

# The Effects of an Academic Language Program on Student Reading Outcomes

NCEE 2022-0071  
U.S. DEPARTMENT OF EDUCATION

*A Publication of the National Center for Education Evaluation at IES*



**U.S. Department of Education**

Miguel Cardona

*Secretary*

**Institute of Education Sciences**

Mark Schneider

*Director*

**National Center for Education Evaluation and Regional Assistance**

Matthew Soldner

*Commissioner*

Marsha Silverberg

*Associate Commissioner*

Tracy Rimdzius

*Project Officer*

The Institute of Education Sciences (IES) is the independent, nonpartisan statistics, research, and evaluation arm of the U.S. Department of Education. The IES mission is to provide scientific evidence on which to ground education practice and policy and to share this information in formats that are useful and accessible to educators, parents, policymakers, researchers, and the public.

We strive to make our products available in a variety of formats and in language that is appropriate for a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other IES product or report, we would like to hear from you. Please direct your comments to [ncee.feedback@ed.gov](mailto:ncee.feedback@ed.gov).

This report was prepared for the Institute of Education Sciences (IES) under Contract ED-IES-15-C-0050 by MDRC. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

August 2022

This report is in the public domain. Although permission to reprint this publication is not necessary, it should be cited as:

Corrin, W., Zhu, P., Shih, M., Brown, K. T., Teres, J., Darrow, C., Nichols, A., & Lack, K. (2022). *The Effects of an Academic Language Program on Student Reading Outcomes* (NCEE 2022-007). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. <http://ies.ed.gov/ncee>.

This report is available on the Institute of Education Sciences website at <http://ies.ed.gov/ncee>.



# The Effects of an Academic Language Program on Student Reading Outcomes

**August 2022**

**William Corrin**

**Pei Zhu**

**Miki Shih**

**Kevin Thaddeus Brown, Jr.**

**Jed Teres**

MDRC

**Catherine Darrow**

**Austin Nichols**

**Kelly Lack**

Abt Associates

---

Helping English learners and students from economically disadvantaged backgrounds read and perform academically as well as their more advantaged peers remains a struggle for many schools seeking equitable outcomes for their students. Research suggests that an increased focus on learning of academic language—the formal language used to read, write, listen, and speak about the topics studied in school—may be one way to boost student and school success. This study investigated WordGen Elementary, a program designed to improve fourth- and fifth-grade students’ ability to understand and communicate academic language and their general reading skill. The program included a curriculum of supplemental reading, speaking, and writing activities for students to engage in over a school year, as well as ongoing professional development to support teachers’ delivery of the curriculum in their classrooms. About sixty schools from six districts around the country agreed to participate and were assigned at random to implement the program or continue with their typical language instruction programs and practices. The study compared the average reading performance of these two groups of schools to assess the program’s effectiveness.

## Key Findings

- ***The program did not affect the academic language skills or reading performance of students overall.***
- ***Despite the program’s attention to instruction relevant to the specific learning needs of English learners and students from economically disadvantaged backgrounds, the program had no effect on academic language skills or reading performance of either group.***
- ***The training and support provided to teachers during the year did not change most aspects of instruction that were targeted by the program, which might explain the lack of effects on student outcomes.***

---

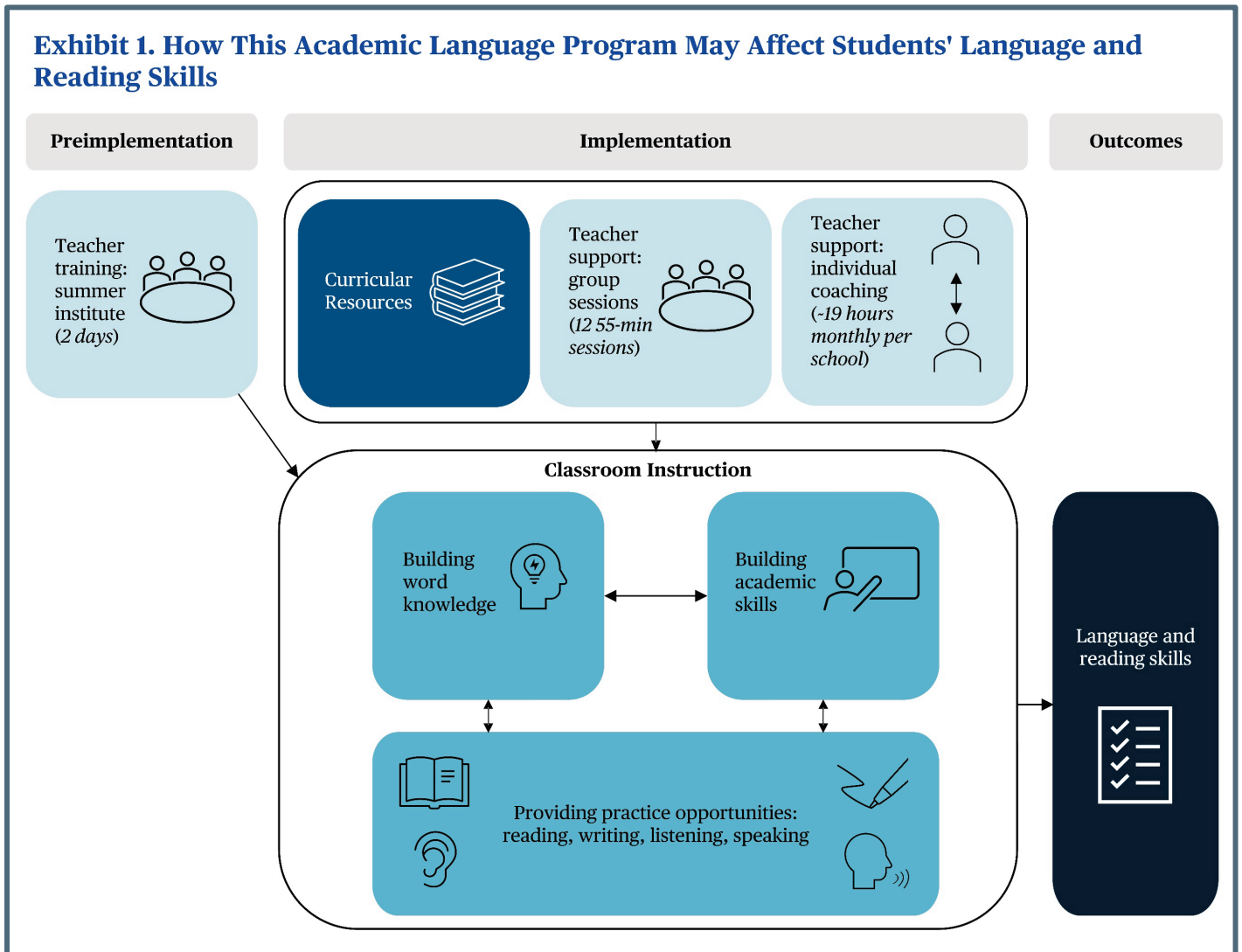
## INTRODUCTION

Significant differences in achievement between English learners, students from economically disadvantaged backgrounds, and their more advantaged peers have persisted for decades. The latest National Assessment of Educational Progress (NAEP) Reading Report Card indicates that only 10 percent of fourth-grade English learners and 21 percent of economically disadvantaged students are proficient readers, compared with 35 percent of students overall.<sup>1</sup> Such differences in reading achievement persist in middle school and are apparent in other academic content areas such as math.<sup>2</sup> While improvement for all students is needed, the gaps are also important to close so as not to leave behind English learners who make up about 10 percent of all students in American public schools and those from low-income households who constitute just over half of the public-school student population.<sup>3</sup>

Lack of familiarity with the formal language of school, or academic language, can be an obstacle to students’ academic success, especially for children learning English or growing up in poverty. These students often have limited exposure to academic language due to, for example, limited access to resources, such as varied reading materials, that support reading development.<sup>4</sup> English learners also face the challenge of developing proficiency in English at the same time as they are trying to master the academic content of subjects like social studies, math, and

science. Therefore, academic language support may help these students strengthen their English skills and better learn academic content as subject area content becomes more demanding during the upper-elementary grades when students are preparing to enter middle school.<sup>5</sup>

Helping students understand and use academic language is the key goal of WordGen Elementary, the program evaluated in this study.<sup>6</sup> At the center of the program is a curriculum comprised of 12 two-week units, with 40-50-minute lessons each school day. Each unit focuses on a topic meant to engage students (such as “*who should decide what we eat?*” and “*why do we fight?*”).<sup>7</sup> As part of the study program, staff provided intensive training and support to introduce participating teachers to the curriculum and help them integrate the daily lessons into their English language arts, math, science, and social studies instruction by supplementing or replacing similar activities that they would typically use.<sup>8</sup> The study also paid for coaches to support professional learning communities, answer teachers’ questions, and model how the curriculum should be used. Successful implementation in the classroom over one year was expected to improve, as a necessary first step, students’ language and reading skills (see Exhibit 1).



While prior studies suggest programs with some similar instructional elements could improve students' academic vocabulary and content area knowledge, not enough is known about whether this kind of program leads to the kinds of literacy outcomes that educational leaders care about most—better reading comprehension and reading achievement.<sup>9</sup> The current study not only examined these outcomes, but also was implemented in a set of selected schools serving large proportions of English learners and students from economically disadvantaged backgrounds. Information on effective ways to improve these outcomes for these types of schools and students is particularly important in efforts to recover from educational disruptions stemming from the Coronavirus pandemic. Box 1 provides an overview of the study design, which was largely implemented during the 2017-2018 school year (hereafter referred to as the program year). More details about how the study was carried out are in Appendix B.

## Box 1. Study Design

### Who participated?

- Fifty-eight elementary schools from six large urban districts in five states and over 3,000 fourth- and fifth-graders in these schools participated.<sup>10</sup>
- Less than 30 percent of the students in the study performed at or above proficiency levels on state English language arts (ELA) and math assessments prior to the program year. Compared with schools nationally, study schools on average served a higher proportion of students eligible for free or reduced-price lunch (82 percent vs. 71 percent) and a higher proportion of English learners (36 percent vs. 17 percent).<sup>11</sup> (See Exhibit B.3 and Exhibits B.12, B.15, and B.18 in Appendix B for details.)

