common approach for scaling up effective educational practice, schools adopt evidence-based programs to be implemented with fidelity. An alternative approach assumes that programs should be adapted to local contexts. In this randomized trial of a reading intervention, we study a scaffolded sequence of implementation in which schools first develop proficiency by implementing the program with fidelity before implementing structured adaptations. We find evidence supporting the scaffolded sequence: A fidelity-focused approach promoted learning and instructional change more so for teachers inexperienced with the intervention, while a structured adaptive approach was more effective for teachers experienced with the intervention. Students benefited more from the structured adaptive approach but only when their teacher had prior experience with the fidelity-focused version.*

**KEYWORDS:** cluster-randomized trial, fidelity of implementation, flexible implementation, scale-up

An important and enduring question in education is that of how research-based instructional practices can be brought to scale (Coburn, 2003; Elmore, 1996). In one common model for scaling up education interventions,

---

DAVID M. QUINN is an assistant professor of education at the Rossier School of Education, University of Southern California, 3470 Trousdale Parkway, Los Angeles, CA 90089; e-mail: [quinnmd@usc.edu](mailto:quinnmd@usc.edu). His research focuses on measuring and explaining inequalities in educational outcomes by race/ethnicity and socioeconomic status; he is particularly interested in the role that teachers can play in closing achievement gaps and improving student outcomes.

JAMES S. KIM is a professor at Harvard University. His research examines the effectiveness of literacy reforms and interventions in improving student outcomes.

schools adopt evidence-based programs, and administrators do their best to ensure that teachers implement those programs with fidelity (the “fidelity-focused” approach). Scholars have argued, however, that consistently achieving positive results across the dynamic and varied settings of schools requires a different model, in which teachers adapt a program’s instructional practices to better fit their context while adhering to core program principles (McDonald, Keesler, Kauffman, & Schneider, 2006). While this approach of “structured adaptive” implementation has the potential to improve the fit between instructional practices and local conditions, adaptation also introduces the possibility that teachers will alter practices in ways that make them less effective (McLaughlin & Mitra, 2002). A third model for managing program implementation combines the fidelity-focused and structured adaptive approaches into a “scaffolded sequence” designed to prepare teachers to make effective rather than counterproductive adaptations (McMaster et al., 2014; Slavin, Madden, & Datnow, 2007). In this scaffolded sequence, teachers first internalize a program’s theory and become proficient with its procedures by implementing the program with fidelity. After mastering the program as designed, teachers transition into an adaptive phase of program implementation during which structures guide their tailoring of the program to their context. However, little empirical work has been done to test the effects of such a scaffolded sequence of program implementation on teacher and student outcomes.

In this study, we investigate the effectiveness of the scaffolded sequence of program implementation in the context of a randomized trial of READS for Summer Learning, a summer literacy intervention for elementary school students that includes school- and home-based components. We randomly assigned 27 schools to one of two conditions: (a) a *core fidelity-focused condition* (Core READS), in which teachers were expected to faithfully implement program procedures used in previous experiments, or (b) a *core treatment plus structured teacher adaptations condition* (Adaptive READS), in which teachers were afforded opportunities to adapt program components so as to better meet the needs of their students (without contradicting core program principles). In other work, we found a positive main effect on student reading comprehension of Adaptive READS compared to Core READS (Kim et al., 2017). This left open the question, however, of what effect a scaffolded sequence of program implementation might have on teachers’ and students’ learning. The present study addresses this question.

At baseline for this study, some teachers had already participated in Core READS due to a within-school random assignment experiment in the previous year in which teachers were assigned to Core READS or a business-as-usual control group. Consequently, a subset of teachers in the present study followed the scaffolded sequence of a fidelity-focused implementation approach in the first year (Core READS), followed by a structured adaptive approach the next year (Adaptive READS). As we describe in more detail

below, this feature allows us to examine the effectiveness of the scaffolded sequence by testing whether the effects of the structured adaptive condition (vs. the fidelity-focused condition) depended on whether the teacher had prior experience with the fidelity-focused version of the program.

We begin by describing the fidelity-focused and structured adaptive approaches and how the approaches differ in their underlying assumptions and the demands they place on implementers. We then discuss theory suggesting that the optimal approach to scaling up educational programs may be the scaffolded sequence of program implementation in which schools first develop proficiency in implementing an evidence-based program with fidelity and then implement structured adaptations to fit their local contexts. After describing how we test the scaffolded hypothesis in this study, we present our results and end by discussing the implications of our findings.

## **Background**

Bringing an educational program to scale involves implementing that program in numerous and varied contexts. While randomized controlled trials have helped the field to identify programs that can work for some students in some circumstances, we have much to learn about the complicated process of scaling up programs and sustaining them over time, a process that involves engaging teachers in deep learning, changing instructional practice, and shifting ownership over educational programs from developers to teachers (Bryk, Gomez, Grunow, & LeMahieu, 2015; Coburn, 2003). The fidelity-focused and structured adaptive approaches to implementation represent two schools of thought for how programs should be scaled up.

### **The Linear Model of Research and Fidelity of Program Implementation**

A common approach for bridging the research-practice gap at scale is that of the “linear model” of scientific research (Coburn & Stein, 2010). As applied to education, the linear model describes a sequence that begins with basic research conducted in the social science disciplines, followed by applied research motivated by problems of educational practice, which eventually leads to the codification of professional knowledge and the dissemination of best practices. Often, these practices take the form of educational programs, or “set[s] of replicable instructional events” (Popham, 1967, p. 402). The programs that are shown through research to have positive effects for students become candidates for scale-up (O’Donnell, 2008). This has been the model championed by the Institute of Education Sciences since 2002 (Coburn & Stein, 2010).

A key concept related to the linear model in education research is that of program implementation fidelity. In the applied research phase, it is only meaningful to measure the effect of an educational treatment if that treatment is well defined and administered faithfully (Dane & Schneider, 1998;

Dusenbury, Brannigan, Falco, & Hansen, 2003; Murnane & Nelson, 2007). When studies indicate that a program has positive effects for students, practitioners are encouraged to implement the program with fidelity because it was under this condition that the positive effects were observed (Dusenbury et al., 2003; LaChausse, Clark, & Chapple, 2014). Bolstering the argument for implementation fidelity is research showing that lower fidelity predicts lower effectiveness and that fidelity is lower in field experiments than in the lab (Hulleman & Cordray, 2009).

Scholars have conceptualized fidelity in various ways (for reviews, see Dane & Schneider, 1998; Mowbray, Holter, Teague, & Bybee, 2003; O'Donnell, 2008), but in its most basic sense, fidelity can be thought of as “the degree to which teachers and other program providers implement programs as intended by the program developers” (Dusenbury et al., 2003, p. 240). Some researchers argue that any change to the original program procedures represents a departure from fidelity regardless of the extent to which the change adheres to foundational program principles (Domitrovich et al., 2009; Sherin & Drake, 2009). However, others accept some program adaptations within the bounds of fidelity as long as the changes are not so drastic as to compromise the program's integrity or effectiveness (Hall & Loucks, 1978 as cited in O'Donnell, 2008). In this study, we adopt the former conception of fidelity (i.e., strict adherence to researcher-designed program procedures) as a contrast to the structured adaptive implementation approach described in the next section.

However important treatment fidelity may be to program evaluation, success stories from the fidelity-focused approach to scale-up are somewhat rare as programs that show promise during initial efficacy trials often (but not always) fail to replicate positive effects in large-scale effectiveness studies (Coalition for Evidence-based Policy, 2013). This may happen for a variety of reasons. First, counterfactuals vary across contexts, and any given intervention may not be superior to all business-as-usual practices. Second, the instructional practices that comprise a program may only be effective for students with certain characteristics or in certain contexts, or perhaps only certain teachers in certain contexts are able to effectively implement these programs (Slavin, 2002).

### **The Structured Adaptive Approach to Program Implementation**

As an alternative to the fidelity-focused approach to instructional scale-up, some scholars argue that educational treatments need not be thought of as recipes to be strictly followed; rather, bringing a program to scale may be more a process of instituting practices across schools that bear a “family resemblance” with one another (Elmore, 1996). In this conception, scaling up educational treatments requires balancing program fidelity with program adaptation (Castro, Barrera, & Martinez, 2004; Dane & Schneider, 1998;

Ferrer-Wreder, Adamson, Kumpfer, & Eichas, 2012; McDonald et al., 2006; McLaughlin, 1990; U.S. Department of Health and Human Services, 2002). According to this line of thinking, programs may have the best chance of improving educational outcomes at scale if the “core components” of the program are kept intact while practitioners adapt the intervention so as to make it more compatible with their context (e.g., more responsive to the particular needs or interests of students or teachers) (Castro et al., 2004; Ferrer-Wreder et al., 2012; McDonald et al., 2006; McLaughlin, 1990). This requires that we consider practitioners’ depth of learning, change in practice, and ownership over the program as important elements of scale-up beyond the count of schools that are “doing” the program (Coburn, 2003).

