The Impact of Word Knowledge Instruction on Literacy Outcomes in Grade 5

Appendix A. Data, intervention, sample, and methodology

See https://go.usa.gov/xsAkf for the full report.

Appendix A. Data, intervention, sample, and methodology

This appendix describes the study data, intervention, analytic sample, and analyses used to assess the impact of Word Knowledge Instruction (WKI) on student literacy outcomes in grade 5.

Study data

This study used data from both primary and secondary sources.

- The primary data consisted of researcher-developed, short-term measures (real-word decomposition, nonword derivation, and inferencing of word meanings) that were administered by participating teachers in spring 2019 and served to assess short-term outcomes.

- The secondary data consisted of administrative data from the participating school district on students whose parents consented to their participation in the study:

  o Scores on the August 2018 administration of the i-Ready Reading test, which served as the pretest for the short-term outcomes.

  o For the long-term vocabulary outcome, scores on the August 2018 administration of i-Ready Vocabulary served as the pretest for the i-Ready Vocabulary outcome (administered April 2019).

  o For the long-term reading comprehension outcome, scores on the late spring 2018 administration of the Florida Standards Assessment English Language Arts (FSA–ELA) served as the pretest for the FSA-ELA outcome (administered late spring 2019).

  o Student demographic characteristics for the 2018/19 school year (eligibility for the national school lunch program, English learner status, and race/ethnicity).

Outcome measures

This section describes the researcher-developed, short-term outcome measures and the long-term outcome measures used in study analyses.

Researcher-developed, short-term outcome measures. Three researcher-developed short-term outcome measures were used for the study to assess students' morphological awareness: real-word decomposition, nonword derivation, and inferencing of word meanings. The real-word decomposition and the nonword derivation measures were adapted from measures used in an evaluation of a program of morphological awareness instruction in California that consisted of 45 minutes of daily, whole-class word knowledge instruction for 20 weeks with grade 6 students (Lesaux et al., 2014). Lesaux and colleagues derived both measures from previous

research (Carlisle, 2000, and Carlo et al., 2004, for real-word decomposition, and Nagy et al., 2006 and Tyler & Nagy, 1989, for nonword derivation). The inferencing of word meanings outcome measure was constructed directly from science or social studies passages taken from the grade 5 core reading program in the Florida school district.

A paper test form was used for assessing student performance on all researcher-developed, short-term outcomes. Study team members entered student responses into the study database, where a database algorithm automatically scored them as correct or incorrect. Study team members then checked each other's input and corrected any data entry errors.

Real-word decomposition. For assessing student performance on the real-word decomposition outcome measure, teachers read aloud a word (for example, *divisible*) containing a suffix (*-ible*) that was taught in WKI and asked students to extract the base word (*divide*) and write it in the blank (for example, *"Please ___ the cake into small pieces."*). Although each item included a suffix that was taught in WKI, the base word was not included in WKI activities associated with the taught suffix. In addition, slightly more than half of the suffixes (11) could also have been taught to the control group because the state-adopted core English language arts (ELA) materials were available to all teachers and these materials included instruction on those suffixes. Therefore, this outcome is not considered to be strongly aligned to the WKI treatment program.

Student responses were scored as either correct or incorrect. Internal consistency (Cronbach's alpha) reliability for the study sample was .85, which exceeds the .50 threshold for an outcome measure established by What Works Clearinghouse (2020b; table A1). The real-word decomposition outcome measure correlated significantly and moderately with the other measures of vocabulary and reading comprehension, with convergent validity correlations ranging from .57 to .60. A correlation above r = .85 would suggest that the two measures were measuring the same construct, whereas a nonsignificant correlation would indicate no relationship between the measures.

Nonword derivation. For assessing student performance on the nonword derivation measure, teachers read aloud a sentence (for example, "The man is a great ____.") and asked students to complete the sentence by choosing the nonsense base word that had an appropriate suffix (for example, *tranter*) from among the four answer choices (*tranter*, *tranting*, *trantitious*, and *trantiful*). This outcome measure was not considered to be strongly aligned to WKI because none of the nonsense base words were taught in WKI and the answer choices included both suffixes that were taught in WKI and suffixes that were not.

Internal consistency (Cronbach's alpha) reliability for the current study sample was .76. The nonword derivation measure correlated significantly and moderately with the other vocabulary and reading comprehension measures, with convergent validity correlations ranging from .47 to .52 (see table A1).

Inferencing of word meanings. The inferencing of word meanings outcome measure was constructed from science and social studies texts that are part of the core grade 5 ELA program. The researchers selected 15 sentences that included a word with an affix that was taught in WKI and a base word that was not included in WKI activities associated with the taught affix. For example, the word *tireless* (in the sentence "Soon, Anthony became a *tireless* advocate of women's rights in all possible ways.") includes a WKI taught affix *(-less)* with a base word that is not included in WKI instruction (*tire)*.

For assessing student performance on inferencing of word meanings, teachers read aloud each of the selected sentences and asked students to choose the word that meant the same as the italicized target word in the sentence from among four choices. Answer choices for the example cited above included the synonym *energetic* as well as the words *exhausted*, *reckless*, and *weaken*. Although each target word included an affix that was taught in WKI, 60 percent of the affixes were also taught in the core ELA program and the selected sentences were taken

from passages in the core ELA program available to teachers of both the treatment and the control groups. Therefore, this outcome is not considered to be strongly aligned to the WKI treatment program.

Internal consistency (Cronbach's alpha) reliability for the study sample was .61. The inferencing of word meanings outcome measure correlated significantly and moderately with the other two vocabulary and reading comprehension measures, with convergent validity correlations ranging from .56 to .57 (see table A1).

Table A1. Internal consistency reliability and convergent validity correlations for the three researcher-developed, short-term measures, 2018/19

| Measure | Reliability | | Convergent validity correlation with | | | |
	Cronbach's alpha	Number of students	i-Ready Vocabulary	Number of students	Florida Standards Assessment English Language Arts	Number of students
Real-word decomposition	.85	3,365	.57	2,448	.60	2,303
Nonword derivation	.76	3,365	.47	2,451	.52	2,305
Inferencing of word meanings	.61	3,361	.57	2,448	.56	2,302

Note: All correlations were significant at $p < .001$.
Source: Authors' analysis of school district data for 2018/19.

Long-term outcome measures. Two long-term outcome measures that covered the two target domains—vocabulary and reading comprehension—were used in the current study: i-Ready Vocabulary scores and FSA–ELA scores.[1] The i-Ready Vocabulary assessment was administered three times a school year (late August, early December, and early April). Scores from the i-Ready Vocabulary assessment were provided for the 2018/19 school year, with the first administration of the i-Ready Vocabulary assessment serving as the pretest. Data from the FSA–ELA were provided for the 2017/18 and 2018/19 school years, with scores from the 2017/18 administration serving as the pretest. Correlations between baseline and outcome measures are reported in table A2.