### How was the study conducted?

- Within each district, about half of the schools with similar proportions of English learners were randomly assigned either to participate in the academic language program (hereafter referred to as the “program group”) or not to participate (the “non-program group”) in the fall of 2017.<sup>12</sup>
- The resulting two groups of schools were similar in student composition, characteristics, and student achievement at the start of the study (see Exhibit B.4 in Appendix B, Section I).
- The study estimated the program's effects by the differences in average student outcomes between these two groups of schools.

### What data and measures were used?<sup>13</sup>

- **Student outcomes** include scores on the Core Academic Language Skills Instrument (CALs-I) that measures students' academic language skills, scores on the reading comprehension part of the Gates-MacGinitie reading test (GMRT) that measures students' reading comprehension skill, and performance on the state ELA test that measures students' general reading skills. The study administered the CALs-I and GMRT tests to fourth- and fifth-grade students in the study schools in the spring of the program year and collected their state ELA scores from district records for the program year and the year after the program ended (the follow-up year).<sup>14</sup>

(continued)

## Box 1 (continued)

- **Implementation support measures** include the amount of initial training and ongoing support offered to teachers and their attendance rates in these activities. These measures were constructed using information from periodic teacher training attendance logs, program provider and coach reports, and online teacher surveys administered in the fall/early winter and the spring of the program year.
- **Teachers' classroom practice measures** include an overall composite of instructional practices that were generally important for academic language development as well as three sub-scores (building word knowledge, building academic skills, and providing practice opportunities). Data used to construct these scores were collected through classroom observations of teachers in the program and non-program schools conducted twice during the program year.<sup>15</sup>

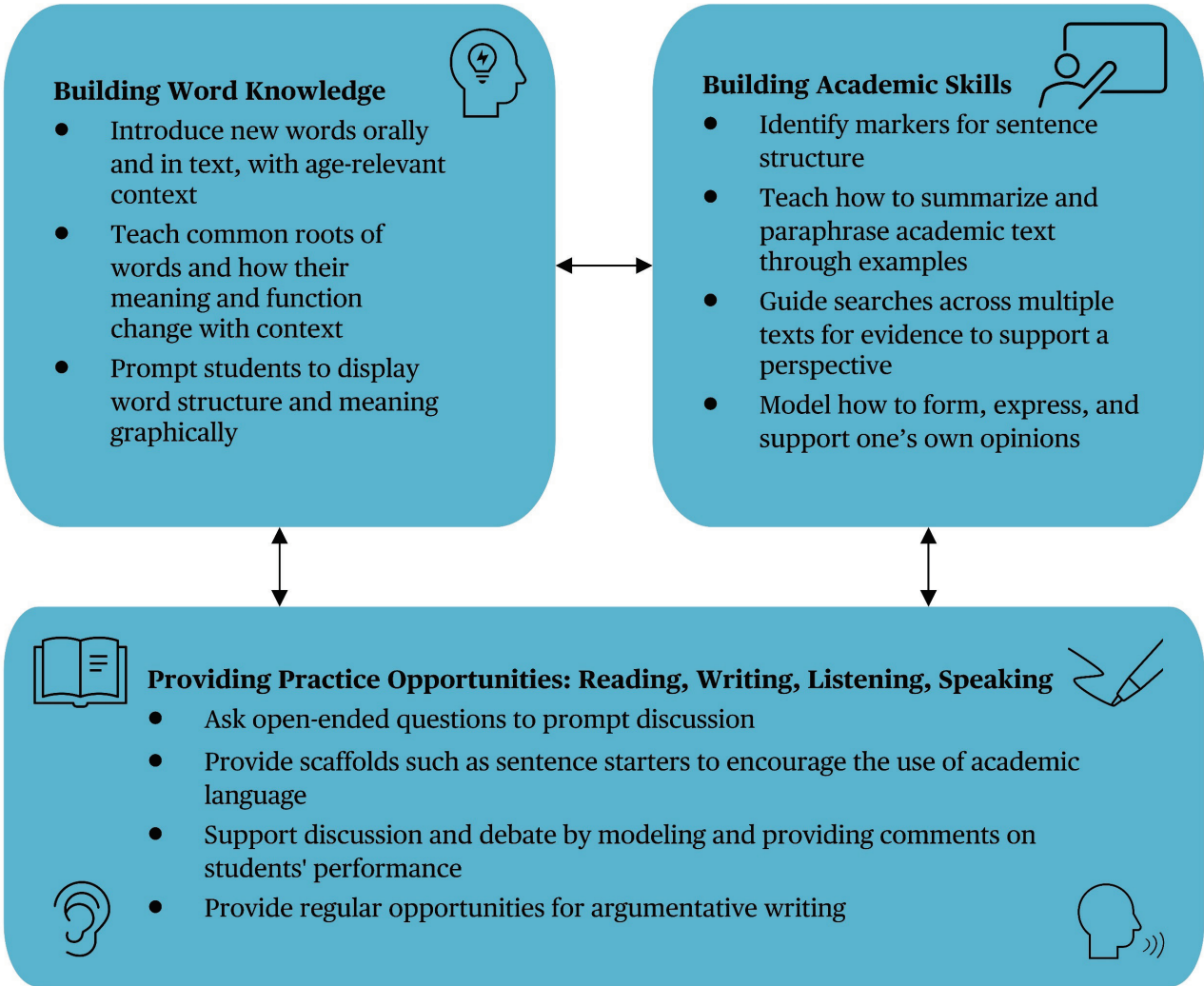
## NO EFFECT OF THE PROGRAM ON STUDENTS' LANGUAGE AND READING PERFORMANCE OVERALL

The curricular units in this program were designed to build all students' academic word knowledge and academic skills through engagement with a variety of texts, word-learning activities, writing tasks, and opportunities for discussion and debate. To effectively deliver these units in the classroom, teachers were expected to use instructional practices that directly target the three core instructional components of the program that focused on building word knowledge, building academic skills, and providing opportunities to practice such knowledge and skills. Exhibit 2 provides examples of such practices for each component.

The program provided specific guidance to facilitate teachers' integration of these practices into the curricular units. For example, to teach the word "*isolated*," the program suggested that teachers first define the word with age-appropriate language ("*having little contact or being separated from others*"). Teachers then prompt students to think of a time when they felt isolated from friends or family to help them understand the meaning through personal connections. From a set of pictures, students then pick the one that depicts someone who is isolated. Other forms of the word ("*isolation*" and "*isolate*") that can be found in the unit would also be introduced to help students' understanding of the text. These activities were expected to help students become independent word learners and support their comprehension and use of the focus words across different contexts.<sup>16</sup>

The program also provided tips on teaching academic skills and suggested various activities to encourage student practice. For instance, teachers could have students summarize perspectives presented in reading materials in writing on flashcards and then discuss with a partner to identify evidence to support their summarization. These skills are essential for students' understanding of academic content and are especially important for fourth- and fifth-graders transitioning from learning to read to reading to learn.<sup>17</sup> In addition, the program encouraged the use of targeted words through teacher activities such as tallying the correct use of the newly learned words in writing or speaking throughout the classroom and rewarding students for high usage. Delivery of these practices was expected to improve students' academic language, reading comprehension, and eventually lead to growth in general reading skills for all students.<sup>18</sup>