Effective structured adaptive program implementation likely requires more than simply giving teachers permission to depart from fidelity. Rather, teachers are likely to be better positioned to make effective adaptations if they are supported in doing so through program management structures (Sailors et al., 2014). While research on such structures is rare, one recent article (Lemons, Fuchs, Gilbert, & Fuchs, 2014) reported an evaluation of the effectiveness of structured autonomy compared to a fidelity focus for teachers implementing Peer-Assisted Learning Strategies in Reading (PALS), an evidence-based peer-tutoring program for reading. The structure came through researchers’ classification of program activities as “core,” meaning teachers could not adapt them, and “noncore,” meaning teachers could omit or adapt them. This study showed promising results for the effect of the structured adaptive approach to program implementation; however, teachers self-selected into the fidelity-focused and structured adaptive conditions, precluding causal inference. Additionally, it was not possible to test the scaffolded hypothesis, that structured adaptive management is more effective when preceded by fidelity-focused management.

### **Fidelity Versus Structured Adaptation: Contrasting Assumptions and Demands**

The fidelity-focused and the structured adaptive approaches to program implementation make different assumptions and different demands on the practitioners implementing any given program. In the fidelity-focused approach, the teacher’s job is to achieve the program ideal envisioned by the program developers, in other words, to follow the instructional recipe as closely as possible. Teachers are not expected to diagnose, prescribe, design, or innovate. This approach lends itself well to what has been called the “control” approach to instructional management (Rowan, 1990), in which the teacher’s instructional tasks are predetermined and well defined and the school administrator’s role is to ensure that the teachers execute those tasks. Given that job tasks are prescribed, there is little need for teachers to share innovations with one another or engage in group problem

solving or decision making. Consequently, this approach typically lacks work structures or learning experiences that are collaborative in nature. Under a fidelity-focused approach, the types of learning experiences that support teachers' implementation more often follow the traditional workshop model of professional development, in which the goal is for teachers to learn program-specific procedures and the importance of implementing them faithfully (LaChausse et al., 2014). The approach assumes that the linear model of scientific research will optimize student learning, that is, that research will uncover what works in education and if we can just get teachers to follow these effective practices, students will learn more.

In contrast, teachers under the structured adaptive approach to program implementation must be able to recognize what is working and what is not working about an intervention. When a program isn't working, teachers must diagnose the problem by integrating their general instructional knowledge with specific knowledge about their context and their students, then devise solutions, test those solutions, and repeat the process as needed. As such, the adaptive approach assumes that teachers have the knowledge and skill necessary to plan and execute program adaptations that will improve intervention outcomes, as opposed to adapting the program in ways that will render it less effective. This requires more from the teacher in terms of critical thinking and knowledge about the instructional theory of the program. Without deep knowledge of the program, teachers cannot make adaptations that are consistent with the program theory, so they are unlikely to improve the program (Penuel, Gallagher, & Moorthy, 2011). In the worst case, teachers' adaptations may even be harmful (McLaughlin & Mitra, 2002).

The adaptive approach also calls for more teacher autonomy over instructional practice (as opposed to administrator control) and for collaborative work structures, given the increased potential payoff of exchanging information about innovations and engaging in group problem-solving (Elmore, 1996). As such, learning experiences built around peer collaboration and experimentation may be more useful for teachers under the structured adaptive approach compared to the fidelity-focused approach. One implication of the adaptive model including collaborative work structures is that the experiences of individual teachers will depend on the colleagues with whom they are collaborating. Teachers may learn more and improve their implementation more when they are collaborating with expert colleagues (Frank, Zhao, Penuel, Ellefson, & Porter, 2011).

### **The Scaffolded Sequence of Program Implementation**

The distinct demands made of teachers by the fidelity-focused and structured adaptive approaches to program implementation, and the contrasting work structures that support teachers in meeting those demands, suggest that these approaches may be better-suited for different sets of

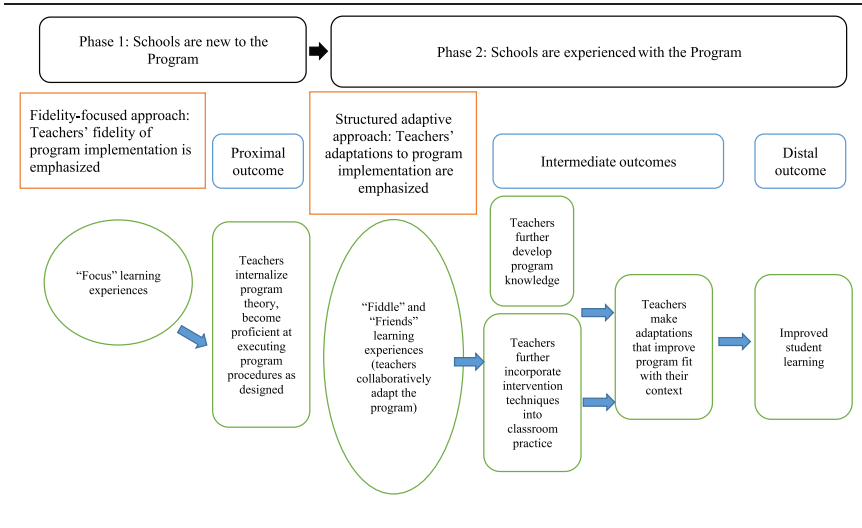
circumstances, where their respective assumptions are met (Berman, 1980). Teachers with less curricular knowledge may perform better under a stricter fidelity approach, while more experienced or effective teachers may be capable of improving the program under the freedom offered by an adaptive approach. Furthermore, skilled teachers who are used to making instructional decisions can sometimes feel frustrated by rigid implementation requirements, making the fidelity-focused approach a poor fit (Meyer, Miller, & Herman, 1993; Murnane & Nelson, 2007; Rowan, 1990).

A Vygotskian perspective (Vygotsky, 1978) suggests that organizing these approaches to program implementation into a scaffolded sequence may optimize structured results. That is, rather than choosing between a fidelity-focused versus adaptive approach when adopting a new program, a school might begin with a fidelity-focused phase of program implementation, followed by a structured adaptive phase. In other words, the scaffolded sequence may be thought of as a structure for chronologically organizing the fidelity-focused and adaptive approaches. The logic behind this scaffolding is that implementing a program with fidelity enables teachers new to the program to understand how the various components work together as designed and to execute them proficiently. After mastering the program as designed through this period of fidelity, teachers are prepared to make structured adaptations that may enhance program effectiveness because they possess the foundational knowledge and skills necessary for designing and executing effective adaptations (McMaster et al., 2014; Slavin et al., 2007). While the theory behind this scaffolded sequence is sound, no experimental studies have directly tested this approach.

### *Scaffolded Learning for Teachers*

Frank and colleagues (2011) devised the mnemonic *focus, fiddle, friends* to describe a learning process for teachers that aligns well with the scaffolded sequence outlined above. According to Frank et al.'s (2011) progression, teachers who are unfamiliar with a particular educational approach first learn about the approach through *focused* professional development and direct instruction. This enables teachers to develop an understanding of an educational program and its theory of action and achieve a basic level of implementation. Teachers then *fiddle* with these techniques by experimenting with variations on the techniques to determine what works best for them and their students. Through this process, teachers develop specific and high-level questions, requiring them to receive expert assistance from experienced *friends* (colleagues). Such discussions with colleagues also help to spread useful innovations and keep the program alive.

Frank et al. (2011) tested this conceptual model using data from schools in which teachers were being encouraged to incorporate computer technology into their classroom instruction. Consistent with the model, the



**Figure 1. Conceptual model for the scaffolded sequence of program implementation.**

researchers found that teachers who were initially infrequently using computer technology experienced greater implementation gains when they received focused professional development on technology use. Teachers who were initially at medium levels of implementation benefitted most from having opportunities to experiment with the technology, while teachers initially at a high level of implementation benefitted from interacting with colleagues about computers (though the highest-level implementers experienced implementation gains from all types of learning experiences). The authors suspected that fiddle experiences may be more effective when they followed focus learning experiences but were unable to empirically test this.

*Differentiated Learning and The Scaffolded Sequence*

The results of the Frank et al. (2011) study suggest that teachers may learn the most, and improve implementation the most, when approaches to program implementation are differentiated according to teachers' intervention-related experience or knowledge. Such differentiation fits well with the model of fidelity-focused and structured adaptive program implementation as a scaffolded sequence. In Figure 1, we merge these ideas into a graphical conceptual model. We test this approach in the present study through the methods described later.

As seen under the Phase 1 heading of Figure 1, when schools initially adopt an intervention, teachers experience a period of fidelity-focused



program implementation in which their primary intervention-related learning comes through explicit instruction on how to implement the program faithfully. These *focused* learning experiences enable new teachers to internalize the program theory and achieve basic proficiency with the instructional procedures as designed by the program developers (the Phase 1 proximal outcome).

According to this conceptual framework, after teachers develop sufficient mastery of the program, they are ready for the second phase, which emphasizes teachers' structured adaptations to program implementation. In this phase, teachers collaboratively (i.e., with friends) design adaptations (fiddle) that they believe will make the program more effective for their students. Through these fiddle and friends learning experiences, teachers acquire a deeper understanding of the program and how its active ingredients interact with their context, which enables teachers to more successfully incorporate the program techniques into their classroom practice (intermediate outcomes).

During Phase 2, one source of teacher learning may be peer effects facilitated through the collaborative implementation structures. That is, individual teachers may have greater success with the program when their collaborators are more expert with the program (Jackson & Bruegmann, 2009) because more expert teachers are more likely to accurately diagnose program difficulties, design effective solutions they can share with colleagues, and provide colleagues with high-quality feedback.

Finally, teachers' experimentation with program adaptations leads to an enacted program that fits better with their context. The improved tailoring of the program to the context then leads to the distal outcome of improved student learning.