Table A2. Correlations between baseline and outcome measures, 2017/18 and 2018/19

Measure	Number of students	Correlation
Real-word decomposition[a]	2,173	.65
Nonword derivation[a]	2,173	.55
Inferencing of word meanings[a]	2,173	.61
i-Ready Vocabulary	2,208	.74
Florida Standards Assessment English Language Arts	2,075	.81

Note: All correlations were significant at $p < .001$.
a. The i-Ready Reading test was used as the baseline measure.
Source: Authors' analysis of school district data for 2017/18 and 2018/19.

The i-Ready measures. The i-Ready Reading is a computer-adaptive multiple-choice assessment for grades K–12 that is administered three times a year: in the fall (late August), winter (early December), and spring (early April; Curriculum Associates LLC, 2018). The i-Ready Reading composite score comprises the Vocabulary, Informational Text Comprehension, and Literacy Text Comprehension subtests. Test-retest reliability for the i-Ready Reading is .86 and marginal reliability is .97 (Curriculum Associates LLC, 2018). The i-Ready Vocabulary subtest assesses students on academic and domain-specific vocabulary, word relationships, word-learning strategies, prefixes,

---

[1] i-Ready Reading posttest scores were not used as a long-term outcome because i-Ready Reading was highly correlated with the FSA–ELA (correlation of .83) and because the district was most interested in evaluating the impact of WKI on FSA–ELA.

suffixes, word roots, and use of reference materials. Marginal reliability for the i-Ready Vocabulary subtest is .89 (Curriculum Associates LLC, 2018).

Florida Standards Assessments English Language Arts (FSA–ELA). The FSA–ELA is the annual standards-based, criterion-referenced English language arts assessment that is used to measure student proficiency on the state's ELA standards. Scores on the FSA–ELA are reported as a developmental scale score and range from 240 to 412. Marginal reliability is .88 for the grade 4 FSA–ELA and .89 for grade 5 (Florida Department of Education, 2018).

Correlations among all study measures are reported in table A3.

Table A3. Correlations among all study measures, 2017/18 and 2018/19

Measure	Pretest			Outcome measure					
	i-Ready Vocabulary	i-Ready Reading	FSA–ELA	i-Ready Vocabulary	i-Ready Reading	FSA–ELA	Real word decom-position	Nonword derivation	Inferencing of word meanings
Pretest									
i-Ready Vocabulary	1								
i-Ready Reading	.90	1							
FSA–ELA	.72	.83	1						
Outcome									
i-Ready Vocabulary	.74	.77	.69	1					
i-Ready Reading	.77	.85	.79	.91	1				
FSA–ELA	.72	.82	.81	.73	.83	1			
Real-word decomposition	.60	.65	.63	.59	.64	.64	1		
Nonword derivation	.49	.55	.55	.50	.56	.57	.61	1	
Inferencing of word meanings	.57	.61	.57	.61	.64	.61	.53	.51	1

FSA–ELA is Florida Standards Assessment English Language Arts.
Note: All correlations were significant at $p < .001$.
Source: Authors' analysis of school district data for 2017/18 and 2018/19.

Sample

The study took place in a large school district in central Florida. Recruitment of schools and teachers into the study started in spring 2018 through a virtual recruitment meeting presenting the study to 46 principals. Principals were invited if at least 60 percent of students in their elementary school were eligible for the national school lunch program (an indicator of poverty) and the school had at least two grade 5 ELA teachers. Following this recruitment meeting, 42 principals agreed to have their school participate. The principals provided the names of ELA teachers in their schools ($n = 104$) who agreed to be randomly assigned during the 2018/19 school year either to implement WKI instruction within the standard ELA instruction block (treatment group) or to teach only the standard ELA instruction block (business-as-usual control group; figure A1).

In late spring 2018 the study team used Microsoft Excel to randomly assign the participating ELA teachers within each school (school served as the blocking variable) to either implement WKI instruction within the ELA instruction block or continue with the business-as-usual ELA instruction block only. Specifically, each ELA teacher within each school was assigned a random number. ELA teachers were then sorted in descending order based on the assigned random number within each school; the first half were assigned to WKI and the remaining to the business-as-usual control. In the 12 schools with an odd number of participating ELA teachers, the treatment group was randomly assigned one more teacher than the control group.

The random assignment process resulted in 58 treatment teachers (serving 1,967 grade 5 students) and 46 control teachers (serving 1,551 grade 5 students). Eight teachers serving 178 students (5 teachers serving 110 students in the treatment group and 3 teachers serving 68 students in the control group) were originally classified as ELA

teachers in departmentalized schools[2] but were reclassified as math/science teachers after random assignment. Therefore, these teachers were not considered eligible to participate in this study (see figure A1). This reclassification also resulted in the loss of one school. An additional ELA teacher (serving 22 students) who was allocated to the treatment group chose not to participate for health reasons and was removed from the analyses. Of the remaining 3,318 students (representing 95 ELA teachers across 41 schools), 824 did not have parental consent to participate (335 in the treatment group and 489 in the control group) and 207 students withdrew from participating schools (131 in the treatment group and 76 in the control group). Two additional schools (3 teachers and 73 students in the treatment group) were removed from the analyses because of reclassification of teachers from the control group. All of the remaining 39 schools in the analytic sample had both treatment and control teachers represented. Overall attrition for teachers was 4.2 percent,[3] with differential attrition of 7.5 percent (table A4). Overall individual nonresponse rates for students ranged from 37.1 percent to 41 percent, and differential nonresponse rates for students ranged from 6.2 percent to 7.4 percent.

Table A4. Sample attrition from the Word Knowledge Instruction and control group, 2018/19

Sample	Treatment group (Word Knowledge Instruction)			Control group			Teacher-level attrition (%)		Student-level nonresponse (%)	
	Schools	Teachers	Students	Schools	Teachers	Students	Overall	Differential	Overall	Differential
Assigned	42	58	1,967	42	46	1,551	na	na	na	na
Assigned and eligible	42	53	1,857	42	43	1,483	na	na	na	na
Overall analytic[a]	39	49	1,296	39	43	918	4.2	7.5	37.1	6.7
Real-word decomposition	39	49	1,279	39	43	894	4.2	7.5	38.2	7.4
Nonword derivation	39	49	1,279	39	43	894	4.2	7.5	38.2	7.4
Inferencing of word meanings	39	49	1,279	39	43	894	4.2	7.5	38.2	7.4
i-Ready Vocabulary	39	49	1,294	39	43	914	4.2	7.5	37.2	6.9
FSA–ELA	39	49	1,214	39	43	861	4.2	7.5	41.0	6.2