# Exhibit 2. Instructional Practice Examples for the Three Core Components

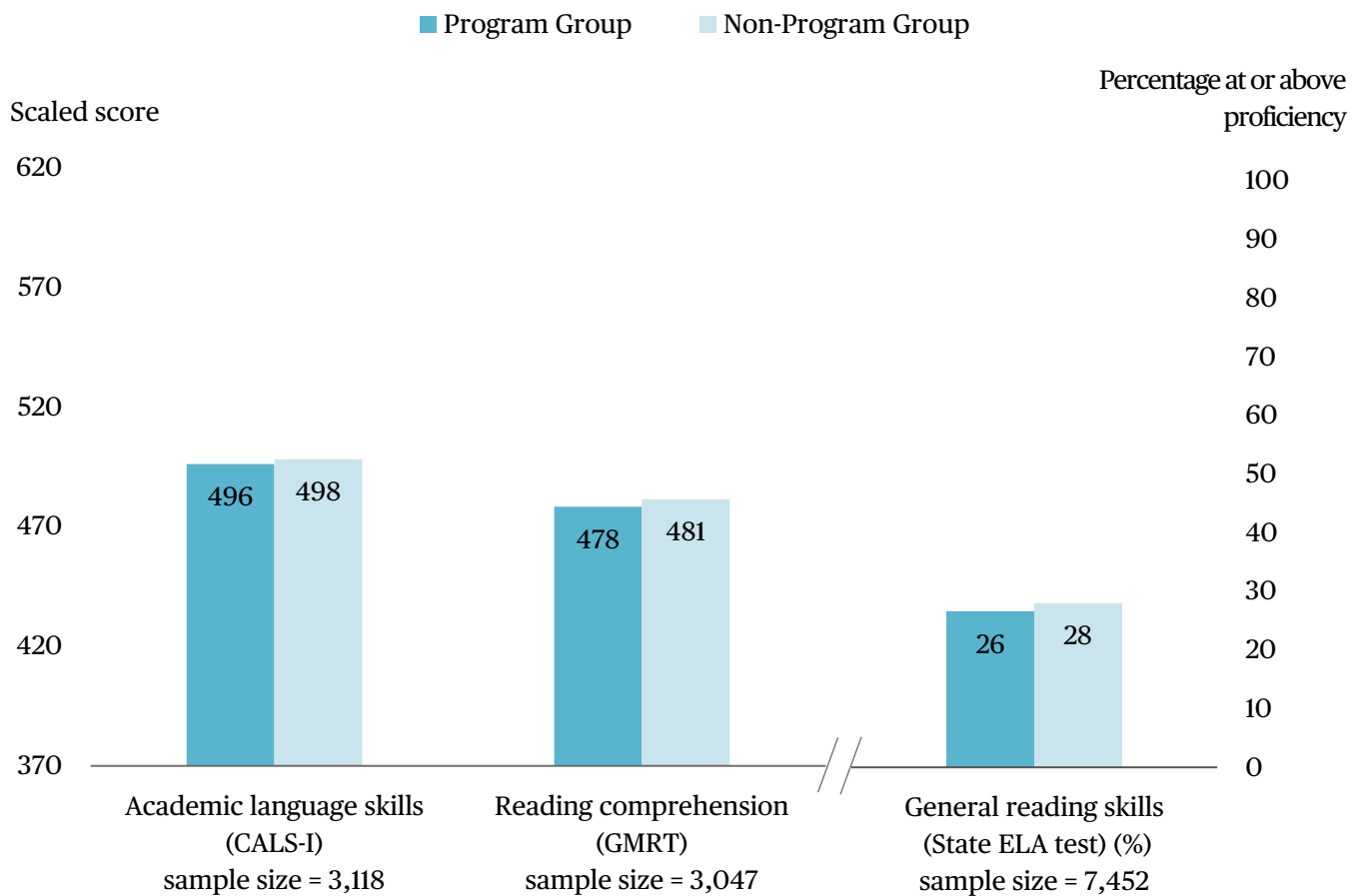


- ***The program did not affect students' academic language skills during the program year.***

At the end of the program year, students' academic language skills (CALIS-I) scores were similar in schools in the program and non-program groups (see Exhibit 3). The average score on the academic language assessment is equivalent to the sixty-first percentile for the program group and the sixty-sixth percentile for the non-program group.<sup>19</sup> This assessment is a broader measure of academic language proficiency than those used in prior studies of academic language programs, which typically focused on measuring the learning of specific vocabulary taught by the individual programs studied.<sup>20</sup> Because the broader set of skills is important to school leaders who might adopt an academic language improvement program, and to reduce the burden on participating schools and students, the study did not administer a test of program-specific vocabulary. Thus, it is not possible to see if prior findings of the positive effects of similar programs on vocabulary are replicated in these schools.



### Exhibit 3. No Effect on Students' Reading and Language Test Performance



SOURCES: Scores from the Core Academic Language Skills Instrument (CALS-I) test and the reading comprehension part of the Gates-MacGinitie reading test (GMRT) administered in the spring of 2018; district records data for student background information and state standardized reading test scores from the 2016-2017 and 2017-2018 school year.

NOTES: 58 study schools participated in the CALS-I test, while 57 schools participated in the GMRT test. The analyses for CALS-I and GMRT scores include all students who consented and responded to these tests. The analyses for the state reading test include all students with valid state test scores in the program year.

ELA = English Language Arts.

None of the estimated impacts are statistically significant at the 0.05 level. The estimated effect on CALS-I translates into an effect size of -0.06 standard deviations; the estimated effect on GMRT translates into an effect size of -0.08 standard deviations; the estimated effect on the percentage of students who scored at or above state proficiency level translates into an effect size of -0.03 standard deviations.

- **The program also did not affect students' reading comprehension and reading achievement.**

Perhaps because the program did not affect students' academic language skills, both the reading comprehension and standardized ELA test scores were similar between the program and non-program schools. At the end of the program year, students in the study sample were about one grade level behind in their reading comprehension skills.<sup>21</sup> Similarly low proportions of students in the program and non-program

schools scored at or above the state-defined reading proficiency level in the spring of the program year (see Exhibit 3).<sup>22</sup> The program also did not have any delayed effects on students' reading skills, as there was no difference in average performance on state ELA tests between the program and non-program schools one year after the training and support of the program ended.<sup>23</sup>

## **ALSO NO EFFECT ON LANGUAGE AND READING PERFORMANCE SPECIFICALLY FOR ENGLISH LEARNERS OR ECONOMICALLY DISADVANTAGED STUDENTS**

Though the program was designed for all students in general education classes, it included curricular units on discussion topics that might be directly relevant to English learners. For example, English learners might have personal knowledge and strong opinions about topics such as “*Should everyone learn a second language?*” and “*Why do communities have different ideas about what brings happiness?*” These topics could help better engage English learners in classroom activities and motivate them to use their language skills.

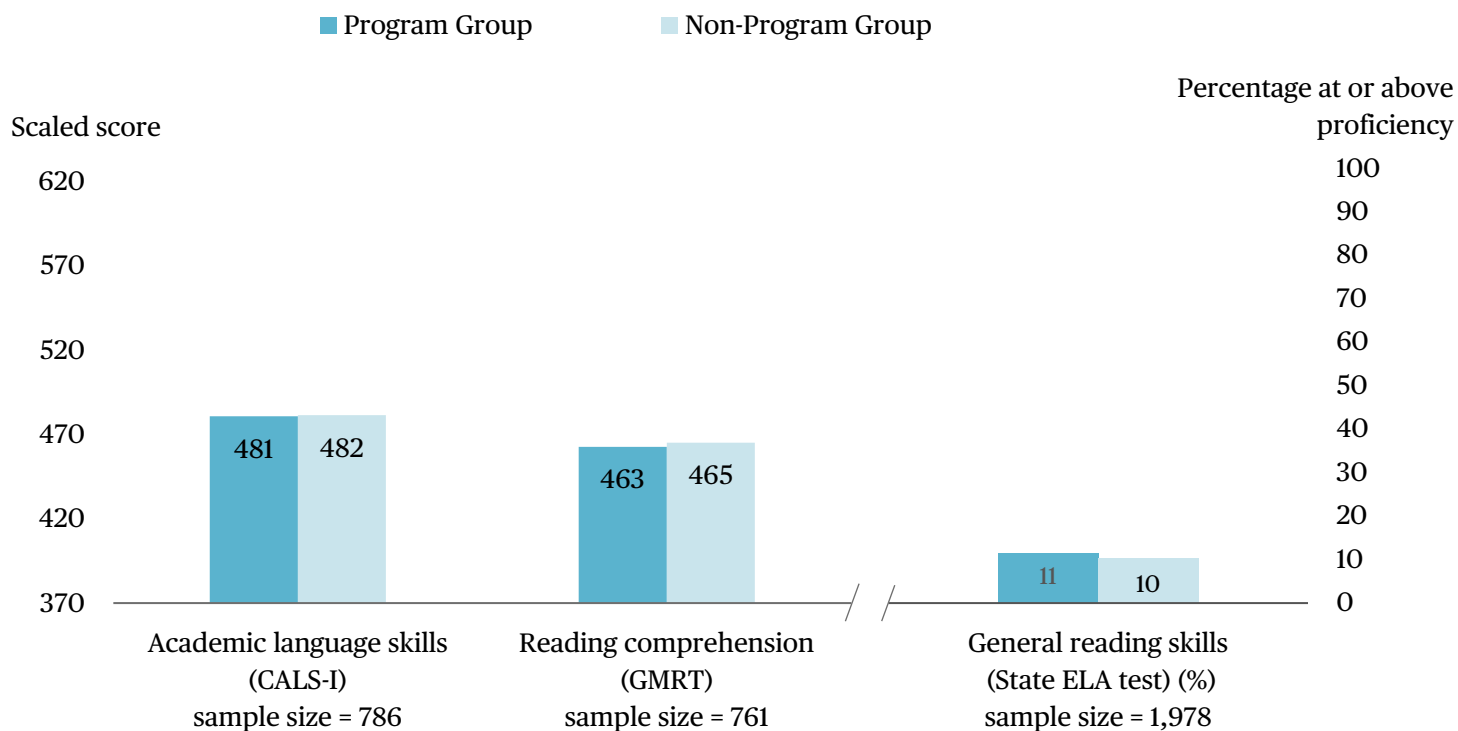
The program also provided additional resources and teaching tips to support English learners in the classroom.<sup>24</sup> In terms of instructional practices, the program featured practices that had been shown to be beneficial for English learners, such as teaching a set of academic vocabulary words intensively across several days, integrating oral and written language instruction into content-area teaching, and providing regular, structured writing opportunities.<sup>25</sup> The program also encouraged teachers to call out close cognates to other languages (for example, “*required*” to “*requerido*” in Spanish) for the benefit of English learners who were native speakers of those or related languages. In addition, the program provided guidance to teachers on how to scaffold discussions and writing exercises through modeling, reviewing strategies, and other collaborative learning approaches. For instance, the curriculum material often prompted teachers to provide sentence frames (“*\_\_ thinks everyone should be included in the community because \_\_,*” “*I agree with \_\_ because \_\_*”) during discussion or writing exercises. Such support could boost English learners' confidence in using their newly acquired academic language skills. All these features represented support above and beyond what English learners might have gotten without the program.

Instructional practices considered beneficial for English learners might also be effective for students from economically disadvantaged backgrounds. Prior research shows that English learners and students from low socioeconomic backgrounds are both at risk for developing academic difficulties in late elementary and middle school.<sup>26</sup> Low English literacy at home and lack of access to resources that support academic language development are present for both student groups. In addition, English learners tend to come from economically disadvantaged backgrounds disproportionately. Therefore, instruction focusing on academic language may be beneficial for both groups.

- ***But the program had no effects on the academic language skills of English learners or students from economically disadvantaged backgrounds.***

English learners and students from economically disadvantaged backgrounds in the program schools performed similarly on the assessment of their academic language skills as their counterparts in the non-program schools. On average, the English learners' scaled score was within a point of 482 on the test, and the economically disadvantaged students' scaled score was within a point of 497, regardless of whether they were in the program schools or the non-program schools (see Exhibits 4 and 5).<sup>27</sup>

### Exhibit 4. No Effect on English Learners' Reading and Language Test



SOURCES: Scores from the Core Academic Language Skills Instrument (CALS-I) test and the reading comprehension part of the Gates-MacGinitie reading test (GMRT) administered in the spring of 2018; district records data for student background information and state standardized reading test scores from the 2016-2017 and 2017-2018 school year.