### **The Present Study and Research Questions**

In this study, we test the scaffolded sequence of program implementation by analyzing data from a randomized trial of READS for Summer Learning, an evidence-based literacy program for elementary school students. In the year prior to the present study, teachers were randomly assigned within schools to either implement a fidelity-focused version of READS or a business-as-usual control condition. In the next year—the year we focus on in this article—we employed school-level random assignment with a subset of these schools. Specifically, we randomly assigned schools to either implement READS under the same fidelity-focused approach as year 1 (Core READS) or implement the program under a structured adaptive approach (Adaptive READS). In other work, we have shown Core READS to be more effective, compared to business as usual, at improving students' reading comprehension (Kim et al., 2016). We have also shown a positive main effect of Adaptive versus Core READS on student outcomes (and some teacher outcomes not examined here; Kim et al., 2017). The

present study makes a unique contribution to the literature beyond these prior studies by offering an empirical test of the scaffolded sequence of program implementation.

Our primary objective in the present study is to test the hypothesis that a scaffolded sequence of program implementation optimizes teacher learning, changes in practice, and ultimately, student learning. We do this by exploiting baseline variation in teachers' prior participation in Core READS. Specifically, we test the scaffolded hypothesis by examining whether the effects of the adaptive condition (vs. the core, or fidelity-focused, condition) differ depending on whether the teacher had prior experience with Core READS (i.e., we compare the Adaptive effect for teachers with prior Core experience to the Adaptive effect for teachers without prior Core experience). According to the scaffolded hypothesis (as outlined previously), the Adaptive condition will be more effective for teachers who had prior experience implementing the intervention with fidelity, while the Core (fidelity-focused) condition will be more effective for teachers who did not have prior intervention experience. To be clear, we did not begin this study by randomly assigning a set of teachers to a "scaffolded" experimental condition; rather, in the present study, we have a fidelity-focused condition (Core READS) and a structured adaptive condition (Adaptive READS), and we test the scaffolded hypothesis by testing whether an interaction exists between condition and teacher prior experience with (Core) READS. Additionally, as an exploratory question, we seek to understand whether, as suggested previously, peer effects facilitated teacher learning under the collaborative program implementation activities of the Adaptive approach. Understanding how contrasting program implementation structures may interact with characteristics of teachers and schools will help education researchers, decision makers, and practitioners design effective adoption processes for school improvement efforts.

In this study, we ask the following research questions:

*Research Question 1* (RQ1; main effects of implementation approach): In the second year of program implementation, does Adaptive READS (i.e., the structured adaptive approach), as compared to Core READS (i.e., the fidelity-focused approach), affect teachers' intervention-related learning and incorporation of intervention techniques into regular classroom practice?

*Research Question 2* (RQ2; scaffolded hypothesis/interaction effects): In the second year of program implementation, do the effects of Adaptive READS (i.e., the structured adaptive approach) versus Core READS (i.e., the fidelity-focused approach) on teacher outcomes depend on teachers' prior experience with Core READS (RQ2a) or on teachers' peers' prior experience with Core READS (RQ2b)?

*Research Question 3* (RQ3; scaffolded hypothesis/interaction effects): In the second year of program implementation, do the effects of Adaptive READS (i.e., the structured adaptive approach) versus Core READS (i.e., the fidelity-focused approach) on student reading comprehension depend on the teacher's prior experience with Core READS?

## Methods

### Procedures

#### *READS for Summer Learning*

READS for Summer Learning is a program designed to narrow income-based reading skill gaps among elementary school students. In this study, we compare two versions of READS executed over the 2014–2015 school year and summer of 2015: Core READS and Adaptive READS. Core READS is an evidence-based program (Kim et al., 2016; White, Kim, Kingston, & Foster, 2014) representing a fidelity-focused approach to management in which teachers receive training and resources to support their adherence to researcher-designed program procedures. Adaptive READS takes a structured adaptive approach by having teachers work collaboratively with their grade-level teams, with guidance from researchers, to adapt READS in ways they believe will increase its effectiveness for their students.

*Core READS.* Students in Core READS receive eight books in the mail over summer vacation that are matched to their reading level and interests. Each book includes a trifold (or paper folded into thirds) that leads students through the “READS reading routine.” This routine, which is designed to engage students and scaffold their reading, includes a pre-reading activity, which focuses students’ attention on important text structures, and a post-reading comprehension check. Students are expected to mail back completed trifolds (with postage prepaid).

Core READS teachers attended a two-hour training during which they learned how to implement six scripted lessons at the end of the school year that prepare students for the summer activities. During the training, teachers received an overview of the program procedures and materials, watched video clips of teachers implementing key lesson components, and practiced delivering the lessons. To bridge the home and the school, students and their families are invited to a READS Family Night (RFN) in the spring. At this event, parents learn about READS and the trifolds. Also in the spring, students complete a reading comprehension assessment and reading interest survey; this information is used in an algorithm to match books to students. Prior to summer break, students receive copies of the two books used in the end-of-year lessons. Over the summer, the families of students who do not return trifolds receive phone calls with reminders and inquiries about additional support they may need to complete the trifold activities.

*Adaptive READS.* In keeping with previous research suggesting that teachers must understand a program’s theory in order to make productive adaptations (Penuel et al., 2011), teachers at schools assigned to Adaptive READS attended an orientation session in November 2014 in which they

learned the underlying research-based principles of READS. Teachers received school-specific data from a previous year of (Core) READS implementation (e.g., data on trifold return rates and RFN attendance) and examined these data with their grade-level teams to develop hypotheses about ways the program may be improved in their school. After this initial meeting, teachers could elect to earn district professional development credit by completing six online modules in December designed to teach them more about the research-based principles underlying READS (81% of teachers surveyed in the spring participated in the modules). Teachers then attended two additional formal meetings—one in January and one in February—to finalize a plan, based on the data and the research-based principles, for how they would adapt READS. Each school submitted their adaptation plan, allowing us to code the types of adaptations that teachers made. At a majority of Adaptive schools, teachers modified program components so as to improve student and family engagement with the program. For example, a common adaptation included locally developed plans to improve outreach to parents and increase attendance at the family literacy event. A majority of schools also organized a new fall READS event. Additionally, a majority of Adaptive schools modified READS lessons through extensions and/or substitutions, and a majority of Adaptive teachers made changes to their students' summer book lists. See Kim et al. (2017) for detail on these and other adaptations.

Teachers in Adaptive READS received a \$600 stipend for participating; teachers in Core READS received \$300. See Table 1 for a comparison of Core and Adaptive READS.

### Setting, Design, and Participants

In Figure 2, we present a graphic illustrating this study's random assignment procedures. In the school year prior to the present study (the 2013–2014 school year), teachers were randomly assigned within participating elementary schools to a treatment group that implemented Core READS or to a business-as-usual control condition. From this set of schools, 27 high-poverty schools from seven North Carolina school districts were determined eligible and were successfully recruited to participate in the present study over the next school year (to be eligible, schools needed to be 75% or more free or reduced-price lunch, due to prior research showing that READS was effective with this population). Consenting schools were then matched within district based on school poverty level and performance on the state standardized test (in one case, due to the odd number of schools, a triad was formed). Within each matched pair (or triad), one randomly selected school was assigned to Adaptive READS for the 2014–2015 school year; the other schools were assigned to Core READS.

Given that teachers in the previous year were randomly assigned within schools, some teachers had prior experience with Core READS at baseline of

*Table 1*  
**Operationalization of READS Core Components in Core and Adaptive READS**

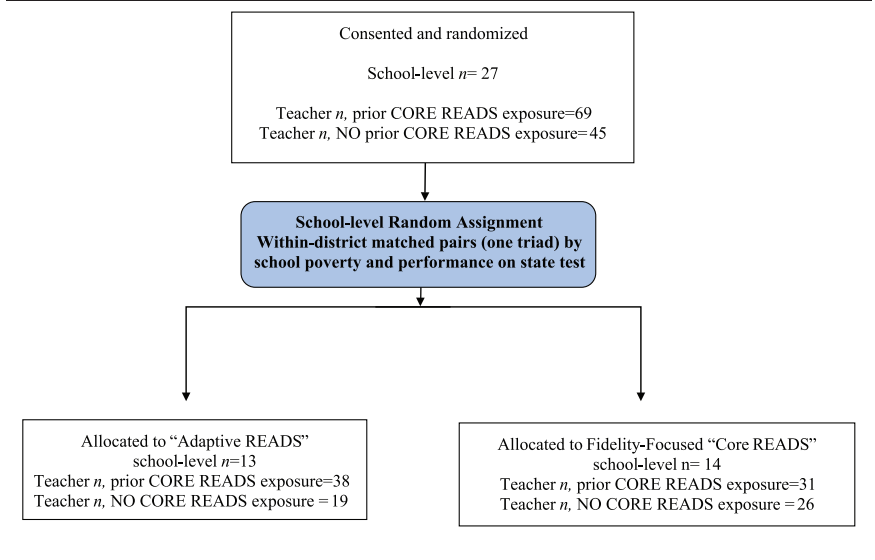
Core Component	Operationalization in Core READS	Operationalization in Adaptive READS	Potential for Acceptable Adaptation
<p>Lessons and summer materials: Students learn a comprehension routine (READS Reading Routine) to use with summer books</p> <p>Family engagement activities: Families learn about READS and how they can encourage their children to participate</p>	<ul style="list-style-type: none"> <li>• Teachers must deliver 6 scripted lessons over 6 consecutive school days</li> <li>• Teachers must distribute fliers to recruit families to the event</li> <li>• Teachers must attend the event</li> <li>• A trained facilitator demonstrates the routine and describes how families can encourage participation over the summer</li> </ul>	<ul style="list-style-type: none"> <li>• Teachers must teach at least 6 READS lessons (scripts are optional)</li> <li>• Teachers must prepare students to use the routine independently over the summer</li> <li>• Each school must host at least 1 family event where parents learn about READS</li> <li>• Teachers must recruit families to the event; they may distribute fliers and/ or use other strategies</li> <li>• Teachers must attend the event; they may take on additional responsibilities at the event(s)</li> </ul>	<ul style="list-style-type: none"> <li>• Make procedural changes to facilitate lesson implementation</li> <li>• Make content changes to address student engagement in or understanding of the routine</li> <li>• Use knowledge of families to increase attendance</li> <li>• Make procedural changes to facilitate implementation</li> <li>• Make content changes to address student and family engagement in or understanding of READS</li> </ul>

