Decoding Success: Phonics-based Instruction and Student Achievement

ALEX HOLLINGSWORTH, OHIO STATE UNIVERSITY AND NBER ALBERTO ORTEGA, INDIANA UNIVERSITY AND NBER NIKOLAOS ZIROGIANNIS, INDIANA UNIVERSITY

A federal grant that encouraged school districts to teach kindergarten through third-grade reading using phonics-based curricula improved reading scores. These gains persisted through at least the eighth grade. Treated students outperformed peers on questions that required basic decoding skills and those with more words, a benefit that also drove improvements in math scores. Students exposed to the phonics program were also less likely to drop out of high school and more likely to graduate on time. The program further caused students to graduate from four-year colleges at higher rates and increased earnings in young adulthood. The program cost was modest at about $132 per student per year. These costs were repaid by increased tax revenue from higher earnings by the time the phonics-exposed students turned 31.

Can a government program change how students are taught to read, and does it matter in the long run?

Learning to read is a foundational skill, yet fewer than one in three US fourth graders reads at a proficient level. Reading pedagogy has been a source of great debate among teachers and education scholars, with the science of reading and whole-language approaches representing two sides of the debate. The former relies on phonics and teaches children to map letters to sounds, while the latter emphasizes context clues and text immersion. Recently, phonics-based approaches have surged in popularity. Since 2013, more than 40 states have passed laws mandating adoption of phonics-based science of reading curricula. This is despite the fact that the largest attempt to incorporate phonics in US elementary education was cut in 2008 after a government evaluation found no statistically significant evidence that it improved reading comprehension.

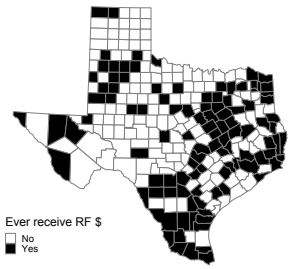
In a new study (Hollingsworth, Ortega, and Zirogiannis, 2026), we study this program. Reading First, created by the No Child Left Behind Act in 2002, sent roughly $1 billion per year to states to promote phonics-based instruction in kindergarten through third grade. Texas, the focus of our study, applied eligibility criteria to distribute $630 million to underperforming school districts.

We use Texas administrative records on 7.7 million third graders to follow public school students from third grade through college and into the labor market. About 2.1 million of them attended third grade in a Reading First–eligible district after the program began. To identify the effects of the program, we exploit the eligibility formula Texas used to allocate Reading First funds. School districts in Texas were eligible only if at least 13.6% of their third graders had failed the reading test in 2002 and if they met other income-based criteria. A separate funding formula determined how much money each eligible district would receive.

We find that Reading First increased third-grade reading test scores by 0.1 standard deviations, with the money districts received spent mainly on reading materials and teacher salaries. The effect persisted through eighth grade and spilled over to math scores (through math questions that require more reading). Students exposed to Reading First were more likely to graduate from high school and to complete a bachelor's degree, and once in the labor market they earned $509 more per year at age 25.

KEY FINDINGS

- **Reading and math scores improved.** Reading First improved both reading and math scores by about 10% of a standard deviation.
- **The gains lasted through eighth grade and out-lived the funding.** The effect persisted as students progressed through elementary and middle school, even after the funding ended.
- **The money changed what was taught, not who taught it.** Districts bought reading materials and paid existing teachers more rather than hiring new teachers. Students gained most on longer exam questions and on questions that test comprehension of a whole passage.
- **Important long-run effects.** Exposed students are more likely to finish high school on time and obtain a bachelor's degree. Once in the labor market, they earn $509 more per year at age 25.
- **The program paid for itself.** Reading First cost about $400 per child and was a very cost-effective policy considering the reading gains it made possible. Given the earnings effects we find, the tax revenue it generated exceeded the program's cost by the time Reading First students reached age 31.



Texas counties with at least one school district that received Reading First funds (black). The program reached districts in every region of the state.