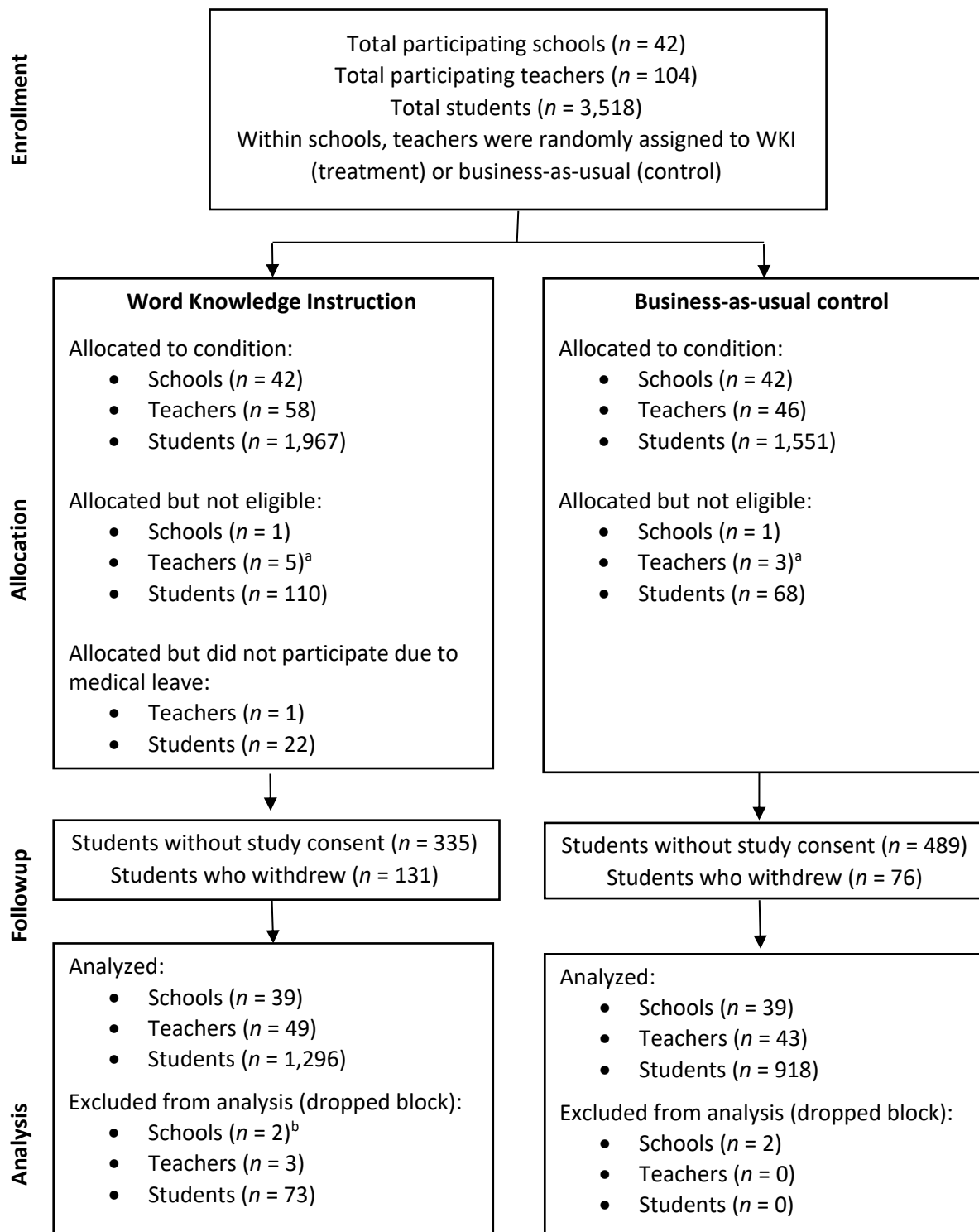FSA–ELA is Florida Standards Assessment English Language Arts.
na is not applicable.
a. Students in the overall analytic sample have scores on at least one of the outcome measures. All 39 schools in the analytic sample had both treatment and control group teachers represented.
Source: Authors' analysis of school district data for 2017/18 and 2018/19.

---

[2] Ninety percent of participating schools used a departmentalized approach to instruction, with one teacher responsible for ELA instruction and another responsible for math/science instruction within each grade.
[3] Teachers who were reclassified from ELA teachers to non-ELA teachers were not included in the teacher-level attrition calculations (What Works Clearinghouse, 2020b).

Figure A1. Consolidated standards of reporting trials diagram for a study on the impact of Word Knowledge Instruction for grade 5 students, 2018/19

Enrollment

Total participating schools (*n* = 42)
Total participating teachers (*n* = 104)
Total students (*n* = 3,518)
Within schools, teachers were randomly assigned to WKI (treatment) or business-as-usual (control)

Allocation

Word Knowledge Instruction

Allocated to condition:
- Schools (*n* = 42)
- Teachers (*n* = 58)
- Students (*n* = 1,967)

Allocated but not eligible:
- Schools (*n* = 1)
- Teachers (*n* = 5)[a]
- Students (*n* = 110)

Allocated but did not participate due to medical leave:
- Teachers (*n* = 1)
- Students (*n* = 22)

Business-as-usual control

Allocated to condition:
- Schools (*n* = 42)
- Teachers (*n* = 46)
- Students (*n* = 1,551)

Allocated but not eligible:
- Schools (*n* = 1)
- Teachers (*n* = 3)[a]
- Students (*n* = 68)

Followup

Students without study consent (*n* = 335)
Students who withdrew (*n* = 131)

Students without study consent (*n* = 489)
Students who withdrew (*n* = 76)

Analysis

Analyzed:
- Schools (*n* = 39)
- Teachers (*n* = 49)
- Students (*n* = 1,296)

Excluded from analysis (dropped block):
- Schools (*n* = 2)[b]
- Teachers (*n* = 3)
- Students (*n* = 73)

Analyzed:
- Schools (*n* = 39)
- Teachers (*n* = 43)
- Students (*n* = 918)

Excluded from analysis (dropped block):
- Schools (*n* = 2)
- Teachers (*n* = 0)
- Students (*n* = 0)

a. These English language arts teachers were reclassified to math/science instruction and were no longer eligible to participate in this study.

b. These schools were removed from analyses because control group teachers were reclassified from English language arts teachers to math/science teachers.

Source: Authors compilation.

Following What Works Clearinghouse (WWC) standards, the study team assessed baseline equivalence on achievement and demographic characteristic variables for students in the analytic sample: 1,296 in treatment and 918 in control (What Works Clearinghouse, 2020b). Tables A5 and A6 present descriptive data on these variables for students in treatment and control groups, as well as effect size differences between these groups. To assess baseline equivalence between students in treatment and control groups, frequencies and means for each group were entered into a WWC study review guide spreadsheet to derive effect size differences. The absolute value of effect size differences on baseline variables ranged from 0.06 to 0.10 (see table A5). According to WWC, effect size differences greater than 0.05 and less than 0.25 fall within the "adjustable range," meaning that differences warrant the inclusion of the baseline variables in the analytic model. Following WWC standards, baseline variables were included in all analytic models, and therefore all outcomes are likely to meet WWC standards with reservations. Effect size differences on demographic variables were in the acceptable range—the absolute value of the effect size difference was smaller than 0.05.