*(continued)*

Table 1 (continued)

Core Component	Operationalization in Core READS	Operationalization in Adaptive READS	Potential for Acceptable Adaptation
Summer books: Students receive 10 free books over the summer that are matched to their reading level and interests	<ul style="list-style-type: none"> <li>Teachers play no role in this component</li> <li>Student reading level is measured using Lexile framework</li> <li>Students complete a reading preferences survey</li> <li>A computer algorithm matches books to students (20 books matched; 8 matched books and 2 lesson books sent over the summer)</li> <li>Books are distributed over the summer</li> </ul>	<ul style="list-style-type: none"> <li>Student reading level and preferences are measured as in core, and a computer algorithm matches books to students</li> <li>Teachers may move student Lexile bands up/down by 100 points, generating new book lists</li> <li>Teachers may rearrange ordering of books, making matches more or less likely</li> <li>Teachers may opt to send students additional matched books rather than the 2 lesson books</li> <li>Teachers may opt to have some books delivered at the end of the school year</li> </ul>	<ul style="list-style-type: none"> <li>Address measurement error in reading test and/or survey by using knowledge of students to improve computer book matches</li> <li>Increase the chances that students receive their books over the summer</li> </ul>
Summer nudges: Students and families receive nudges (reminders, incentives) over the summer to encourage participation in READS	<ul style="list-style-type: none"> <li>Teachers play no role in this component</li> <li>Families receive tips and reminders via text or phone calls (sent by research team)</li> <li>Students receive prizes for turning in trifolds (sent by research team)</li> </ul>	<ul style="list-style-type: none"> <li>Families receive tips and reminders via text or phone calls; teachers can personalize the tips</li> <li>Students receive prizes for turning in trifolds</li> <li>Teachers can create additional nudges to remind and/or incentivize students over the summer</li> </ul>	<ul style="list-style-type: none"> <li>Capitalize on personal relationships with students to encourage participation</li> <li>Devise “nudges” that more successfully engage students and families in READS over the summer</li> </ul>

Source. Adapted from Kim et al. (2017).



**Figure 2. School-level randomization to Adaptive versus Core READS.**

the present study, while others did not (in the first year, teachers were not aware of the future possibility of participating in a structured adaptive version of READS). Consequently, the present study includes four types of teachers: (a) teachers who did not participate in Core READS in the past but participated in Core READS for this study; (b) teachers who did not participate in Core READS in past and participated in Adaptive READS for this study; (c) teachers who had previously participated in Core READS, followed by a second year of Core READS in this study; (d) teachers who had previously participated in Core READS, followed by Adaptive READS for this study. This last group participated in the full scaffolded sequence. According to the scaffolded hypothesis described earlier, Adaptive READS will be more effective (compared to Core) among teachers who had previous (Core) READS experience, and Core READS will be more effective (compared to Adaptive) among teachers who had no prior exposure to READS. As described in more detail below, we can therefore test the scaffolded hypothesis by testing whether the effect of Adaptive READS (compared to Core READS) on select outcomes differed depending on whether the teacher had previous experience with Core READS.

In both conditions for this study, all teachers (regardless of prior experience) were required to participate in their condition's learning structures (i.e., all Core teachers participated in lesson training for the present study, regardless of whether they had previously participated in Core READS; all

Adaptive teachers participated in the working group meetings, regardless of whether they had previously participated in Core READS).

## Measures

Teachers completed a web-based survey in the spring of 2015 with questions about their intervention experiences. The survey included original items and items adapted from previously validated surveys. In developing the survey, we went through several rounds of review with external experts and piloted items with teacher consultants (for text from selected survey items, see Appendix A in the online version of the journal).

### *Teachers' Literacy-Related Learning*

We measured teachers' literacy-related learning in areas related to the intervention with an index created by averaging five survey items ( $\alpha = .84$ ) and standardizing those averages to a mean of 0 and standard deviation of 1. Items comprising the index were presented to teachers following the introduction, "In this set of questions, we'd like you to think about your literacy-related learning this school year. This learning could have taken place in any setting." Teachers were then asked several questions with the stem "How much did you learn this school year about each of the following?" The areas teachers were asked about were areas related to the READS components: "matching books to students for independent reading," "teaching students a reading comprehension routine," "engaging students' families in student literacy," "supporting students' independent reading," and "increasing students' engagement in reading." Answer choices were *nothing*, *very little*, *some*, *quite a bit*, and *a tremendous amount*, which were scaled from 0 (*nothing*) to 4 (*a tremendous amount*). A principal components analysis (PCA) revealed only one component with an eigenvalue above 1, which positively weighted all items. For interpretive clarity, we present the results using the mean-based index, but results replicate with a PCA-derived index (see Appendix B in the online version of the journal).

### *READS-Related Literacy Activities in Teachers' Regular Classroom Practice*

We created a scale to measure changes in teachers' literacy practices by averaging teachers' responses on five relevant survey items ( $\alpha = .85$ ) and standardizing the index to a mean of 0 and standard deviation of 1. Teachers were asked a series of questions about the extent to which they incorporated new literacy strategies or followed READS-based principles in their regular classroom instruction (i.e., outside of READS). The areas asked about were the same five areas described previously for teachers' literacy-related learning, and READS was not explicitly referenced as a source of the change in practice. For example, one question asked, "This school



year, to what extent did you incorporate new strategies for supporting students' independent reading into your regular classroom practice (i.e., outside of your planned READS activities)?" Answer choices were *not at all*, *very little*, *some*, *quite a bit*, and *a tremendous amount*, again with scores ranging from 0 (*not at all*) to 4 (*a tremendous amount*). A principal component analysis revealed only one component with an eigenvalue over 1, which positively weighted all items. Again, we present analyses using the mean-based index, but results replicate with the PCA-derived index (see Appendix B in the online version of the journal).

We inquired about teachers' literacy practices outside of READS, as opposed to their practices as part of READS, for two primary reasons. First, our interest is in the broad impact of the scaffolded approach on teachers' instruction. Compared to teachers' behaviors during READS, teachers' incorporation of READS-related principles into their instruction outside of READS is a better indicator of the spread of instructional principles, which is an important dimension of scale-up (Coburn, 2003). Second, because teachers in neither the Core nor Adaptive condition were encouraged to incorporate READS-inspired practices into their regular classroom instruction, teachers' practices outside of READS provide an "apples to apples" comparison, unlike teachers' practices during READS (given that Core teachers were required to implement the program with fidelity while Adaptive teachers were encouraged to adapt).

### *Student Reading Comprehension*

Students took the reading comprehension section of the Iowa Test of Basic Skills (ITBS), Level 10, Form C in fall of 2015 as a posttest. The ITBS is a reliable assessment with reported KR-20 coefficients above .93 (Hoover et al., 2003). We use the ITBS developmental standard score metric, standardized to a sample mean of 0 and standard deviation of 1, to allow the treatment effect coefficient to be interpreted as an effect size.

### **Analytic Plan**

#### *RQ1: Main Effects of Program Implementation Approach on Teacher Outcomes*

To test for Adaptive-Core differences on teachers' literacy-related learning and classroom literacy practices, we fit ordinary least squares (OLS) regression models of the form:

$$Y_{is} = \beta_1 ADAPTIVE_s + \beta_2 EXP_{is} + \sum \beta_k X_s + \varepsilon_{is}, \quad (1)$$

where  $Y_{is}$  is the standardized index score for teacher  $i$  in school  $s$  on either the learning index or the classroom practices index,  $ADAPTIVE$  is a binary indicator variable expressing whether school  $s$  was randomly assigned to

the Adaptive READS condition,  $EXP$  is a binary indicator expressing whether teacher  $i$  had prior experience implementing (Core) READS, and  $x_s$  is a set of dummy variables indicating to which randomization bloc (i.e., within-district pair or triad) school  $s$  belonged. We use cluster-robust standard errors to account for residual dependence within schools.<sup>1</sup>

*RQ2 and RQ3: Interaction Effects Between Teacher Intervention Experience and Program Implementation Approach (Scaffolded Hypothesis)*

*RQ2a and RQ2b: Teacher outcomes.* To test whether the effect of Adaptive READS differed for teachers with and without prior READS experience (RQ2a), we added an interaction term to Model 1:

$$Y_{is} = \beta_1 ADAPTIVE_s + \beta_2 EXP_{is} + \beta_3 (ADAPTIVE_s \times EXP_{is}) + \sum \beta_k x_s + \varepsilon_{is}, \quad (2)$$

where all other terms are as defined previously.<sup>2</sup> To test whether the effect of Adaptive differed depending on the extent to which teachers' peers were experienced with the intervention (RQ2b), we added to Model 1 the interaction between  $ADAPTIVE$  and the number of teacher  $i$ 's READS team colleagues who had prior experience with READS, the main effect of the number of other teachers with READS experience, and the main effect of the total number of teachers on the READS team in school  $s$ .