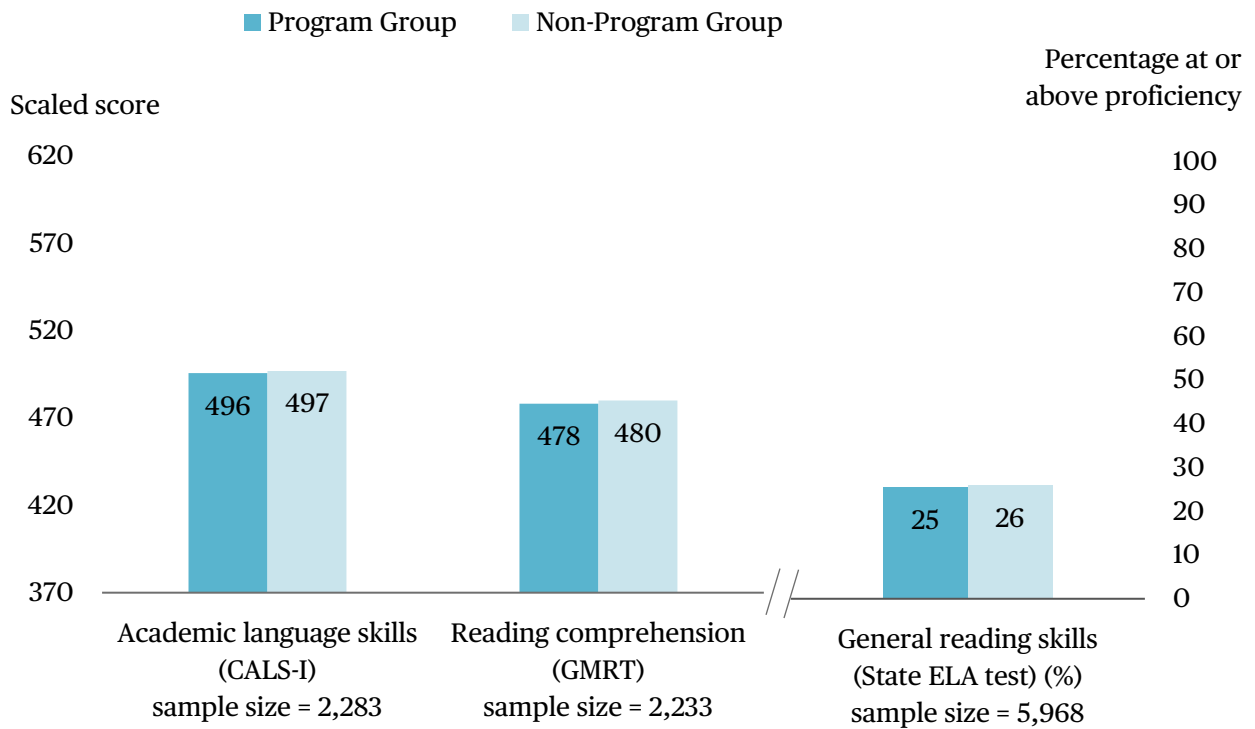
NOTES: 58 study schools participated in the CALS-I test, while 57 schools participated in the GMRT. The analyses for CALS-I and GMRT scores include all English learners who consented and responded to these tests. The analyses for the state English Language Arts (ELA) test include all English learners with valid state test scores in the program year.

None of the differences are statistically significant at the 0.05 level. The estimated effects for English learners translate into effect sizes of -0.02, -0.06, and 0.02 for the CALS-I, GMRT, and state ELA test, respectively.

- ***The program also had no effects on reading comprehension or reading achievement for either of these two groups of students.***

At the end of the program year, the reading comprehension and general reading performance of English learners and students from economically disadvantaged backgrounds was similar across the program and non-program groups (see Exhibits 4 and 5). As was true for students overall, the reading achievement test scores were similar for these two important groups of students in the follow-up year.<sup>28</sup>

## Exhibit 5. No Effect on Reading and Language Test Performance for Students from Disadvantaged Backgrounds



SOURCES: Scores from the Core Academic Language Skills Instrument (CALS-I) test and the reading comprehension part of the Gates-MacGinitie reading test (GMRT) administered in the spring of 2018; district records data for student background information and state standardized reading test scores from the 2016-2017 and 2017-2018 school year.

NOTES: 58 study schools participated in the CALS-I test, while 57 schools participated in the GMRT test. The analyses for CALS-I and GMRT scores include all students from disadvantaged backgrounds who consented and responded to these tests. The analyses for the state English Language Arts (ELA) test include all students from disadvantaged backgrounds with valid state test scores in the program year.

None of the differences are statistically significant at the 0.05 level. The estimated effects for students from disadvantaged backgrounds translate into effect sizes of -0.04, -0.05, and -0.01 for the CALS-I, GMRT, and state ELA test, respectively.

## CHALLENGES IN DELIVERING TRAINING AND SUPPORT TO TEACHERS ON PROGRAM CONTENT AND INSTRUCTIONAL PRACTICES LIMITED CHANGES IN CLASSROOM INSTRUCTION EXPECTED TO AFFECT STUDENT OUTCOMES

The program provider planned intensive training and sustained support activities during the program year to support teachers' implementation of the curriculum in classrooms.<sup>29</sup> The initial training conducted by the program provider was intended to orient teachers and study-hired local coaches to the program.<sup>30</sup> Coaches and

teachers were expected to attend monthly webinars, each with a follow-up session to reflect on how the training could support the implementation of specific lessons. The coaches were expected to help teachers integrate the program into their daily instruction and support students’ practice in using academic language through classroom discussion and debates. Because these opportunities for student practice were central to the program’s success, but can be challenging to do well, coaches were expected to observe and provide teachers with feedback, as needed, throughout the year (see Exhibit 6).<sup>31</sup>

### Exhibit 6. Initial Training and Ongoing Support Plan for Coaches and Teachers

	Coaches		Teachers	
	To receive...	From...	To receive...	From...
<b>Initial training (before program start)</b>	4 days of training	Program provider	2 days of training <sup>a</sup>	Program provider + coaches
<b>Ongoing support (throughout the program year)</b>	Monthly check-in	Program provider	Six 55-min guidance sessions and six 55-min reflection sessions <sup>b</sup>	Coach-facilitated discussion on pre-recorded guidance from program experts
	Regular email communication	Program provider	Tailored support for emerging challenges <sup>c</sup>	Coaches

NOTES: <sup>a</sup>Every teacher in the four districts that started as expected would attend two days (16 hours) of training as originally planned and every teacher in the two late-start districts would attend one day of condensed training (8 hours) that covered the same contents.

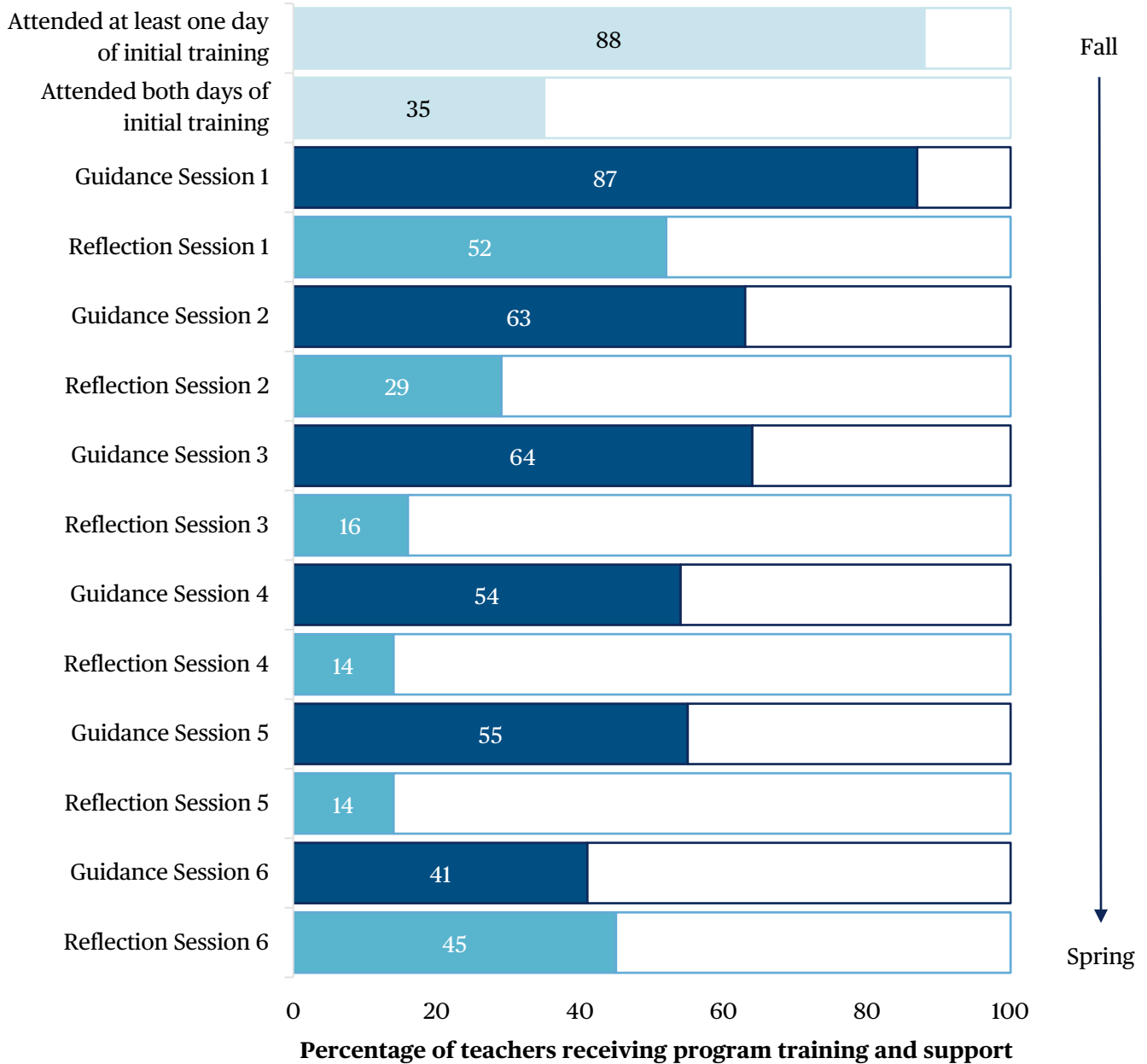
<sup>b</sup>Due to scheduling issues, not all six sessions were delivered in all school districts (see Appendix C, Section III for details).