Table A5. Baseline equivalence for analytic sample on achievement variables, 2017/18 and 2018/19

	Treatment group students (Word Knowledge Instruction)			Control group students			
Achievement variable	Number	Mean	Standard deviation	Number	Mean	Standard deviation	Effect size[a]
i-Ready Reading[b]	1,279	551.70	45.60	894	555.54	46.72	−0.08
i-Ready Vocabulary	1,294	546.14	47.52	914	549.10	49.10	−0.06
FSA–ELA	1,214	306.52	19.54	861	308.45	19.89	−0.10

FSA–ELA is Florida Standards Assessment English Language Arts.
Note: The sample included 49 teachers in the Word Knowledge Instruction group and 43 teachers in the control group, representing 39 schools.
a. Following What Works Clearinghouse procedures, Hedges' g was used to calculate the effect sizes. See What Works Clearinghouse (2020a) for more information.
b. i-Ready Reading was used as the pretest baseline measure for the three short-term outcome measures (real-word decomposition, nonword derivation, and inferencing of word meanings).
Source: Authors' analysis of school district data for 2017/18 and 2018/19.

Table A6. Baseline equivalence for analytic sample on student demographic characteristics, 2018/19

	Treatment group students Word Knowledge Instruction ($n = 1,296$)		Control group students ($n = 918$)		Effect size[a]	Total ($n = 2,214$)	
Student demographic characteristic	Mean	Standard deviation	Mean	Standard deviation		Mean	Standard deviation
Eligible for national school lunch program	0.85	0.35	0.86	0.35	−0.05	0.85	0.35
English learner student	0.20	0.40	0.20	0.40	0	0.20	0.40
Asian	0.02	0.14	0.02	0.14	0	0.02	0.14
African American	0.21	0.41	0.22	0.41	−0.04	0.21	0.41
Hispanic	0.54	0.50	0.54	0.50	0	0.54	0.50
White, non-Hispanic	0.19	0.39	0.18	0.39	0.04	0.18	0.39
Other	0.04	0.21	0.04	0.19	0	0.04	0.20

Note: The analytic sample included 49 teachers in the Word Knowledge Instruction group and 43 teachers in the control group, representing 39 schools.
a. Following What Works Clearinghouse procedures, a Cox index was used to calculate the effect sizes for dichotomous variables. See What Works Clearinghouse (2020a) for more information.
Source: Authors' analysis of school district data for 2018/19.

Intervention

WKI consists of 15-minute lessons taught 4 days a week for 20 weeks as part of grade 5 ELA instruction. Teachers began implementing WKI lessons on September 10, 2018, just after i-Ready Reading pretests were administered, including the i-Ready Vocabulary subtest. WKI focuses on 20 prefixes and suffixes, which together are called

affixes. The intervention includes words with suffixes that result in derived adjectives (for example, *achieve* + *able* = *achievable*) and derived nouns (for example, *equip* + *ment* = *equipment*), as well as prefixes (*re-*, *trans-*, *con-*). Instruction highlights the phonological (*electric* + *ity* = *electricity*), orthographic (*decide* + *sion* = *decision*), and phonological and orthographic (*theory* + *al* = *theoretical*) shifts that occur when affixes connect to base words. Instruction also highlights Spanish and English cognates and activities using connective words. WKI uses such evidence-based practices as frequent exposure to words containing target affixes with repetition and explicit instruction; active practice composing sentences using targets within meaningful linguistic contexts; links to existing knowledge by pairing targets with synonyms, expansions, and associations and contrasting targets to antonyms to enhance breadth and depth of vocabulary; active practice in authentic contexts by defining the meaning of target words in connected text; visuals and manipulatives to show derivational conversions and depict word meanings; and regular reviews with checks for understanding (Bowers et al., 2010; Kieffer & Lesaux, 2007, 2010; Lesaux et al., 2014).

WKI instruction is intended to cover only 15 minutes of the 120-minute daily ELA instructional block four days a week. For the remaining 105 minutes of the ELA block and on the fifth day, teachers used the district's state-adopted core reading program, other district approved instructional materials, and other teacher-selected supplemental materials. Therefore, it is likely that students of teachers in the WKI treatment group also received instruction focused on word parts outside of the WKI lessons. In other words, the WKI curriculum uses a small fraction of the total time allocated to ELA instruction, so the treatment group might have received most of the same instruction that the control group received on affixes plus the extra instruction provided by WKI.

Training. WKI developers provided in-person training during the summer of 2018 and ongoing support via a website. Specifically, teachers assigned to the treatment group were trained over a two-day period in July 2018 for approximately 6 hours a day and were sent home with a teacher manual and a set of student workbooks to further familiarize themselves with the WKI strategies and materials. Teachers received an honorarium of $25 an hour to attend the summer training. Additionally, all teachers were enrolled in a CANVAS website so that they could access the training PowerPoint presentations, videos of teachers implementing WKI lessons, and additional resources. WKI teachers unable to attend summer training (five ELA teachers) were required to pass quizzes covering the summer training materials on the website. If they struggled with the quizzes, they received help until they passed. Class sets of student materials were mailed directly to each school in August 2018.

Once WKI started in September 2018, developers supported teachers in multiple ways on an as-needed basis. Teachers were encouraged to request support from the developers through email or in person during a scheduled classroom observation. If classroom observations indicated that teachers were struggling to implement WKI lessons with fidelity, developers worked with them to improve implementation. On a monthly basis developers visited several treatment classrooms to observe instruction and address teachers' questions or concerns. Any teacher concerns addressed to the study team were referred to the developers.

Business-as-usual control

In general, teachers in the business-as-usual control group conducted 120 minutes of ELA instruction daily, using the district's state-adopted core reading program, other district-approved instructional materials, and other teacher-selected supplemental materials. The amount of instruction focused on understanding word parts varied across these instructional materials. The district's core reading program (HMH's *Journeys*) included instruction on 21 affixes in the vocabulary component and 6 additional affixes in the spelling component (9 of these affixes overlap with WKI). In addition, the district made a supplementary *Language and Literacy* professional book available in its teacher portal that included instruction on an additional 15 affixes (7 of these affixes overlap with WKI). Collectively, students in the business-as-usual control group were potentially exposed to 42 affixes, including 16 affixes that also appeared in the WKI lessons. However, WKI included explicit instruction in morphological

awareness in four lessons on each affix, whereas the state-adopted core reading program included only one lesson on an affix or group of affixes and emphasized vocabulary building.