*RQ3: Student Reading Comprehension*

As noted previously, in other work, we report a significant positive main effect of Adaptive READS (compared to Core READS) on student reading comprehension (Kim et al., 2017). In the present study, our interest is in testing the scaffolded hypothesis by examining whether the effect of Adaptive READS on student reading comprehension differed depending on teachers' prior experience with Core READS. We use OLS regression to model the fall ITBS posttest score of student  $i$  in teacher  $t$ 's classroom in school  $s$  as:

$$ITBS_{its}^{(fall)} = \beta_1 ADAPTIVE_s + \beta_2 EXP_{is} + \beta_3 (ADAPTIVE_s \times EXP_{is}) + \beta_4 PRETEST_{its} + \sum \beta_k x_s + \varepsilon_{its}, \quad (3)$$

where  $PRETEST$  is a pre-randomization pretest covariate (student score on spring 2014 state reading test) included to improve precision, other variables are as defined earlier, and standard errors are clustered by school.

We also present models showing the main effect of Adaptive READS on student reading comprehension overall and the main effect of Adaptive READS separately for the subgroups of students whose teachers were new to READS and those whose teachers had READS experience. Because the

outcome is standardized to a mean 0 and *SD* of 1, the coefficients on *ADAPTIVE* in these models can be interpreted as effect sizes.

## Results

### Descriptive Statistics

In Table 2, we present descriptive statistics by condition for school-, teacher-, and student-level baseline characteristics (top panel) and outcome variables (bottom panel). As seen, random assignment was successful in creating groups of schools that were similar in terms of student percent free or reduced-price lunch and mean reading achievement. Teachers in both conditions were similar in terms of experience in the field of education and master's degree attainment. A slightly higher proportion of teachers in the Adaptive condition had prior experience with READS, and Adaptive READS teachers were slightly more likely to be Black or female relative to Core READS teachers (some demographic information was collected through a fall survey; note incomplete sample sizes). All active teachers at the time of spring survey administration submitted a survey. Among students in the analytic sample, pretest reading scores did not differ by condition. Additionally, among students with pretest scores, attrition did not differ significantly by condition (19.31% for Core READS, 17.32% for Adaptive READS).

### Teacher Outcomes

#### *Literacy-related learning*

*Main effect of program implementation approach (RQ1).* In Table 3, we present models predicting teachers' standardized scores on the literacy learning index. In the first column, we see that condition did not have a significant main effect on teachers' self-reported learning (with a nonsignificant advantage for Adaptive READS of .17 *SD*). Unsurprisingly, teachers across conditions with READS experience reported learning less than teachers who were new to READS (effect size [ES] =  $-.59$  *SD*).

*Interaction between intervention experience and program implementation approach (RQ2a, scaffolded hypothesis).* In the second column, we find that the effect of Adaptive READS differed significantly depending on whether teachers had past experience with READS. While the effect of Adaptive READS was negatively-signed and not significant for teachers new to READS, the effect of Adaptive READS was positive for teachers with past READS experience (ES =  $-.29 + .75 = .46$  *SD*,  $p = .052$ ). These results are consistent with the theory behind the scaffolded model as they show that the management structures were not equally effective for teachers with different levels of experience with the intervention. Specifically, the structures and activities of Adaptive READS—as compared to the structures

**Table 2**  
**Descriptive Statistics by Condition**

	Core READS (Control)		Adaptive READS (Treatment)		Adjusted T-C Difference	p Value		
	Mean	SD	Mean	SD				
<b>Background characteristics</b>								
<b>School-level variables</b>								
School percent free or reduced-price lunch	84.61	10.39	14	85.81	6.5	13	1.38	.54
Average score on fourth-grade state reading test	441.72	3.08	14	442.08	3.18	13	0.12	.87
Percent of fourth graders scoring proficient or above on state reading test	41.84	14.61	14	43.66	13.34	13	0.79	.84
<b>Teacher-level variables</b>								
Number years working in field of education	9.74	7.76	53	10.46	6.71	54	1.19	.14
Number of years teaching in current grade level (Grade 4)	4.41	4.96	54	4.08	4.03	53	-0.3	.52
Number of years working at current school	4.61	5.73	54	4.81	5.42	54	0.26	.65
Worked with READS before this school year? (1 = yes, 0 = no)	0.55		56	0.68		56	0.13	.05
Have or working toward master's degree? (1 = yes, 0 = no)	0.54		54	0.56		54	0.02	.67
Female (1 = yes, 0 = no)	0.88		56	0.96		56	0.08	.05
Black (1 = yes, 0 = no)	0.2		54	0.31		54	0.12	.19
White (1 = yes, 0 = no)	0.67		54	0.57		54	-0.1	.23
<b>Student-level variables</b>								
Student end of grade pretest scores	435.34	10.34	518	435.55	10.1	611	-0.07	.92
<b>Outcomes</b>								
Literacy Learning Index (standardized)	-0.04	1.09	56	0.01	0.96	56	0.09	.65
How much did you learn this school year about: matching books to students for independent reading?	3.3	0.87	56	3.59	0.89	56	0.32	.01
How much did you learn this school year about: teaching students a reading comprehension routine?	3.61	0.97	56	3.59	0.87	56	0.01	.97
How much did you learn this school year about: engaging students' families in student literacy?	3.54	0.76	56	3.5	0.85	56	-0.01	.95
How much did you learn this school year about: supporting students' independent reading?	3.66	0.94	56	3.56	0.83	55	-0.08	.62
How much did you learn this school year about: increasing students' engagement in reading?	3.66	0.86	56	3.7	0.85	56	0.07	.67

(continued)

**Table 2 (continued)**

	Core READS (Control)		Adaptive READS (Treatment)		Adjusted T-C Difference	<i>p</i> Value
	Mean	<i>SD</i>	Mean	<i>SD</i>		
READS-related Literacy Practices Index (standardized)						
Over the past 2 months, to what extent did you guide students in selecting books for independent reading?	-0.06	1.14	0.12	0.91	0.18	.43
This school year, to what extent did you incorporate new strategies for teaching reading comprehension into your regular classroom practice?	3.27	0.94	3.39	0.97	0.16	.41
Over the past 2 months, how much emphasis did you place on engaging students' families in student literacy?	3.48	0.99	3.71	0.82	0.2	.29
This school year, to what extent did you incorporate new strategies for supporting students' independent reading?	3.16	0.91	3.21	0.85	0.1	.49
This school year, to what extent did you incorporate new strategies for engaging students in independent reading?	3.48	1.03	3.62	0.85	0.11	.65
	3.48	1	3.55	0.81	0.08	.68

*Note.* Means and *SD* are unadjusted. *p* value is for test of the null hypotheses that T-C = 0 (standard errors clustered at the school level). Variables listed under indices in "Outcomes" panel represent index components (measured at posttest). Adj. T-C Diff = difference estimated from regression that controls for fixed effects of randomization blocs.

*Table 3*  
**Regression Models Predicting Teachers' Literacy-Related Learning**

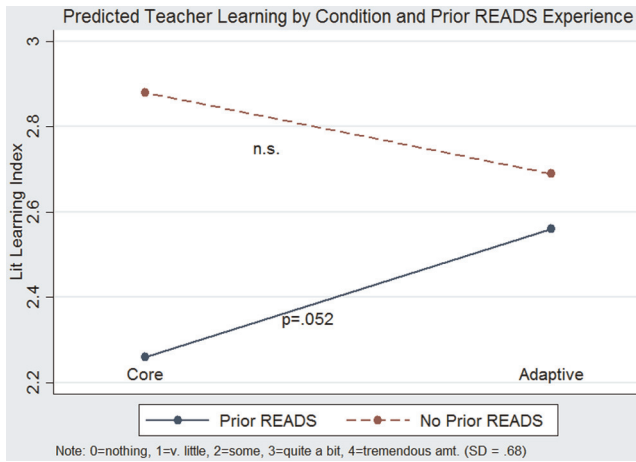
	(1) Literacy Learning Index (Standardized)	(2) Literacy Learning Index (Standardized)	(3) Literacy Learning Index (Standardized)
Adaptive READS	0.171 (0.199)	-0.288 (0.270)	-0.497 (0.479)
READS experience	-0.587*** (0.148)	-0.945*** (0.180)	-0.622** (0.179)
Adaptive READS × READS Experience		0.750* (0.307)	
Adaptive READS × Number of Other Experienced Teachers on Team			0.285 (0.181)
Number of other experienced teachers on team			-0.230 (0.123)
Number of teachers on team			0.189* (0.0756)
<i>N</i>	112	112	112
<i>R</i> <sup>2</sup>	0.250	0.278	0.283

*Note.* Cluster-robust standard errors in parentheses. All models control for fixed effects of randomization blocs. School-level *n* = 27.

\**p* < .05. \*\**p* < .01. \*\*\**p* < .001.

and activities of Core READS—were more beneficial for teachers who had previously participated in Core READS.