<sup>c</sup>Includes one-on-one coaching sessions, support calls or web conferences, computer-based support in forms such as email exchanges, other group-based coaching, and other activities. Together with the Guidance and Reflection sessions, a coach was expected to provide 20 hours of support per month per school.

- ***The amount of training and ongoing support received by the teachers was less than intended.***

The initial training provided by the program provider was generally well attended. Specifically, 88 percent of the program-school teachers attended at least one day of initial training, and 35 percent of them attended two days of training (see Exhibit 7). In the four districts where the full two-day training was offered, 74 percent of teachers attended both days. In the two districts that joined the study late, the provider had to condense the two-day training into one day to accommodate those districts’ schedules.<sup>32</sup> In those two districts, 87 percent of teachers attended the one-day training. Overall, around 81 percent of the program-school teachers attended the full amount of initial training provided to them (two days for regular-start districts and one day for late-start districts).

### Exhibit 7. Receipt of Teacher Training and Professional Support Sessions Was Initially Strong, but Was Difficult to Sustain Throughout the School Year



SOURCES: Teacher training attendance records collected in summer/fall of 2017 and teacher-scheduled coaching attendance records collected throughout the 2017-2018 school year.

In contrast to the initial training, all districts experienced some challenges with participation in the ongoing training and supports. Over the program year, the average teacher attendance rate for the guidance sessions was 61 percent and for the reflection sessions was 28 percent, with participation in both starting higher but dropping over time (see Exhibit 7).<sup>33</sup> Coaches experienced difficulties with scheduling the ongoing support sessions, and both coaches and teachers reported the competing demands of other programs and testing as challenges they faced when delivering or attending the professional development activities.<sup>34</sup>

In addition, coaches reported delivering 44 hours of coaching *in total* during the program year across the study schools to which they were assigned (ranging from one to six schools), far below the targeted level of 20 hours per school per month.<sup>35</sup>

- ***Integrating the program’s daily lessons into existing curricula was a challenge.***

Both the program provider and coaches reported that making time to deliver the program lessons was one of the main challenges for teachers in the program schools, with 9 of the 11 coaches stating that there were too many demands on instructional time from their core subject matter curriculum and standardized testing.<sup>36</sup> Partially due to these challenges, teachers in the program schools made it through only about two-thirds (7.7) of the 12 units during the program year, on average.<sup>37</sup>

- ***The program changed word knowledge instruction in classrooms, as intended, but did not change other more challenging instructional practices.***

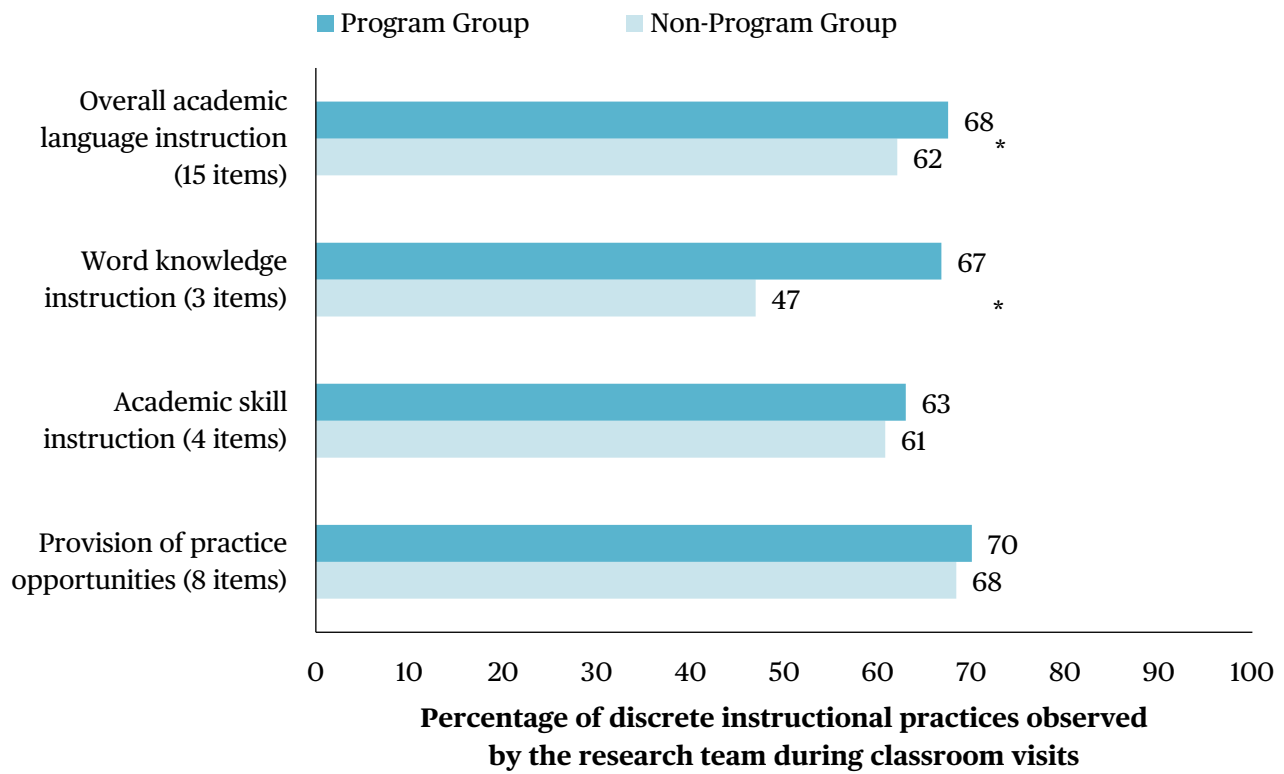
The study team observed that, overall, teachers in the program schools used more instructional practices that support the development of academic language than their counterparts in the non-program schools. However, the difference was largely driven by an increase in the use of practices focused on building students’ word knowledge (see Exhibit 8).<sup>38</sup> On average, the study team observed the program-school teachers using 5 percentage points more of the 15 instructional practices on the classroom observation checklist in their classrooms than teachers in the non-program schools. Of the three components that were the focus of the program, the two groups of teachers differed the most in their word knowledge instructional practices: teachers in the program schools used 20 percentage points more of the measured word knowledge practices than their counterparts in the non-program schools. In contrast, the two groups of teachers did not differ in their academic skills instruction or their provision of opportunities to practice academic language.<sup>39</sup>

These latter two, academic skills instruction and providing opportunities to practice academic language, are likely more challenging than word knowledge instruction to implement well, particularly to accommodate English learners and struggling readers.<sup>40</sup> Specifically, the program expected teachers to give students more agency in the classroom so that they could actively practice and better develop academic vocabulary and academic skills, an instructional approach that is particularly challenging.<sup>41</sup> Indeed, the program provider reported that it was hard for teachers to make this shift to more active student participation.

Additionally, more than half of the program-school teachers reported that it was difficult for them to modify activities or work to accommodate the specific needs of English learners and struggling readers.<sup>42</sup> The proportions of teachers reporting such issues were similar in the non-program schools, indicating that the program did not help teachers accommodate the needs of English learners and other struggling readers above and beyond what was happening in the non-program schools.

Furthermore, the additional training and ongoing support for academic language instruction that program teachers reported receiving compared to non-program teachers were only associated with increased use of word knowledge instruction. The training and support were not consistently associated with instruction on academic skills development or opportunities for student practice.<sup>43</sup> This lack of association between professional development and the use of more demanding practices might suggest that teachers need more extensive or different types of support to develop the pedagogical skills necessary to manage a classroom in which students are more actively involved in driving the activities.

### Exhibit 8. The Program Affected Academic Language Instruction Overall, but the Effect Is the Result of Word Knowledge Instruction



SOURCE: Classroom observations conducted in the spring of 2018.

NOTES: The sample includes 165 fourth- and fifth-grade classrooms from 57 study schools. Teachers' use of instructional practices important for academic language development is measured by the percentage of such practices overall and in each of the three instruction components as recorded during classroom observation.

\* indicates that the estimated difference between the program and non-program schools is statistically significant at the 0.05 level.

- ***Changes in word knowledge instruction alone might not have been enough to change the key student outcomes.***

The study found a weak relationship between the use of program-specific word knowledge instruction and average student outcomes in the program schools, suggesting that more of this type of instruction might not improve those key measures. However, students in the program schools where teachers did more of the other key instructional components—academic skill instruction or opportunities for students to practice academic language—had higher average scores on the academic language skill test than students in those program schools where teachers used those practices less.<sup>44</sup> These analyses suggest the importance of improving all three areas of instruction, consistent with existing research, which shows that programs focused on vocabulary instruction often yield significant effects on word-level outcomes (such as vocabulary, decoding, and fluency) but not on other reading outcomes, like comprehension, unless vocabulary instructional practices are integrated with other strategies.<sup>45</sup>



## LESSONS LEARNED AND LOOKING FORWARD

This study highlights the challenges of integrating supplementary materials into the school day and implementing the kinds of academic language classroom practices that research suggests are beneficial for students from families with low incomes or whose first language is not English. Given the persistent differences in achievement between the groups targeted in this study and their peers—differences which may be exacerbated by the pandemic’s disruptions to the education system—it is important to draw insights that might help refine this or similar programs or make them easier to adopt.