Classroom observations by the study team

The study team conducted two classroom observations (fall 2018 and spring 2019) during the 20-week intervention in both treatment and control ELA classrooms using the same observation form. The observations captured the entire duration of the ELA block and served four purposes:

- Quantify fidelity within WKI classrooms, including adherence to lesson sequence and script and quality of implementation.

- Describe any instruction on affixes beyond the 15-minute lessons in the WKI classrooms.

- Determine potential contamination within the ELA control classrooms.

- Describe instruction in the ELA control classrooms.

The categories describing ELA instruction were derived from the Florida State Standards. Observers were instructed to code the start time of each instructional and transitional activity. For each instructional activity, observers identified the general focus of ELA instruction (for example, comprehension, fluency, and vocabulary) and noted any instances of instruction relevant to this study (for example, prefix/suffix, base/root, parts of speech, or connectives). During WKI lessons observers noted the specific lesson being taught, adherence to the lesson sequence and script, quality of implementation, and lesson duration.

Observers were trained over a two-day period in August 2018 to use the observation tool. Observers were trained to achieve better than 80 percent reliability on the tool before conducting classroom observations. Inter-rater reliability was monitored during the observation windows by randomly selecting 40 percent of observations to be coded by two observers. Agreement between pairs of observers averaged 93 percent (standard deviation of 6 percent) and ranged from 70 percent[4] to 100 percent. Following each inter-rater observation, the pair of observers reconciled any disagreements and created a single consensus observation form for that ELA teacher.

Word Knowledge Instruction. Fidelity ratings for the WKI portion of each observation (adherence to lesson sequence and script, quality of instruction, and lesson duration) were averaged to create overall fidelity ratings. Table A7 reports the mean fidelity ratings across all teachers implementing WKI. The three components of fidelity showed adequate implementation: adherence to lesson sequence and script (mean = 85 percent, standard deviation = 15 percent); quality of instruction (mean = 3 on a 1–5 scale, standard deviation = 0.46); and lesson duration in minutes (mean = 21.32, standard deviation = 5.03). Program coverage was also adequate based on completion of WKI activities in student workbooks. Forty-seven of the 49 WKI teachers covered all 20 affixes. Two WKI teachers covered only 15 of the 20 affixes: one teacher went on leave mid-year and no one took over WKI in the class, and the other teacher moved through the program more slowly.

---

[4] For one pair of observers agreement was 70 percent. This occurred only because very few instructional activities were coded, and one observer failed to code a 10-minute segment of time that the other observer had coded. The next lowest percent agreement was 80 percent.

Table A7. Descriptive data for the components of fidelity for teachers implementing Word Knowledge Instruction, 2018/19

Component of fidelity	Mean	Standard deviation	Minimum	Maximum
Adherence to lesson sequence and script (percent)	85	15	45	100
Quality of Word Knowledge Instruction lessons (1–5 scale)	3	0.46	1.08	3.89
Duration of Word Knowledge Instruction lessons (minutes)	21.32	5.03	12	35.5

Note: *n* = 49.
Source: Authors' analysis of classroom observation data.

Business-as-usual control. Classroom observations confirmed that none of the business-as-usual control teachers used any WKI materials during the ELA instructional block.

Analysis

A three-level hierarchical linear model (HLM) with students nested in teachers and teachers nested in schools was used to estimate treatment effects. The HLM accounts for student and teacher sources of variability in the outcomes. Because teachers within each school were randomly assigned to a treatment or control group, each block or school can be viewed as a mini-experiment. Therefore, school was modeled as Level 3 in the HLM, and heterogeneity in the treatment effect across schools was modeled as a random effect. This three-level specification enables exploring whether the treatment effect varied across the 39 schools in the analytic sample, and if so, by how much.

For research questions 1 and 2, the following three-level multilevel model was used:

Level 1 (student)

$$Y_{ijk} = \pi_{0jk} + \pi_{1jk}(Pretest)_{ijk} + \pi_{2jk}(NSLP)_{ijk} + \pi_{3jk}(EL)_{ijk} + e_{ijk}$$

Level 2 (teacher)

$$\pi_{0jk} = \beta_{00k} + \beta_{01k}(Treatment)_{jk} + \beta_{02k}(Average\ student\ pretest)_{jk} + r_{0jk}$$
$$\pi_{1jk} = \beta_{10k}$$
$$\pi_{2jk} = \beta_{20k}$$
$$\pi_{3jk} = \beta_{30k}$$

Level 3 (school)

$$\beta_{00k} = \gamma_{000} + u_{00k}$$
$$\beta_{01k} = \gamma_{010} + u_{01k}$$
$$\beta_{02k} = \gamma_{020}$$
$$\beta_{10k} = \gamma_{100}$$
$$\beta_{20k} = \gamma_{200}$$
$$\beta_{30k} = \gamma_{300}$$

where Y_{ijk} represents the outcome score for student *i* in ELA teacher *j*'s class in school *k*. The outcomes used to address each research question are reported in table A8.

Table A8. Pretest covariates for each outcome, 2017/18 and 2018/19

Outcome measure	i-Ready Vocabulary	i-Ready Reading	Florida Standards Assessment English Language Arts
Real-word decomposition		X	
Nonword derivation		X	
Inferencing of word meanings		X	
i-Ready Vocabulary	X		
Florida Standards Assessment English Language Arts			X

Source: Authors compilation.

In the Level 1 model the student outcome score is modeled as a function of the following student characteristics: pretest score (see table A8), eligibility for the national school lunch program (NSLP), and English learner (EL) student status, where a value of 1 indicates program eligibility or English learner student status and a value of 0 indicates lack of eligibility or status. Student characteristics, including pretests and demographic characteristic variables, were grand mean centered so that π_{0jk} represents the adjusted mean outcome score for ELA teacher j in school k. Lastly, e_{ijk} represents the random student effect that is assumed to be normally distributed with a mean of 0 and constant variance σ^2.

In the Level 2 model the adjusted mean outcome score π_{0jk} for ELA teacher j in school k is modeled as varying randomly across ELA teachers. The coefficient of the treatment indicator β_{01k} (ELA teachers implementing WKI have a score of 1 and ELA teachers in the control group have a score of 0) is the key parameter of interest and represents the expected difference in outcome scores between the WKI treatment condition and the business-as-usual control condition in school k, with other covariates controlled for in the model. Student pretest scores were aggregated by ELA teacher and included as a covariate in this model (grand mean centered). Lastly, r_{0jk} is the random teacher effect, which represents the deviation of ELA teacher j's classroom in school k and is assumed to be normally distributed with a mean of 0 and variance $\tau_{\pi00}$.