We present this interaction graphically in Figure 3. The y-axis represents the original (unstandardized) teacher literacy learning index, with teachers' expected values reported on the original index scale for interpretive purposes. The x-axis contrasts teachers in the Core condition to teachers in the Adaptive condition. The dashed line connects expected values for teachers with no prior READS experience, while the solid line connects expected values for teachers with prior READS experience (expected values derived holding values for bloc dummy variables constant at their means). As can be seen, all subgroups of teachers have predicted values falling between 2 and 3, which represent “some learning” and “quite a bit of learning,” respectively. Examining the dashed line, we see that among teachers with no READS experience, the Core condition promoted more learning (though not by a statistically significant amount). The solid line shows that among



**Figure 3. Interaction between treatment condition and teacher experience predicting intervention-related teacher learning. Difference in slopes is statistically significant at  $p < .05$ .**

teachers with prior READS experience, the Adaptive condition promoted more learning ( $p = .052$ ). Again, the difference-in-differences, which tests the scaffolded hypothesis, is statistically significant.

*Peer effects (RQ2b).* In the third column of Table 3, we test whether Adaptive READS was more effective at promoting learning for teachers in schools in which a greater number of other teachers had experience with READS. The interaction was not statistically significant.

#### *READS-Related Literacy Activities Outside of READS*

*Main effect of program implementation approach (RQ1).* In Table 4, we present models predicting the extent to which teachers incorporated READS-related principles and practices into their regular classroom instruction. As seen in Column 1, the Adaptive condition had no significant main effect on this outcome (with a coefficient of  $.23 SD$ ), and teachers' prior experience with READS also did not significantly predict their READS-related literacy practices outside of READS ( $-.33 SD$ ).

*Interaction between intervention experience and program implementation approach (RQ2a, scaffolded hypothesis).* In Column 2, we again see that the effect of Adaptive READS differed significantly depending on whether the teacher had prior READS experience. The effect of Adaptive was

*Table 4*  
**Regression Models Predicting Teachers' Use of READS-Related Literacy Activities Outside of READS**

	(1)	(2)	(3)
	READS-Related Literacy Practices Index (Standardized)	READS-Related Literacy Practices Index (Standardized)	READS-Related Literacy Practices Index (Standardized)
Adaptive READS	0.227 (0.227)	-0.309 (0.272)	-0.456 (0.663)
READS experience	-0.330 (0.199)	-0.748** (0.260)	-0.346 (0.219)
Adaptive READS × READS Experience		0.875* (0.363)	
Adaptive READS × Number of Other Experienced Teachers on Team			0.319 (0.275)
Number of other experienced teachers on team			-0.206 (0.157)
Number of teachers on team			0.0942 (0.172)
<i>N</i>	112	112	112
<i>R</i> <sup>2</sup>	0.111	0.149	0.131

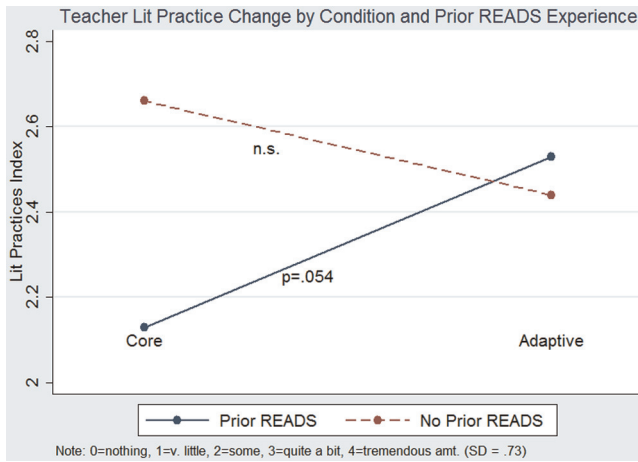
*Note.* Cluster-robust standard errors in parentheses. All models control for fixed effects of randomization blocs. School-level  $n = 27$ .

\* $p < .05$ . \*\* $p < .01$ .

negatively signed and statistically zero for teachers new to READS and was positive and marginally significant for teachers with READS experience ( $ES = -.309 + .875 = .57 SD, p = .054$ ).

In Figure 4, we present this interaction graphically, again using model-predicted values on the original (unstandardized) survey scale (and again, holding values for bloc dummy variables constant at their means). All groups' predicted values fall between 2 and 3 ("some" and "quite a bit" of use of the READS-related practices during regular classroom instruction). From the dashed line, we see that among teachers with no prior READS experience, Core READS teachers used more READS practices (though the difference is not statistically significant). The solid line illustrates that among teachers with READS experience, the Adaptive condition promoted more incorporation of READS practices into regular instruction ( $p = .054$ ). Again,





**Figure 4. Interaction between treatment condition and teacher experience predicting teachers' changes in literacy practices. Difference in slopes is statistically significant at  $p < .05$ .**

the difference-in-differences, which tests the scaffolded hypothesis, is statistically significant.

*Peer effects (RQ2b).* In the third column of Table 4, we find that the effect of Adaptive READS on individuals' literacy practices did not differ depending on the number of a teacher's colleagues who had READS experience.

### Student Reading Comprehension

In Table 5, we present models predicting student fall reading comprehension. As reported elsewhere (Kim et al., 2017), Adaptive READS had a significant main effect on students' fall reading comprehension posttest compared to Core READS (Column 1).<sup>3</sup> The results in Columns 2 through 4 show that this main effect was driven by the subgroup of students whose teachers had READS experience.

### *Interaction Between Teachers' Intervention Experience and Program Implementation Approach (RQ3, Scaffolded Hypothesis)*

In Columns 2 and 3, we present the Adaptive effects for the subgroups of students taught by teachers who were new to READS and those taught by teachers who had previous READS experience, respectively. The effect of Adaptive READS for students taught by inexperienced READS teachers

Table 5

**Regression Models Predicting Student Reading Comprehension Posttest Scores.**

	(1) All Teachers	(2) Teachers New to READS	(3) Teachers With READS Experience	(4) Interaction
Adaptive READS	0.114* (0.0460)	-0.0557 (0.0679)	0.226*** (0.0397)	-0.0425 (0.0700)
Teacher READS experience	0.0465 (0.0532)			-0.105 (0.0683)
Adaptive READS × Teacher READS Experience				0.281*** (0.0718)
<i>N</i>	1,129	470	659	1,129
<i>R</i> <sup>2</sup>	0.534	0.569	0.531	0.538

Note. Cluster-robust standard errors in parentheses. All models control for student pretest and fixed effects of randomization blocs. Teacher-level *n* = 91; school-level *n* = 27.

\**p* < .05. \*\*\**p* < .001.

was not significant (ES = -.06; Column 2), while the effect size for students taught by teachers with READS experience was large and statistically significant (ES = .23). Furthermore, as indicated in Column 4 through the interaction between adaptive and teacher READS experience, the Adaptive effects were statistically different for students whose teachers did and did not have READS experience. These results are consistent with the hypothesis that the benefits of the Adaptive condition for experienced teachers reported earlier translate into more effective instructional experiences for students.

### Discussion

In this study, we tested the hypothesis that a scaffolded sequence of educational program implementation would improve intervention outcomes for teachers and students. In the scaffolded sequence, practitioners new to a program first experience a phase of fidelity-focused implementation, the intention of which is for them to internalize the program theory and procedures by implementing the program as designed by the researchers. After developing proficiency with the program in this way, practitioners transition into a structured adaptive approach to program implementation, during which they make changes that they believe will improve the fit of the program with their context (without compromising core program principles). We found evidence in favor of our hypothesis that a fidelity-focused approach to implementing READS (i.e., Core READS) would lead to more

learning and changes in practice for teachers new to READS, while a structured adaptive approach (i.e., Adaptive READS) would be more effective for teachers who had previously experienced the fidelity-focused version of READS. We also found evidence that these effects for teachers mattered for students, such that the Adaptive condition had positive effects on student reading comprehension only when those students were taught by teachers with prior READS experience. As discussed below, the extent to which these results might generalize to other types of interventions is unclear; however, collectively our findings are consistent with the scaffolded hypothesis.

### **The Scaffolded Sequence**

This study adds experimental evidence to the small set of nonexperimental studies bringing evidence to bear on the effectiveness of a scaffolded sequence of program implementation. Our findings are consistent with Lemons et al. (2014), who found (in a nonexperimental setting) that after implementing a fidelity-focused version PALS, teachers who self-selected into an adaptive PALS condition demonstrated better outcomes compared to teachers who continued in the fidelity-focused condition. Our findings also suggest that the differentiation of teacher learning activities suggested by Frank et al.'s (2011) focus, fiddle, friends framework is well aligned with the stages of the scaffolded sequence. In their observational study, Frank et al. found that teachers who were inexperienced with classroom technology reforms improved their implementation most when their learning experiences centered around information delivery (focus), while teachers who had already reached a basic level of implementation improved most when they experimented with adaptations to the techniques (fiddle) and consulted with colleagues (friends). Accordingly, the professional learning component of Core READS focused on basic information delivery and practice implementing the scripted lessons, while the learning component of Adaptive READS consisted of experimentation and collaboration. Our pattern of results, showing that the effects of these conditions depended on teachers' prior experience with (Core) READS, offers experimental evidence in favor of fusing the focus, fiddle, friends differentiated learning concepts with the scaffolded sequence of program implementation.