- **Improving the integration of program activities.** The program was designed to be flexible and could either supplement or replace existing practices across subject areas. However, teachers found it difficult to fit program activities into their existing instructional time, even with support. The program may need to provide easier or more explicit ways to facilitate integration and achieve its desired effects. Schools seeking to adopt the program also need to commit time and resources for teachers to participate in professional development activities and to fit the program curriculum into daily instruction.
- **Providing enough support for teachers.** This study planned intensive professional development and coaching over a single school year. While most teachers received the intended content of the initial training, their receipt of ongoing support during the school year fell well short of the intended amount due to access to schools and scheduling challenges. If the planned level of support is needed to change practice, then it may not be feasible to fully implement the program’s instructional approaches within a single school year. In fact, a prior study of the same academic language program found positive effects on reading comprehension only after the second year of implementation when the intended amount of training and support were delivered, but not after the first year.<sup>46</sup>
- **Strengthening the focus on instructional practices expected to improve student outcomes.** Existing research indicates that instruction on academic word knowledge is not enough to affect students’ academic language skills and reading ability.<sup>47</sup> Although the program emphasized activities that were hypothesized to lead to improved student reading comprehension, such as providing students with opportunities to use academic language through class discussion and debate, only word knowledge instruction increased in the program schools.<sup>48</sup> Studies of other programs that provided intensive professional development to improve instruction have similarly found impacts on some practices, but not on others—typically the most challenging practices.<sup>49</sup> Professional development providers may want to consider different ways to support teachers’ implementation of the more difficult practices with future research examining whether they lead to improved student outcomes.

## ENDNOTES

---

<sup>1</sup>National Assessment of Educational Progress (2019a); National Assessment of Educational Progress (2019b). Economically disadvantaged students are defined by eligibility for the National School Lunch Program (NSLP).

<sup>2</sup>For example, only 4 percent of eighth-grade English learners and 20 percent of economically disadvantaged students are reading at or above proficient level, compared with 34 percent of overall students (National Assessment of Educational Progress, 2019b). Similarly, the proportion of fourth-graders with at or above proficiency math performance is 16 percent among English learners, 26 percent among students eligible for the NSLP, and 41 percent overall (National Assessment of Educational Progress, 2019a).

<sup>3</sup>National Center for Education Statistics (2021). Also see Batalova and McHugh (2010), Bidwell (2015), and Walker (2015).

<sup>4</sup>Research suggests that underdeveloped academic language is one reason these students are at particular risk for poor academic outcomes (Kieffer, 2010).

<sup>5</sup>For review, see Lesaux, Koda, Siegel, and Shanahan (2006); Bailey and Heritage (2008); Foorman et al. (2015); Guerrero (2004); Hakuta, Goto Butler, and Witt (2000); Honig (2010); Shanahan and Shanahan (2008); Kim, Hsin, and Snow (2018). Reading interventions can help struggling readers in upper elementary grades to catch up on their reading skills. Research has shown that the growth rate in reading declines as grade level increases. The annual spring-to-spring growth in reading for Grade 4 is 0.40 standard deviations in effect size, and for Grade 5 is 0.32 standard deviations (Bloom, Hill, Black, and Lipsey, 2008, Table 1). On average, the effect of reading interventions on struggling readers in Grades 4 and 5 was 0.30 standard deviations in effect size, on par with their expected annual growth rate (Scammacca, Roberts, Vaughn, and Stuebing, 2015, Table 4).

<sup>6</sup>The WordGen Elementary program was selected through a competitive process. See Appendix A, Section I for a discussion of this process.

<sup>7</sup>See Appendix A, Section II for more details about the content of the program.

<sup>8</sup>See Appendix A, Section III for more details about the intended training and support for the teachers.

<sup>9</sup>Truckenmiller, Park, Dabo, and Wu Newton (2019) reviewed eight studies on academic language instructional programs that meet the What Works Clearinghouse (WWC) standard without reservation and found evidence that these programs had statistically significant or substantively important impacts on academic vocabulary, morphology, and content-area comprehension, content-area achievement for diverse classrooms with both English learners and non-English learner struggling readers. These studies evaluated instructional elements such as academic vocabulary instruction, instruction within content area, peer discussion, and writing. Among the eight studies of academic language programs reviewed by Truckenmiller, Park, Dabo, and Wu Newton (2019), only one reported a significant impact on reading comprehension. Jones et al. (2019) studied the same program as the current study and found positive program effects on reading comprehension after two years of implementation.

<sup>10</sup>The study recruited and randomly assigned 70 schools in the summer and early fall of 2017. After the random assignment and prior to the start of the program in the fall of 2017, 12 schools (4 program schools and 8 non-program schools) decided that they could not accommodate the evaluation requirements and withdrew from the study. Statistical tests show no systematic difference between the schools that declined to participate in the study and those that remained. Section I of Appendix B presents details of the recruitment process and the comparisons among these school samples. Exhibit C.1 provides specific sample sizes for each analysis sample.

<sup>11</sup>The national sample includes all public regular elementary schools that are eligible for the school-wide Title I program, serving students in Grades 4 and 5, and which are not a magnet, charter, or virtual school based on school information from the 2016-2017 school year.

<sup>12</sup>For districts where the proportions of English learners varied widely among study schools, schools with similar proportions of English learners were blocked together within the district. In addition, schools in one district were also blocked by when they agreed to participate in the study. Random assignment was conducted within each block. Overall, separate random assignment was carried out in 11 such blocks. Within each block, roughly half of the study schools were randomly assigned to participate in the program. The other half were assigned to continue with usual practices.

<sup>13</sup>Appendix B provides details of the data-collection activities and measures used in the evaluation.

<sup>14</sup>Due to logistic concerns, the team did not administer the CALS-I and GMRT tests at baseline. See Appendix B, Section II, “Study Administered Tests,” for a discussion of the reasons.

<sup>15</sup>The study randomly selected three classrooms across the fourth and fifth grades in each school to be observed by certified observers twice during the program year. The observers used two academic language instructional practice checklists in each observation: one checklist covered instructional practices that are generally considered important for academic language learning and were used to measure the alignment with these practices for teachers in both the program and non-program schools. The other checklist covered program-specific instructional practices and were used to measure program-school teachers’ implementation fidelity. See Appendix B, Section III for details of these measures, including a comparison between the alignment score and the implementation fidelity score.

<sup>16</sup>Beck, Perfetti, and McKeown (1982); Jenkins, Stein, and Wysocki (1984); and Bolger, Balass, Landen, and Perfetti (2008) show that introducing key words repeatedly in authentic and varied contexts supports word learning and comprehension; Carlo et al. (2004) demonstrates that teaching about derivational morphology and common roots could foster students’ word learning on their own.

<sup>17</sup>Diazgranados, Selman, and Dionne (2016).

<sup>18</sup>See Lawrence and Snow (2010) for a review of this literature.

<sup>19</sup>See Table 14 in Barr and Uccelli (2016) for the conversion from scaled scores to percentile ranks; see Table 1 in Barr, Uccelli, and Phillips Galloway (2019) for characteristics of the CALS-I norming sample; see Exhibit B.12 in Appendix B, Section III for characteristics of the CALS-I study sample.

<sup>20</sup>Most of the research on academic language proficiency has focused almost exclusively on academic vocabulary (Lawrence et al., 2012; Lesaux, Kieffer, Faller, and Kelley, 2010; Nagy and Townsend, 2012). One recent study of the WordGen Elementary program used the same broad assessment of students’ general academic language skills in addition to their program specific vocabulary assessment (Jones et al., 2019).

<sup>21</sup>See Exhibit C.1 in Appendix C, Section I for details on the grade equivalents of the average scores. The state tests used in the study districts all include components beyond reading comprehension. The specific components vary by test but generally include analysis and interpretation of literacy text and informational text, writing, and knowledge of language conventions.

<sup>22</sup>The sample of students used in this analysis includes all students enrolled in the study schools and had valid state test scores. This sample is larger than the sample used for the analysis of CALS-I and GMRT, which is restricted to students who consented to the study data collection and took the tests. Limiting the sample to be consistent across outcomes does not change the pattern of the findings. See Appendix D for detailed impact analysis findings for the alternative sample. The estimated program impacts do not vary significantly across the six study districts, even though the magnitude of the estimated impact appeared to be large in some districts.

Therefore, there was no strong evidence that the program might have worked in some places but not others. Exhibit C.2 in Appendix C, Section I presents the district-level impact estimates for all three outcomes.

<sup>23</sup>These students were fourth- and fifth-graders in study schools during the program year and were fifth- and sixth-graders in the follow-up year (2018-2019 school year). See Exhibit C.4 in Appendix C, Section I for details of these findings.

<sup>24</sup>These resources include newsletters for families, vocabulary cards and slides in multiple languages, and more, all available from the provider's website (<https://www.serpoinstitute.org/wordgen-weekly/wge-with-english-learners>). Information can be downloaded from the provider's website (<https://irp-cdn.multiscreensite.com/7a45b809/files/uploaded/Tips%20for%20Supporting%20ELs%20-%20One%20Page.pdf>).

<sup>25</sup>Baker et al. (2014).

<sup>26</sup>Kieffer (2010); Uccelli, et al. (2015); Neuman, Kaefer, and Pinkham (2017).

<sup>27</sup>The average scaled scores for English learners translate into thirty-seventh and twenty-eighth percentiles for fourth- and fifth-graders. The average scaled scores for students from economically disadvantaged backgrounds translate into fifty-seventh and sixty-fifth percentile for fourth- and fifth-graders, respectively. As noted before, however, the norming sample used for the conversion differs from the study sample in student background characteristics (see Barr and Uccelli, 2016; Barr, Uccelli, and Phillips Galloway, 2019).

<sup>28</sup>See Exhibit C.4 in Appendix C, Section I for details of the findings for the follow-up year.

<sup>29</sup>See Appendix A, Section III for the proposed training content.

<sup>30</sup>The majority of coaches served in a district- or school-related capacity prior to the evaluation; however, a few coaches who were not already district employees were recruited for the study.