In the Level 3 model the average outcome β_{00k} and the treatment effect in each school β_{01k} are modeled as random effects. In the equation for the average outcome in each school, the parameter γ_{000} represents the mean outcome score for the population of schools, and u_{00k} is the random school effect that represents the deviation of school k's mean from the grand mean and is assumed to be normally distributed with a mean of 0 and variance $\tau_{\beta00}$. In the equation for the treatment effect in each school, the parameter γ_{010} represents the treatment effect for the population of schools after covariates in the model are controlled for. Lastly, u_{01k} is a random effect that is assumed to be normally distributed with a mean of 0 and variance $\tau_{\beta11}$. This random effect represents the variance in the treatment effect across schools.

All models were estimated using HLM software version 7.03. Before any models were estimated, an unconditional model was estimated for each outcome to calculate the proportion of variance in an outcome that is accounted for by differences between students, between teachers, and between schools for each level modeled in the three-level HLM (table A9).

Table A9. Fraction of variation by level for each outcome measure, 2017/18 and 2018/19

Measure	School	Teacher	Student
Real-word decomposition	0.05	0.07	0.88
Nonword derivation	0.05	0.03	0.92
Inferencing of word meanings	0.03	0.07	0.89
i-Ready Vocabulary	0.03	0.06	0.92
Florida Standards Assessment English Language Arts	0.01	0.07	0.92

Note: The fraction of variation at the school level is the school intraclass correlation coefficient. The numbers do not sum to 1.00 due to rounding.
Source: Authors' analysis of school district data 2017/18 and 2018/19.

Effect sizes were computed for all treatment effects using Hedges' *g* formula:

$$g = \frac{\hat{\beta}_{01}}{\sqrt{\dfrac{(n_t - 1)\hat{S}_t^2 + (n_c - 1)\hat{S}_c^2}{(n_t + n_c - 2)}}}$$

where $\hat{\beta}_{01}$ is the estimated treatment effect obtained from the three-level impact model, n_t is the number of students in the treatment group, n_c is the number of students in the control group, $\hat{S}_t^2$ is the outcome unadjusted student-level standard deviation for the treatment group, and $\hat{S}_c^2$ is the outcome unadjusted student-level standard deviation for the control group.

Results from the multilevel models for real-word decomposition, nonword derivation, and inferencing of word meanings outcomes (research question 1) are reported in table A10. Results from the multilevel models for i-Ready Vocabulary and FSA–ELA (research question 2) are reported in table A11. Sample sizes, unadjusted means and standard deviations, adjusted means, and effect size estimates by outcome for research questions 1 and 2 are reported in table A12.

Table A10. Impact of Word Knowledge Instruction on researcher-developed, short-term outcome measures, 2018/19

Model	Real-word decomposition			Nonword derivation			Inferencing of word meanings		
	Coefficient	Standard error	p-value	Coefficient	Standard error	p-value	Coefficient	Standard error	p-value
Fixed effects									
Intercept	11.71	0.16	<.001	8.81	0.13	<.001	7.79	0.09	<.001
Teacher-level covariates									
Word Knowledge Instruction	1.05	0.17	<.001	0.18	0.14	.221	−0.05	0.13	.675
Average student pretest	0.02	0.01	.006	0.01	0.01	.022	0.01	0.00	.023
Student-level covariates									
Pretest	0.06	0.00	<.001	0.04	0.00	<.001	0.03	0.00	<.001
Eligibility for national school lunch program	−0.58	0.21	.006	−0.60	0.19	.002	−0.52	0.13	<.001
English learner student status	−0.43	0.21	.038	−0.64	0.18	<.001	−0.23	0.12	.054
Random effect	Variance	χ^2 (df)	p-value	Variance	χ^2 (df)	p-value	Variance	χ^2 (df)	p-value
Level 1	10.77			8.94			3.68		
Level 2	0.00	13.95 (13)	.377	0.00	8.70 (13)	>.500	0.16	25.28 (13)	.021
Level 3	0.51	85.14 (38)	<.001	0.18	52.19 (38)	.062	0.02	37.89 (38)	>.500
Word Knowledge Instruction achievement effect	0.27	55.41 (38)	.034	0.08	41.30 (38)	.328	0.06	44.69 (38)	.211
Number of students	2,173			2,173			2,173		
Deviance	11,391.59			10,962.63			9,079.91		
Number of parameters	11			11			11		

Note: The sample included 49 teachers in the Word Knowledge Instruction group and 43 teachers in the control group, representing 39 schools. The standard deviations of the outcomes are much smaller than the standard deviations of the pretests (see tables A12 and A5).
Source: Authors' analysis of school district data for 2018/19.

Table A11. Impact of Word Knowledge Instruction on long-term outcome measures, 2017/18 and 2018/19

Model	i-Ready Vocabulary			Florida Standards Assessment English Language Arts		
	Coefficient	Standard error	*p*-value	Coefficient	Standard error	*p*-value
Fixed effects						
Intercept	566.92	1.28	<.001	316.88	0.60	<.001
Teacher-level covariates						
Word Knowledge Instruction	−0.71	1.81	.696	0.44	0.80	.591
Average student pretest	0.17	0.06	.012	0.10	0.06	.132
Student-level covariates						
Pretest	0.68	0.02	<.001	0.84	0.02	<.001
Eligibility for national school lunch program	−5.60	2.05	.006	−1.58	0.81	.050
English learner student status	−7.91	1.92	<.001	−1.63	0.76	.033
Random effects	Variance	χ^2 (df)	*p*-value	Variance	χ^2 (df)	*p*-value
Level 1	1,003.96			143.63		
Level 2	16.54	11.98 (13)	>.500	5.87	32.62 (13)	.002
Level 3	1.46	42.30 (38)	.290	0.84	42.38 (38)	.287
Word Knowledge Instruction achievement effect	18.22	51.61 (38)	.069	2.05	46.45 (38)	.163
Number of students	2,208			2,075		
Deviance	21,576.32			16,271.61		
Number of parameters	11			11		

Note: The sample included 49 teachers in the Word Knowledge Instruction group and 43 teachers in the control group, representing 39 schools. The standard deviations of the outcomes and pretests are similar (see tables A12 and A5).
Source: Authors' analysis of school district data for 2017/18 and 2018/19.

Table A12. Posttest scores for students receiving Word Knowledge Instruction and students in the business-as-usual control group, 2017/18 and 2018/19

Measure	Word Knowledge Instruction				Business-as-usual control				Effect size[a]
	Number of students	Unadjusted mean	Adjusted mean	Standard deviation	Number of students	Unadjusted mean	Adjusted mean	Standard deviation	
Short-term outcomes									
Real-word decomposition	1,279	12.63	12.76	4.39	894	11.93	11.71	4.66	0.23
Nonword derivation	1,279	8.98	9.62	3.67	894	8.97	8.81	3.69	0.05
Inferencing of word meanings	1,279	7.71	7.74	2.47	894	7.93	7.79	2.60	−0.02
Long-term outcomes									
i-Ready Vocabulary	1,294	565.79	566.21	47.68	914	568.79	566.92	48.82	−0.01
FSA–ELA	1,214	316.90	317.32	20.78	861	318.05	316.88	21.31	0.02

FSA–ELA is Florida Standards Assessment English Language Arts.