Contrary to expectation, we did not find that Adaptive READS teachers learned more or improved their implementation more when more of their grade-level peers had prior experience with Core READS. The interaction coefficients were signed in the expected direction, however, and the null effects could simply be a matter of statistical power. Alternatively, it could be that the extent to which teachers benefit from the collaborative adaptive structures depends more on their own prior knowledge than their peers' knowledge. Relatedly, perhaps Adaptive READS teachers need to have established mastery of the program through prior experience with Core

READS before they are able to benefit from the expertise of their peers. While this three-way interaction (between Adaptive READS, READS experience, and number of other experienced teachers on team) was not significant (models not shown), power to detect this three-way interaction is low, and the direction of the effects was in the expected direction. In short, future research is needed to replicate this null effect and further probe potential explanations.

### *The Value of the Scaffolded Sequence*

Even in cases in which teachers are expected to implement a program with fidelity, program adaptation may be inevitable (Berman & McLaughlin, 1976; Datnow & Castellano, 2000). Consequently, it is important that teachers' adaptations do not compromise the effectiveness of the program. This requires that teachers have deep enough knowledge of the program theory to avoid detrimental adaptations. The scaffolded implementation sequence is a promising way to help teachers develop this knowledge, and the approach to program implementation could be designed with this pedagogical purpose in mind. Teachers' internalization of the program theory is likely to be improved if the fidelity phase is framed as an opportunity for teachers to learn the program before adapting it, as opposed to being framed as the end goal, where teachers' value comes primarily from the fact that they are executing the program designer's vision. In other words, given that teachers' instincts seem to be to adapt programs, the scaffolded approach can be a way of harnessing and focusing that instinct in a way that maximizes the potential for the adaptations to be productive.

Some scholars argue that implementers may be more likely to sustain programs over time when they adapt them to their context (Dearing, 2008). If a particular program has positive effects on student achievement, then sustainability is desirable. Consequently, program adaptation may serve the important goal of sustainability even if the adaptations do not lead to measurable improvements in short-term intervention effects for students. Furthermore, there is potential for adaptation to lead to cascading effects. If teacher involvement in program decision-making leads to higher teacher morale and improved school culture, this may indirectly improve student outcomes (Lee, Dedrick, & Smith, 1991; Lee & Smith, 1996). Of course, adaptations that sustain a program while rendering it ineffective would be counterproductive. Again, this speaks to the value of providing scaffolds that prepare teachers for making effective adaptations.

### **Local Capacity and Will**

Successful implementation of educational policies depends on both the capacity and the will of the implementers (McLaughlin, 1987). The scaffolded sequence discussed here primarily concerns capacity. Although it is

not easy for policymakers to influence local capacity, it may be easier than building will, given that training and consultation can be provided to improve capacity (McLaughlin, 1987). While more explicit research is needed on what, if any, effect the scaffolded sequence may have on teachers' will to implement programs, theory suggests that positive feedback loops affecting teachers' will may arise. To begin with, will and capacity are related because teachers tend to be more willing to implement a program when they believe the program is effective (Kearns et al., 2010); building teachers' capacity to effectively implement a program may therefore also indirectly build their will to implement the program. Second, skillful teachers can be resentful of being asked to follow a program with fidelity, and involving teachers in the decision-making process is one way of earning teacher buy-in (Berman, 1980; Blakely et al., 1987). By framing the fidelity phase as a temporary scaffold that helps teachers acquire the knowledge and skill necessary for teacher-led adaptation, school leaders provide teachers with additional motivation for learning the program at a deeper level and may therefore improve the effectiveness of the scaffolded approach.

## **Limitations and Future Research**

### *Continuous Versus Binary Constructs*

Throughout this article, we have used binary constructs—fidelity-focused versus structured adaptive program implementation approaches and teachers experienced versus inexperienced with a program. Applying simplifying heuristics is useful when developing and testing theory, but in practice, these concepts may exist as continua rather than binaries (Berman, 1980). It is possible that the process of transitioning from a fidelity-focused to structured adaptive approach may be improved with intermediary scaffolds, such that implementers' authority and decision making increase gradually, perhaps beginning with simple decisions before building up to the full structured adaptive approach. Some researchers have suggested that programs should include built-in adaptation suggestions (Webster-Stratton, Reinke, Herman, & Newcomer, 2011), which could serve as adaptation scaffolds. Similarly, suggestions for alternative learning experiences could be built into training materials so as to enable differentiated scaffolding when preparing teacher implementers (Harn, Parisi, & Stoolmiller, 2013). In short, the scaffolded sequence might be scaffolded differently across settings depending on local needs, and future research is needed to experiment with such variations.

While we have focused on teachers' experience with a particular set of instructional procedures, other dimensions of teacher experience or expertise are likely to be relevant to teachers' success at implementing a program under each of the management approaches. Some teachers who are particularly skilled in general teaching practice may require less time in the fidelity-focused phase or may indeed be capable of bypassing that phase

*Quinn, Kim*

to immediately begin making program adaptations. Other teachers may need a longer fidelity period to gain the skill necessary for effective adaptations. Relatedly, the effect of the structured adaptive approach on student outcomes may differ depending on teachers' overall level of effectiveness in addition to their experience with the particular intervention. We leave this hypothesis to future study.

### *Interventions Beyond READS*

One way in which READS for Summer Learning differs from many other educational interventions is that it consists of both home- and school-based student learning experiences. While teachers play an important role in preparing students for successful program participation over the summer, the key learning experiences for students (i.e., reading the summer books) take place outside of teacher guidance or supervision. It is therefore unclear how the principles underlying the findings in this study might play out in a program that is entirely school-based or primarily teacher-led. On the one hand, we might expect some of the patterns seen in this study to be even more pronounced for interventions in which teachers play a more central role. For example, in a curricular intervention in which student learning relies entirely on teacher-led instruction, effects on student learning may be more sensitive to teachers' learning, instructional practice, and adaptations. On the other hand, adaptations to such a program may need to be more substantial for the adapted intervention to be sufficiently distinct from the intervention as originally designed. Relatedly, teachers may need more program-specific knowledge or general expertise to effectively adapt more complex programs. Finally, the demands made on teachers when collaboratively adapting a year-long curricular intervention may be more taxing compared to what occurred in READS. If teachers oppose frequent collaborative meetings and a more demanding adaptation process, teacher investment in the program may be negatively impacted. Given all of these complexities, future research is needed to build an understanding of how fidelity-focused versus structured adaptive approaches play out with other types of educational interventions across various settings.

### *Methodological Limitations*

One methodological limitation is the self-report nature of the teacher outcomes. This concern is less about whether the effects observed here are trustworthy; randomization reassures us that causal inferences regarding the program effects within each teacher subgroup are warranted, and it seems unlikely that these interaction effects would be seen on teachers' perceptions of their learning and behaviors but not on their actual learning and behaviors. Given that teachers were blind to our hypothesized interaction between experience and condition, it also seems unlikely that social

desirability bias could be driving results. Finally, the student outcomes provide additional support for the conclusion that teachers' actual learning and behaviors were affected. Instead, the limitation here is that we cannot know exactly how these measures of teachers' perceptions might relate to observable behavior change. The precise content of teachers' learning is unknown, and we cannot draw conclusions about whether certain teacher learning is relevant to changing practice and whether certain practices are relevant to improving student outcomes. Relatedly, we do not have measures of teachers' mastery of the program theory, which is hypothesized to enable teachers to make more effective adaptations. To obtain a finer-grained picture of this process, further study will be needed.

## Conclusion

By understanding the circumstances under which a fidelity-focused approach versus a structured adaptive approach to educational program management will generally lead to improved outcomes, practitioners will be better positioned to tailor school improvement efforts to their contexts. The findings in this study provide empirical support for the notion that fidelity-focused and structured adaptive approaches can form an effective scaffolded sequence of program implementation, and point toward new areas of exploration that can inform teacher-implemented instructional programs and educational programs more broadly.

## Notes

This study was made possible by an Investing in Innovation Fund (i3) grant from the U.S. Department of Education (PR/Award No. U396B100195). However, the contents of this article do not represent the policy of the U.S. Department of Education, and the content is solely the responsibility of the authors. David Quinn received support from the Dean's Summer Fellowship at the Harvard Graduate School of Education. We are grateful to Heather Hill, Ebony Bridwell-Mitchell, Andrew Volkert, Margaret Troyer, Celia Gomez, Rebecca Unterman, Ann Mantil, and Beth Schuler for feedback on earlier drafts. Helen Chen Kingston, Mary Burkhauser, and Kirsten Aleman contributed to the design of Adaptive READS. Communities in Schools-North Carolina were crucial to implementation efforts. Any errors or omissions are our own.

<sup>1</sup>We also fit multilevel models with random intercepts for schools as sensitivity analyses; all conclusions are unchanged (see Appendix C in the online version of the journal).

<sup>2</sup>We include only teachers who were in these schools prior to this study's randomization but after the previous year's randomization. For these teachers, prior READS experience is considered a baseline characteristic but not random. Models that include only teachers for whom experience was randomly assigned largely replicate the results presented here, as do models using the analytic sample and fit separately for teachers with and without prior READS experience or models that fully interact experience with all covariates (including bloc).

<sup>3</sup>These models differ slightly from those reported in Kim et al. (2017). In Kim et al. (2017), we fit multilevel models with random effects for schools and use school-mean imputation for pretest controls, leading to different sample sizes across articles. Conclusions from Table 5 in this article replicate when using pretest imputation.