<sup>31</sup>See Lawrence and Snow (2010) for a review of the effect of discussion and debate in the classroom on reading comprehension, writing, and content-area learning.

<sup>32</sup>Two districts joined the study later than the others. Program group teachers from these two districts participated in initial training in October and December 2017. They started teaching the program in their classrooms in mid-October and mid-December 2018. Appendix D, Section III provides detailed discussions about the amount of initial training and ongoing support received by teachers in the four districts that started the program on time, in the two late-start districts, as well as the contrast in fidelity scores between these two groups of districts (Exhibit D.20).

<sup>33</sup>See Exhibits C.10 and C.11 for details of session attendance. The study provided six guidance sessions to all six study districts. However, due to scheduling challenges, the study was not able to schedule some of the reflection sessions in three districts. When focusing only on the provided reflection sessions, the overall attendance rate for teachers was 49 percent across all districts, and it followed the same pattern over time as presented in Exhibit 7.

<sup>34</sup>See Appendix C, Section III, "Challenges to Implementation" for more details.

<sup>35</sup>This number included hours devoted to guidance and reflection sessions, other group-based coaching, one-on-one sessions, support calls or web conferences, computer-based support such as email exchanges, and other activities. Coaching hours were not discretely distributed across schools. For example, coaching time devoted to a single email exchange may be directed to several teachers from different schools. Therefore, on average, each coach provided a total of 44 hours of supports to the schools they worked with during the program year.

<sup>36</sup>In response to an open-ended question about “the biggest challenge for teachers in implementing WordGen” in the coach reports, 9 of the 11 coaches identified competing demands from other programs and testing as one of the main challenges for teachers at various times during the implementation year.

<sup>37</sup>Exhibit C.14 in Appendix C, Section III presents the overall and district-level unit completion rates and available instructional days for the program schools. Other possible explanations for this level of unit completion include the late start of the program in two school districts, and developers’ message to teachers that sometimes spending more time on a unit is more important than covering more units.

<sup>38</sup>A given practice was considered as used by a teacher if it was observed during the 40-minute classroom observation period. Therefore, this measure captured teachers’ use of the practice but did not reflect the quality of the instruction. See Appendix B, Section III for more details about the measures.

<sup>39</sup>Analyses of the classroom observation data from the CLASS-UE also showed no differences in more general classroom management and teaching practices between the program and non-program schools. See Exhibit C.6 for detailed findings.

<sup>40</sup>Cazden and Beck (2003); Lampert (2015).

<sup>41</sup>See Lawrence and Snow (2010) for a review of the effect of discussion and debate in the classroom on reading comprehension, writing, and content-area learning.

<sup>42</sup>See Exhibit D.24 in Appendix D, Section III for details of the findings based on teacher-reported challenges.

<sup>43</sup>These relationships reflected the estimated correlations between the amount of training and support the program-school teachers reportedly received and their use of the core instructional practices emphasized by the program (as measured by fidelity scores). See Exhibit C.7 in Appendix C, Section II for detailed findings of this analysis.

<sup>44</sup>These relationships reflected the estimated correlations between degree of program implementation and student outcome levels among the program schools. See Appendix B, Section III, for details of the analytic approach and Appendix C, Section II for details of findings (Exhibit C.8).

<sup>45</sup>Wright and Cervetti (2017) found very limited evidence that direct teaching of word meanings, even if on large numbers of words, can improve students’ general reading comprehension. They also identified strategies that teach students to self-monitor their understanding of words and to use multiple strategies for understanding unknown words in the text as promising areas for research.

<sup>46</sup>Jones et al. (2019) found impacts on academic language and reading comprehension measures after the second year implementing the same program studied here. During the first year of implementation for that study, the program was still under development so the program provider and coaches were involved in curriculum development while also trying to provide support to teachers. The second year of implementation is a closer match to the planned program year implementation in this study in terms of the readiness of the curriculum for use and the levels of training and support planned for teachers.

<sup>47</sup>Wright and Cervetti (2017) reviewed 36 studies of vocabulary interventions and concluded that “there is very limited evidence that direct teaching of word meanings, even long term, multifaceted interventions of large numbers of words, can improve generalized comprehension.” A recent study of Word Knowledge Instruction also found no impact on students’ reading skills (Foorman et al., 2021).

<sup>48</sup>For example, Mezynski (1983) identified three factors for vocabulary instruction to transfer to comprehension: Amount of practice, breadth of training in word usage, and active processing (p. 273). Sampson, Valmont, and Van Allen (1982) also found that teaching students to engage in different types of semantic and syntactic analysis of texts could lead to positive impacts on generalized comprehension.

<sup>49</sup>For example, several IES studies of programs providing intensive professional development found no or small impacts on more challenging instructional practices compared to impacts on explicit teacher instruction. In addition, these studies also found the amount of high-quality student practice opportunities to be only a small proportion of the observation period in both the treated and untreated groups (Gamse et al., 2008; Garet et al., 2016).

## REFERENCES

- Bailey, Alison L., and Margaret Heritage. 2008. *Formative Assessment for Literacy, Grades K-6: Building Reading and Academic Language Skills Across the Curriculum*. Thousand Oaks, CA: Corwin Press.
- Baker, Scott, Esther Geva, Michael J. Kieffer, Nonie Lesaux, Sylvia Linan-Thompson, Joan Morris, C. Patrick Proctor, Randi Russell, Russell Gersten, Joseph Dimino, Madhavi Jayanthi, Kelly Haymond, and Rebecca Newman-Gonchar. 2014. *Teaching Academic Content and Literacy to English Learners in Elementary and Middle School (NCEE 2014-4012)*. Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education.
- Barr, Christopher D., and Paola Uccelli. 2016. "CALS-I Psychometric Report." Unpublished internal report prepared for the Institute of Education Sciences-funded Catalyzing Comprehension through Discussion and Debate research project.
- Barr, Christopher D., Paola Uccelli, and Emily Phillips Galloway. 2019. "Specifying the Academic Language Skills That Support Text Understanding in the Middle Grades: The Design and Validation of the Core Academic Language Skills Construct and Instrument." *Language Learning* 69, 4: 978-1021. DOI: 10.1111/lang.12365.
- Batalova, Jeanne, and Margie McHugh. 2010. "Number and Growth of Students in US Schools in Need of English Instruction." Migration Policy Institute. Website: [https://www.migrationpolicy.org/sites/default/files/publications/FactSheet%20ELL1%20-%20FINAL\\_0.pdf](https://www.migrationpolicy.org/sites/default/files/publications/FactSheet%20ELL1%20-%20FINAL_0.pdf).
- Beck, Isabel L., Charles A. Perfetti, and Margaret G. McKeown. 1982. "Effects of Long-Term Vocabulary Instruction on Lexical Access and Reading Comprehension." *Journal of Educational Psychology* 74, 4: 506-521. DOI: 10.1037/0022-0663.74.4.506.
- Bidwell, Allie. 2015. "Most U.S. Students Live in or Near Poverty." U.S. News & World Report. Website: <https://www.usnews.com/news/blogs/data-mine/2015/01/16/most-us-students-come-from-low-income-families>.
- Bloom, Howard S., Carolyn J. Hill, Alison Rebeck Black, and Mark W. Lipsey. 2008. "Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions." *Journal of Research on Educational Effectiveness* 1, 4: 289-328.
- Bolger, Donald J., Michal Balass, Eve Landen, and Charles A. Perfetti. 2008. "Context Variation and Definitions in Learning the Meanings of Words: An Instance-Based Learning Approach." *Discourse Processes* 45, 2: 122-159. DOI: 10.1080/01638530701792826.
- Carlo, María S., Diane August, Barry McLaughlin, Catherine E. Snow, Cheryl Dressler, David N. Lippman, Teresa J. Lively, and Claire E. White. 2004. "Closing the Gap: Addressing the Vocabulary Needs of English-Language Learners in Bilingual and Mainstream Classrooms." *Reading Research Quarterly* 39, 2: 188-215. DOI: 10.1598/RRQ.39.2.3.
- Cazden, Courtney B., and Sarah W. Beck. 2003. "Classroom discourse." Pages 165-197 in Arthur C. Graesser and Morton Ann Gernsbacher (eds.), *Handbook of Discourse Processes*. Mahwah, NJ: Lawrence Erlbaum Associates.