Note: The adjusted mean represents the average posttest score after controlling for student prior achievement and demographic characteristics. The adjusted mean for the business-as-usual control group is equal to the coefficient for the intercept, and the adjusted mean for the Word Knowledge Instruction group is equal to the sum of the coefficients for the intercept and Word Knowledge Instruction (see tables A10 and A11). Total points possible are 20 for real-word decomposition, 16 for nonword derivation, and 15 for inferencing of word meanings. The sample included 49 teachers in the Word Knowledge Instruction group and 43 teachers in the control group, representing 39 schools.

a. Following What Works Clearinghouse procedures, Hedges' *g* was used to calculate the effect sizes. See What Works Clearinghouse (2020a) for more information.

Source: Authors' analysis of school district data for 2017/18 and 2018/19.

The Benjamini-Hochberg Linear Step Up procedure was used to control for the false discovery rate for the significant treatment effect found on real-word decomposition (Benjamini & Hochberg, 1995). The critical *p*-value

was calculated using five as the total number of comparisons estimated. The effect of WKI on real-word decomposition remained significant after multiple comparisons were controlled for (table A13).

Table A13. Benjamini-Hochberg linear step-up procedure applied to the significant treatment effects, 2018/19

Outcomes	Model p-value	Rank	Total effects	Critical p-value	Significant after correction
Real-word decomposition	<.001	1	5	.01	Yes
Nonword derivation	.22	2	5	.02	No
Inferencing of word meanings	.68	4	5	.04	No
i-Ready Vocabulary	.70	5	5	.05	No
Florida Standards Assessment English Language Arts	.59	3	5	.03	No

Source: Authors' analysis of school district data for 2018/19.

Exploratory subgroup analyses

To answer research question 3, treatment was added at Level 2 to the Level 1 English learner student status equation (π_{3jk}) to explore whether WKI instruction had a differential effect on outcome performance by English learner student status (represented by the EL x treatment interaction). The following three-level multilevel model was used:

Level 1 (student)

$$Y_{ijk} = \pi_{0jk} + \pi_{1jk}(Pretest)_{ijk} + \pi_{2jk}(NSLP)_{ijk} + \pi_{3jk}(EL)_{ijk} + e_{ijk}$$

Level 2 (teacher)

$$\pi_{0jk} = \beta_{00k} + \beta_{01k}(Treatment)_{jk} + \beta_{02k}(Average\ student\ pretest)_{jk} + r_{0jk}$$

$$\pi_{1jk} = \beta_{10k}$$

$$\pi_{2jk} = \beta_{20k}$$

$$\pi_{3jk} = \beta_{30k} + \beta_{31k}(Treatment)_{jk}$$

Level 3 (school)

$$\beta_{00k} = \gamma_{000} + u_{00k}$$

$$\beta_{01k} = \gamma_{010} + u_{01k}$$

$$\beta_{02k} = \gamma_{020}$$

$$\beta_{10k} = \gamma_{100}$$

$$\beta_{20k} = \gamma_{200}$$

$$\beta_{30k} = \gamma_{300}$$

$$\beta_{31k} = \gamma_{310}$$

In these models the key parameter of interest is γ_{310} and represents the English learner student status by treatment interaction. Results from the multilevel models for real-word decomposition, nonword derivation, and the inferencing of word meanings outcomes are reported in table A14. Results from the multilevel models for

i-Ready Vocabulary and FSA–ELA are reported in table A15. Unadjusted means and standard deviations and adjusted means by outcome for research question 3 are reported in table A16. Baseline equivalence on all outcome measures by English learner student status is reported in table A17.

Table A14. Differential impacts of Word Knowledge Instruction, by English learner student status for researcher-developed, short-term outcomes, 2018/19

Model	Real-word decomposition			Nonword derivation			Inferencing of word meanings		
	Coefficient	Standard error	*p*-value	Coefficient	Standard error	*p*-value	Coefficient	Standard error	*p*-value
Fixed effects									
Intercept	11.83	0.18	<.001	8.92	0.14	<.001	7.84	0.10	<.001
Teacher-level covariates									
Word Knowledge Instruction	0.99	0.19	<.001	0.22	0.16	.176	−0.07	0.14	.619
Average student pretest	0.02	0.01	.007	0.01	0.01	.020	0.01	0.00	.025
Student-level covariates									
Pretest	0.06	0.00	<.001	0.04	0.00	<.001	0.03	0.00	<.001
Eligibility for national school lunch program	−0.58	0.21	.006	−0.60	0.19	.002	−0.52	0.13	<.001
English learner student status	−0.60	0.30	.048	−0.52	0.27	.058	−0.27	0.18	.128
English learner student status x Word Knowledge Instruction	0.29	0.37	.434	−0.20	0.33	.556	0.07	0.22	.756
Random effects	Variance	χ^2 (df)	*p*-value	Variance	χ^2 (df)	*p*-value	Variance	χ^2 (df)	*p*-value
Level 1	10.77			8.94			3.68		
Level 2	0.00	13.73 (13)	.393	0.00	8.69 (13)	>.500	0.16	25.18 (13)	.022
Level 3	0.50	84.51 (38)	<.001	0.19	52.31 (38)	.061	0.02	37.93 (38)	>.500
Word Knowledge Instruction achievement effect	0.29	56.23 (38)	.028	0.08	41.25 (38)	.330	0.07	44.91 (38)	.205
Number of students	2,173			2,173			2,173		
Deviance	11,390.99			10,962.30			9,079.82		
Number of parameters	12			12			12		

Note: The sample included 49 teachers in the Word Knowledge Instruction group and 43 teachers in the control group, representing 39 schools. The standard deviations of the outcomes are much smaller than the standard deviations of the pretests (see tables A16 and A17).
Source: Authors' analysis of school district data for 2018/19.