## References

- Berman, P. (1980). Thinking about programmed and adaptive implementation: Matching strategies to situations. In H. M. Ingram & D. E. Mann (Eds.), *Why policies succeed or fail* (pp 205–227). Beverly Hills, CA: Sage Publications.
- Berman, P., & McLaughlin, M. W. (1976). Implementation of educational innovation. *The Educational Forum*, *40*(3), 345–370. doi:10.1080/00131727609336469
- Blakely, C. R., Mayer, J. P., Gottschalk, R. G., Schmitt, N., Davidson, W. S., Roitman, D. B., & Emshoff, J. G. (1987). The fidelity-adaptation debate: Implications for the implementation of public sector social programs. *American Journal of Community Psychology*, *15*(3), 253–268.
- Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). *Learning to improve: How America's schools can get better at getting better*. Cambridge, MA: Harvard Education Press.
- Castro, F. G., Barrera, M., & Martinez, C. R. (2004). The cultural adaptation of prevention interventions: Resolving tensions between fidelity and fit. *Prevention Science*, *5*(1), 41–45.
- Coalition for Evidence-based Policy. (2013). *Randomized controlled trials commissioned by the Institute of Education Sciences since 2002: How many found positive versus weak or no effects*. Retrieved from <http://coalition4evidence.org/wp-content/uploads/2013/06/IES-Commissioned-RCTs-positive-vs-weak-or-null-findings-7-2013.pdf>
- Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher*, *32*(6), 3–12.
- Coburn, C. E., & Stein, M. K. (2010). *Research and practice in education: Building alliances, bridging the divide*. Blue Ridge Summit, PA: Rowman & Littlefield Publishers, Inc.
- Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review*, *18*(1), 23–45.
- Datnow, A., & Castellano, M. (2000). Teachers' responses to Success for All: How beliefs, experiences, and adaptations shape implementation. *American Educational Research Journal*, *37*, 775–799.
- Dearing, J. W. (2008). Evolution of diffusion and dissemination theory. *Journal of Public Health Management and Practice*, *14*(2), 99–108.
- Domitrovich, C. E., Gest, S. D., Jones, D., Gill, S., & DeRousie, R. M. S. (2010). Implementation quality: Lessons learned in the context of the Head Start REDI trial. *Early Childhood Research Quarterly*, *25*(3), 284–298.
- Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Education Research*, *18*(2), 237–256.
- Elmore, R. F. (1996). Getting to scale with good educational practice. *Harvard Educational Review*, *66*(1), 1–26.
- Ferrer-Wreder, L., Adamson, L., Kumpfer, K. L., & Eichas, K. (2012). Advancing intervention science through effectiveness research: A global perspective. *Child Youth Care Forum*, *41*, 109–117. doi:10.1007/s10566-012-9173-y
- Frank, K. A., Zhao, Y., Penuel, W. R., Ellefson, N., & Porter, S. (2011). Focus, fiddle, and friends: Experiences that transform knowledge for the implementation of innovations. *Sociology of Education*, *84*(2), 138–156.
- Harn, B., Parisi, D., & Stoolmiller, M. (2013). Balancing fidelity with flexibility and fit: What do we really know about fidelity of implementation in schools? *Exceptional Children*, *79*(2), 181–193.



- Hoover, H., Dunbar, S., Frisbie, D., Oberley, K., Ordman, V., Naylor, R., & Shannon, G. (2003). *Iowa Test of Basic Skills guide to research and development*. Itasca, IL: The Riverside Publishing Company.
- 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, 88–110. doi:10.1080/19345740802539325
- Jackson, K. C., & Bruegmann, E. (2009). Teaching students and teaching each other: The importance of peer learning for teachers. *American Economic Journal: Applied Economics*, 1(4), 85–108.
- Kearns, D. M., Fuchs, D., McMaster, K. L., Saenz, L., Fuchs, L. S., Yen, L., . . . Smith, T. (2010). Factors contributing to teachers' sustained use of kindergarten peer-assisted learning strategies. *Journal of Research on Educational Effectiveness*, 3, 315–342. doi:10.1080/19345747.2010.491151
- Kim, J. S., Burkhauser, M. A., Quinn, D. M., Guryan, J., Kingston, H. C., & Aleman, A. (2017). Effectiveness of structured teacher adaptations to an evidence-based summer literacy program. *Reading Research Quarterly*. Advance online publication. doi:10.1002/rrq.178
- Kim, J. S., Guryan, J., White, T. G., Quinn, D. M., Capotosto, L., & Kingston, H. C. (2016). Delayed effects of a low-cost and large-scale summer reading intervention on elementary school children's reading comprehension. *Journal of Research on Educational Effectiveness*, 9(Suppl. 1), 1–22. doi:10.1080/19345747.2016.1164780
- LaChausse, R. G., Clark, K. R., & Chapple, S. (2014). Beyond teacher training: The critical role of professional development in maintaining curriculum fidelity. *Journal of Adolescent Health*, 54, S53–S58.
- Lee, V. E., Dedrick, R. F., & Smith, J. B. (1991). The effect of the social organization of schools on teachers' efficacy and satisfaction. *Sociology of Education*, 64, 190–208.
- Lee, V. E., & Smith, J. B. (1996). Collective responsibility for learning and its effects on gains in achievement for early secondary school students. *American Journal of Education*, 104(2), 103–147.
- Lemons, C. J., Fuchs, D., Gilbert, J. K., & Fuchs, L. S. (2014). Evidence-based practices in a changing world reconsidering the counterfactual in education research. *Educational Researcher*, 43, 242–252.
- McDonald, S., Keesler, V. A., Kauffman, N. J., & Schneider, B. (2006). Scaling-up exemplary interventions. *Educational Researcher*, 35(3), 15–24.
- McLaughlin, M. W. (1987). Learning from experience: Lessons from policy implementation. *Educational Evaluation and Policy Analysis*, 9(2), 171–178.
- McLaughlin, M. W. (1990). The RAND change agent study revisited: Macro perspectives and micro realities. *Educational Researcher*, 19(9), 11–16.
- McLaughlin, M. W., & Mitra, D. (2002). Theory-based change and change-based theory: Going deeper, going broader. *Journal of Educational Change*, 2, 301–322.
- McMaster, K. L., Jung, P., Brandes, D., Pinto, V., Fuchs, D., Kearns, D., . . . Yen, L. (2014). Customizing a research-based reading practice: Balancing the importance of implementation fidelity with professional judgment. *The Reading Teacher*, 68(3), 173–183. doi:10.1002/trtr.1301
- Meyer, A., Miller, S., & Herman, M. (1993). Balancing the priorities of evaluation with the priorities of the setting: A focus on positive youth development programs in school settings. *The Journal of Primary Prevention*, 14(2), 95–113.
- Mowbray, C. T., Holter, M. C., Teague, G. B., & Bybee, D. (2003). Fidelity criteria: Development, measurement, and validation. *American Journal of Evaluation*, 24(3), 315–340.

- Murnane, R. J., & Nelson, R. R. (2007). Improving the performance of the education sector: The valuable, challenging, and limited role of random assignment evaluations. *Economics of Innovation & New Technology*, 16(5), 307–322.
- O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in k–12 curriculum intervention research. *Review of Educational Research*, 78(1), 33–84.
- Penuel, W. R., Gallagher, L. P., & Moorthy, S. (2011). Preparing teachers to design sequences of instruction in earth systems science: A comparison of three professional development systems. *American Educational Research Journal*, 48, 996–1025. doi:10.3102/0002831211410864
- Popham, J. W. (1967). Instructional product development: Two approaches to training. *AV Communication Review*, 15, 402–411.
- Rowan, B. (1990). Commitment and control: Alternative strategies for the organizational design of schools. *Review of Research in Education*, 16, 353–389.
- Sailors, M., Hoffman, J. V., Pearson, P. D., McClung, N., Shin, J., Phiri, L. M., & Saka, T. (2014). Supporting change in literacy instruction in malawi. *Reading Research Quarterly*, 49(2), 209–231. doi:10.1002/rrq.70
- Sherin, M. G., & Drake, C. (2009). Curriculum strategy framework: Investigating patterns in teachers' use of a reform-based elementary mathematics curriculum. *Journal of Curriculum Studies*, 41(4), 467–500.
- Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational Researcher*, 31(7), 15–21.
- Slavin, R. E., Madden, N. A., & Datnow, A. (2007). Research in, research out: The role of research in the development and scale-up of Success for All. In D. K. Cohen, S. H. Fuhrman, & F. Mosher (Eds.), *The state of education policy research* (pp. 261–280). Mahwah, NJ: Lawrence Erlbaum Associates.
- U.S. Department of Health and Human Services. (2002). *Finding the balance: Program fidelity and adaptation in substance abuse prevention*. Washington, DC: Author.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Webster-Stratton, C., Reinke, W. M., Herman, K. C., & Newcomer, L. L. (2011). The Incredible Years teacher classroom management training: The methods and principles that support fidelity of training delivery. *School Psychology Review*, 40(4), 509–529.
- White, T. G., Kim, J. S., Kingston, H. C., & Foster, L. (2014). Replicating the effects of a teacher-scaffolded voluntary summer reading program: The role of poverty. *Reading Research Quarterly*, 49(1), 5–30. doi:10.1002/rrq.62

Manuscript received June 3, 2016

Final revision received May 27, 2017

Accepted June 5, 2017