- Diazgranados, Silvia, Robert L. Selman, and Michelle Dionne. 2016. "Acts of Social Perspective Taking: A Functional Construct and the Validation of a Performance Measure for Early Adolescents." *Social Development* 25, 3: 572-601. DOI: 10.1111/sode.12157.
- Foorman, Barbara R., Sarah Herrera, Jennifer L. Dombek, Carla Wood, Linda Gaughn, and Lynn Dougherty Underwood. 2021. *The Impact of Word Knowledge Instruction on Literacy Outcomes in Grade 5 (REL 2021-083)*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast.
- Foorman, Barbara R., Sharon Koon, Yaacov Petscher, Alison Mitchell, and Adrea J. Truckenmiller. 2015. "Examining General and Specific Factors in the Dimensionality of Oral Language and Reading in 4th-10th Grades." *Journal of Educational Psychology* 107, 3: 884-899. DOI: 10.1037/edu0000026.
- Gamse, Beth C., Robin Tepper Jacob, Megan Horst, Beth Boulay, Fatih Unlu, Laurie Bozzi, Linda Caswell, Chris Rodger, W. Carter Smith, Nancy Brigham, and Sheila Rosenblum. 2008. *Reading First Impact Study Final Report (NCEE 2009-4038)*. Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education.
- Garet, Michael S., Jessica Heppen, Kirk Walters, Toni Smith, and Rui Yang. 2016. *Does Content-Focused Teacher Professional Development Work? Findings From Three Institute of Education Sciences Studies (NCEE 2017-4010)*. Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education.
- Guerrero, Michael D. 2004. "Acquiring Academic English in One Year: An Unlikely Proposition for English Language Learners." *Urban Education* 39, 2: 172-199. DOI: 10.1177/0042085903260915.
- Hakuta, Kenji, Yuko Goto Butler, and Daria Witt. 2000. "How Long Does It Take English Learners to Attain Proficiency?" (January). Stanford: University of California Linguistic Minority Research Institute.
- Honig, Sheryl L. 2010. "A Framework for Supporting Scientific Language in Primary Grades." *The Reading Teacher* 64, 1: 23-32. DOI: 10.1598/RT.64.1.3.
- Jenkins, Joseph R., Marcy L. Stein, and Katherine Wysocki. 1984. "Learning Vocabulary Through Reading." *American Educational Research Journal* 21, 4: 767-787. DOI: 10.3102/00028312021004767.
- Jones, Stephanie M., Maria LaRusso, James Kim, Ha Yeon Kim, Robert Selman, Paola Uccelli, Sophie P. Barnes, Suzanne Donovan, and Catherine E. Snow. 2019. "Experimental Effects of Word Generation on Vocabulary, Academic Language, Perspective Taking, and Reading Comprehension in High-Poverty Schools." *Journal of Research on Educational Effectiveness* 12, 3: 448-483. DOI: 10.1080/19345747.2019.1615155.
- Kieffer, Michael J. 2010. "Socioeconomic Status, English Proficiency, and Late-Emerging Reading Difficulties." *Educational Researcher* 39, 6: 484-486. DOI: 10.3102/0013189X10378400.
- Kim, Ha Yeon, Lisa B. Hsin, and Catherine E. Snow. 2018. "Reducing Academic Inequalities for English Language Learners: Variation in Experimental Effects of Word Generation in High-Poverty Schools." *International Journal of Bilingual Education and Bilingualism* 24, 7: 1024-1042. DOI: 10.1080/13670050.2018.1535574
- Lampert, Magdalene. 2015. *Deeper Teaching*. Students at the Center: Deeper Learning Research Series. Boston: Jobs for the Future.



- Lawrence, Joshua F., and Catherine E. Snow. 2010. "Oral Discourse and Reading." Pages 320-338 in Michael L. Kamil, David Pearson, Elizabet Birr Moje, and Peter Afflerbach (eds.), *Handbook of Reading Research, Vol. IV*. London: Routledge.
- Lawrence, Joshua F., Lauren Capotosto, Lee Branum-Martin, Claire White, and Catherine E. Snow. 2012. "Language Proficiency, Home-Language Status, and English Vocabulary Development: A Longitudinal Follow-Up of the Word Generation Program." *Bilingualism: Language and Cognition* 15, 3: 437-451. DOI: 10.1017/S1366728911000393.
- Lesaux, Nonie, with Keiko Koda, Linda S. Siegel, and Timothy Shanahan. 2006. "Development of Literacy." Pages 75-122 in Diane L. August and Timothy Shanahan (eds.), *Developing Literacy in Second-Language Learners: Report of the National Literacy Panel on Language-Minority Children and Youth*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Lesaux, Nonie K., Michael J. Kieffer, S. Elisabeth Faller, and Joan G. Kelley. 2010. "The Effectiveness and Ease of Implementation of an Academic Vocabulary Intervention for Linguistically Diverse Students in Urban Middle Schools." *Reading Research Quarterly* 45, 2: 196-228. DOI:10.1598/RRQ.45.2.3.
- Mezynski, Karen. 1983. "Issues Concerning the Acquisition of Knowledge: Effects of Vocabulary Training on Reading Comprehension." *Review of Educational Research* 53, 2: 253-279. DOI: 10.3102/00346543053002253.
- Nagy, William, and Dianna Townsend. 2012. "Words as Tools: Learning Academic Vocabulary as Language Acquisition." *Reading Research Quarterly* 47, 1: 91-108. DOI: 10.1002/RRQ.011.
- National Assessment of Education Progress. 2019a. "NAEP Mathematics: National Achievement-Level Results." Website: <https://www.nationsreportcard.gov/mathematics/nation/achievement/?grade=4>.
- National Assessment of Education Progress. 2019b. "NAEP Reading: National Achievement-Level Results." Website: <https://www.nationsreportcard.gov/reading/nation/achievement/?grade=4>.
- National Center for Education Statistics. 2021. "English Language Learners in Public Schools." Website: [https://nces.ed.gov/programs/coe/indicator\\_cgf.asp](https://nces.ed.gov/programs/coe/indicator_cgf.asp).
- Neuman, Susan B., Tanya Kaefer, and Ashley M. Pinkham. 2017. "A Double Dose of Disadvantage: Language Experiences for Low-Income Children in Home and School." *Journal of Educational Psychology* 110, 1: 102-118. DOI: 10.1037/edu0000201.
- Sampson, Michael R., William J. Valmont, and Roach Van Allen. 1982. "The Effects of Instructional Cloze on the Comprehension, Vocabulary, and Divergent Production of Third-Grade Students." *Reading Research Quarterly* 17, 3: 389-399. DOI: 10.2307/747526.
- Scammacca, Nancy K., Greg Roberts, Sharon Vaughn, and Karla K. Stuebing. 2015. "A Meta-Analysis of Interventions for Struggling Readers in Grades 4-12: 1980-2011." *Journal of Learning Disabilities* 48, 4: 369-390.
- Shanahan, Timothy, and Cynthia Shanahan. 2008. "Teaching Disciplinary Literacy to Adolescents: Rethinking Content-Area Literacy." *Harvard Educational Review* 78, 1, 40-59. DOI: 0.17763/haer.78.1.v62444321p602101.
- Truckenmiller, Adrea J., Jiyoung Park, Arfang Dabo, and Yi-Chieh Wu Newton. 2019. "Academic Language Instruction for Students in Grades 4 Through 8: A Literature Synthesis." *Journal of Research on Educational Effectiveness* 12, 1: 135-159. DOI: 10.1080/19345747.2018.1536773.

- Uccelli, Paola, Christopher D. Barr, Christina L. Dobbs, Emily Phillips Galloway, Alejandra Meneses, and Emilio Sanchez. 2015. "Core Academic Language Skills: An Expanded Operational Construct and a Novel Instrument to Chart School-Relevant Language Proficiency in Preadolescent and Adolescent Learners." *Applied Psycholinguistics* 36, 5: 1077-1109. DOI: 10.1017/S014271641400006X.
- Walker, Tim. 2015. "Shameful Milestone: Majority of Public School Students Live in Poverty." National Education Association. Website: <https://www.nea.org/advocating-for-change/new-from-nea/shameful-milestone-majority-public-school-students-live-poverty>.
- Wright, Tanya S., and Gina N. Cervetti. 2017. "A Systematic Review of the Research on Vocabulary Instruction That Impacts Text Comprehension." *Reading Research Quarterly* 52, 2: 203-226. DOI: 10.1002/rrq.163.

## ACKNOWLEDGMENTS

This study represents a collaborative effort among our colleagues at MDRC, Abt Associates, and the Florida Center for Reading Research (FCRR) at Florida State University. We also received advice, guidance, and support from Christopher Barr of Rice University and Paola Uccelli of Harvard University—particularly regarding the administration, scoring, and understanding of the CALS-I, the academic language assessment used in this study.

This study would not have been possible without the participation and collaboration of the six school districts and their schools included in the evaluation. We thank the teachers, instructional coaches, and school leaders for their efforts to implement the WordGen Elementary program. We also appreciate the willingness of school staff to provide information on how they implemented the program. Staff in participating schools and districts completed surveys, provided student records data, and supported the study team as it conducted multiple rounds of data collection activities in the schools.

Suzanne Donovan, Alexandra Huyghe, and their team of Strategic Education Research Partnership (SERP) colleagues and affiliates—including Catherine Snow, Carol O’Connor, Carol Mahedy, and Gladys Aguilar—were essential in supporting teachers, instructional coaches, schools, and districts in the implementation of WordGen Elementary. Throughout the planning and execution of program implementation, the SERP team was responsive and patient with requests from the study team for updates, information, and more.

We also benefited from the advice of the members of our Technical Working Group who provided expert insights into academic language assessments, instruction that facilitates academic language learning, and research methodology. Members of this group over the course of the project have included Amy Crosson, Penn State University; David Figlio, Northwestern University; David Francis, University of Houston; Jeannette Mancilla-Martinez, Vanderbilt University; C. Patrick Proctor, Boston College; Jeff Smith, University of Wisconsin-Madison; and Julie Washington, University of California-Irvine.

In addition, we wish to extend our thanks to Fred Doolittle, Marie-Andrée Somers, John Hutchins, Leigh Parise, and Will Swarts for providing helpful review on the report, and Luisa LaFleur, Joshua Malbin, Carolyn Thomas, and Alice Tufel for editing and preparing the report for publication.

The report would not have been possible without the contributions of many current and prior staff at MDRC, Abt Associates, and FCRR. At MDRC, we thank Alexandra Cazeils, Kelly Granito, Rob Ivry, Bryce Marshall, Vivi Nguyen, Leigh Parise, Shelley Rappaport, Sarah Salimi, Lauren Scarola, and Jacklyn Altuna Willard. At Abt, we thank Jennifer Ash, Erin Bumgarner, Hanna Carr, Carter Epstein, Alyssa Rulf Fountain, Yeqin He, Meredith Kelsey, Rachel McCormick, Brenda Rodriguez, David Rodriguez, and Haisheng Yang. At FCRR, we benefitted from the thoughtfulness and work of Don Compton, Jennifer Dombek, Jaime Elliott, Barbara Foorman, Ann McDonnell, Adrea Truckenmiller, and Yi-Chieh Newton.

## **DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST**

The research team for this evaluation included staff from MDRC and its subcontractors, Abt Associates, Florida Center for Reading Research (FCRR), Christopher Barr of Rice University, and Paola Uccelli of Harvard University. None of the research team members has financial interests that could be affected by findings from this evaluation. None of the seven members of the Technical Working Group, convened two times by the research team over the course of the study to provide advice and guidance, has financial interests that could be affected by findings from the evaluation.