Table A15. Differential impacts of Word Knowledge Instruction, by English learner student status for long-term outcomes, 2017/18 and 2018/19

Model	i-Ready Vocabulary			Florida Standards Assessment English Language Arts		
	Coefficient	Standard error	*p*-value	Coefficient	Standard error	*p*-value
Fixed effects						
Intercept	568.79	1.41	<.001	317.23	0.64	<.001
Teacher-level covariates						
Word Knowledge Instruction	−1.16	1.95	.555	0.38	0.85	.658
Average student pretest	0.17	0.06	.015	0.10	0.06	.137
Student-level covariates						
Pretest	0.68	0.02	<.001	0.84	0.02	<.001
Eligibility for national school lunch program	−5.58	2.05	.006	−1.58	0.81	.050
English learner student status	−9.22	2.86	.001	−1.80	1.15	.116
English learner student status x Word Knowledge Instruction	2.20	3.55	.536	0.29	1.42	.839
Random effects	Variance	χ^2 (df)	*p*–value	Variance	χ^2 (df)	*p*–value
Level 1	1,004.01			143.64		
Level 2	15.83	11.80 (13)	>.500	5.84	32.59 (13)	.002
Level 3	1.66	42.66 (38)	.277	0.85	42.43 (38)	.286
Word Knowledge Instruction achievement effect	18.72	52.01 (38)	.064	2.03	46.49 (38)	.162
Number of students	2,208			2,075		
Deviance	21,575.94			16,271.57		
Number of parameters	12			12		

Note: The sample included 49 teachers in the Word Knowledge Instruction group and 43 teachers in the business-as-usual control group, representing 39 schools. The standard deviations of the outcomes and pretests are similar (see tables A16 and A17).
Source: Authors' analysis of school district data for 2017/18 and 2018/19.

Table A16. Posttest scores of English learner and non–English learner students receiving Word Knowledge Instruction and those in the business-as-usual control group, 2017/18 and 2018/19

Variable	English learner students						Non–English learner students					
	Word Knowledge Instruction			Business-as-usual control			Word Knowledge Instruction			Business-as-usual control		
	Mean	Adjusted mean	SD	Mean	Adjusted mean	SD	Mean	Adjusted mean	SD	Mean	Adjusted mean	SD
Real-word decomposition	9.92	12.51	4.30	8.31	11.23	4.05	13.29	12.82	4.15	12.83	11.83	4.35
Nonword derivation	6.82	8.42	2.94	6.47	8.40	2.90	9.51	9.14	3.63	9.60	8.92	3.61
Inferencing of word meanings	6.36	7.57	2.06	6.09	7.57	2.07	8.05	7.77	2.45	8.38	7.84	2.52
i-Ready vocabulary	533.87	560.61	52.81	527.53	559.57	47.93	573.74	567.63	42.79	579.26	568.79	43.20
FSA–ELA	303.73	316.10	18.13	299.78	315.43	17.59	320.03	317.61	20.15	322.41	317.23	19.75

SD is standard deviation. FSA–ELA is Florida Standards Assessment English Language Arts.
Note: The adjusted mean for non–English learner students in the business-as-usual control group is equal to the coefficient for the intercept, the adjusted mean for English learner students in the control group is equal to the sum of the coefficients for the intercept and English learner student status, the adjusted mean for non–English learner students in Word Knowledge Instruction is equal to the sum of the coefficients for the intercept and Word Knowledge Instruction, and the adjusted mean for English learner students in Word Knowledge Instruction is equal to the sum of the coefficients for the intercept, English learner student status, Word Knowledge Instruction, and the interaction (see tables A14 and A15). The sample included 49 teachers in the Word Knowledge Instruction group and 43 teachers in the control group, representing 39 schools.
Source: Authors' analysis of school district data for 2017/18 and 2018/19.

Table A17. Baseline equivalence for the analytic sample of English learner and non–English learner students by achievement variable, 2017/18 and 2018/19

Variable	English learner students							Non–English learner students						
	Word Knowledge Instruction			Business-as-usual control			Effect size[a]	Word Knowledge Instruction			Business-as-usual control			Effect size[a]
	Number of students	Mean	SD	Number of students	Mean	SD		Number of students	Mean	SD	Number	Mean	SD	
i-Ready Reading[b]	252	514.60	49.23	179	509.48	46.88	0.11	1,027	560.80	39.71	715	567.08	38.95	−0.16
i-Ready Vocabulary	258	508.46	55.06	185	503.37	52.57	0.09	1,036	555.52	40.33	729	560.70	40.61	−0.13
FSA–ELA	233	292.42	16.44	166	289.46	15.81	0.18	981	309.87	18.34	695	312.99	18.00	−0.17

SD is standard deviation. FSA–ELA is Florida Standards Assessment English Language Arts.

Note: The sample included 49 teachers in the Word Knowledge Instruction group and 43 teachers in the control group, representing 39 schools.

a. Following What Works Clearinghouse procedures, Hedges' *g* was used to calculate the effect sizes. See What Works Clearinghouse (2020a) for more information.

b. i-Ready Reading was used as the baseline measure for the three researcher-developed, short-term outcomes (real-word decomposition, nonword derivation, and inferencing of word meanings).

Source: Authors' analysis of school district data for 2017/18 and 2018/19.

References

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B Methodological*, *57*(1), 289–300. http://www.jstor.org/stable/2346101.

Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of Educational Research*, *80*(2), 144–179. https://doi.org/10.3102/0034654309359353.

Carlisle, J. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing, 12*(3), 169–190.

Carlo, M., August, D., McLaughlin, B., Snow, C., Dressler, C., Lippman, D., et al. (2004). Closing the gap: Addressing the vocabulary needs of English language learners in bilingual and mainstream classrooms. *Reading Research Quarterly, 39*(2), 188–206.

Curriculum Associates, LLC. (2018). *i-Ready assessments technical manual*.

Florida Department of Education. (2018). *Florida Standards Assessments 2017-2018: Volume 4 evidence of reliability and validity*. http://www.fldoe.org/core/fileparse.php/5663/urlt/V4-FSA-1718-TechRpt.pdf.

Kieffer, M., & Lesaux, N. (2007). Breaking down words to build meaning: Morphology, vocabulary, and reading comprehension in the urban classroom. *The Reading Teacher*, *61*(2), 134–144. https://doi.org/10.1598/RT.61.2.3.

Kieffer, M., & Lesaux, N. (2010). Morphing into adolescence: Active word learning for English language learners and their classmates in middle school. *Journal of Adolescent & Adult Literacy*, *54*(1), 47–56.

Lesaux, N. K., Kieffer, M. J., Kelley, J. G., & Harris, J. (2014). Effects of academic vocabulary instruction for linguistically diverse adolescents: Evidence from a randomized field trial. *American Educational Research Journal, 51*(6), 1159–1194. http://eric.ed.gov/?id=EJ1045653.

Nagy, W., Berninger, V., & Abbott, R. (2006). Contributions of morphology beyond phonology to literacy outcomes of upper elementary and middle-school students. *Journal of Educational Psychology, 98*(1), 134–147.

Tyler, A., & Nagy, W. (1989). The acquisition of English derivational morphology. *Journal of Memory and Language, 28*(6), 649–667.

What Works Clearinghouse. (2020a). *Procedures handbook, version 4.1*. http://eric.ed.gov/?id=ED602035.

What Works Clearinghouse. (2020b). *Standards handbook, version 4.1*. http://eric.ed.gov/?id=ED602036.

Wood, C., & Bustamante, K. (2017). *Word Knowledge Instruction*.