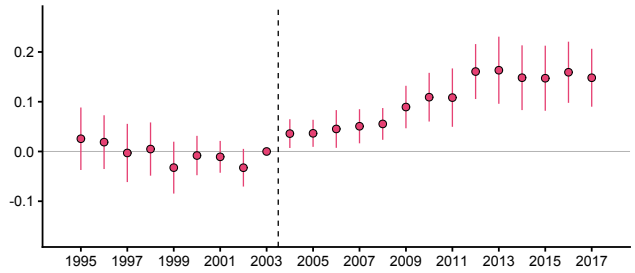
FINDING 1

Third-grade reading scores rose by about 10% of a standard deviation in eligible districts, and the gain grew as more students were exposed to Reading First

Figure 1 shows the gap in third-grade reading scores between eligible and ineligible districts in each year, measured relative to the gap in 2003, the last year before any money arrived. Before 2004, the difference in reading scores between the two groups of districts was constant. The first year of funding (2004) improved third-grade reading scores for eligible districts. The effect grew through 2012 and then leveled off at about 0.15 standard deviations. If we average the effect over 2004–2017, eligibility raised third-grade reading scores by 0.11 standard deviations. Math scores also rose by 0.11 standard deviations (see Finding 3).

The reason for the growing estimates between 2004 to 2012, is that as students spent more of their K-3 years in Reading First classrooms, the effect of the program grew. A third grader tested in 2004 had spent only one year in a Reading First classroom, while students tested in 2007 had four years of exposure to the program. Every third grade class after 2007 received the full impact of Reading First.

Figure 1: Reading scores in eligible districts pulled ahead after 2004 and stayed ahead



This figure shows the difference in standardized third-grade reading scores between Reading First–eligible and ineligible districts in each year, relative to the gap at 2003. The dashed line marks the first year of funding. Bars are 95% confidence intervals.

These results hold when we apply a series of robustness checks including comparing only districts just above and just below the 13.6% fail rate eligibility, dropping the years Texas changed its exam structure, and adding controls for district demographics. We also check other district finance sources and ensure that Reading First was the only revenue source that was different between the eligible and ineligible districts.

When we restrict comparisons to be only among similar schools (e.g., only the poorest schools) or similar students (e.g., only the poorest students) we find similar effects for nearly every group, demonstrating that our findings are not driven by targeting or by a particular type of school or student. We do find some evidence that the effects were largest in schools with above-median shares of economically disadvantaged, Hispanic, and Title I students.

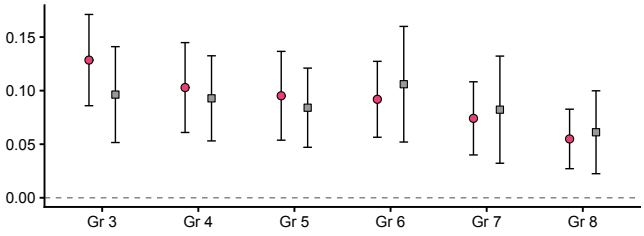
FINDING 2

The gains lasted through eighth grade and outlived the funding

In contrast to many early-childhood interventions, whose benefits fade within a few years, Reading First showed improvements that persisted over time. Students who received phonics-based instruction mandated by Reading First showed reading test improvements, all the way through eighth grade (Figure 2). Students who scored in the bottom decile of the reading distribution in third grade were less likely to be in the same low decile by the time they reached eighth grade.

Furthermore, the positive effects of Reading First persisted even after federal money stopped flowing to school districts and their Reading First spending fell to zero. Third graders in formerly funded districts kept outscoring their peers in every year through 2017 (Figure 1). This is suggestive evidence that Reading First changed how schools approach reading pedagogy in a way that extended beyond the cohort of students exposed to the program during the years it was funded.

Figure 2: Students exposed in early grades still read better in eighth grade



This figure shows the effect of Reading First eligibility on the same students' standardized reading scores in each grade from three through eight. Circles refer to all districts; squares refer only to districts within five percentage points of the 2003 state-wide fail rate of 13.6%. Bars are 95% confidence intervals.

FINDING 3

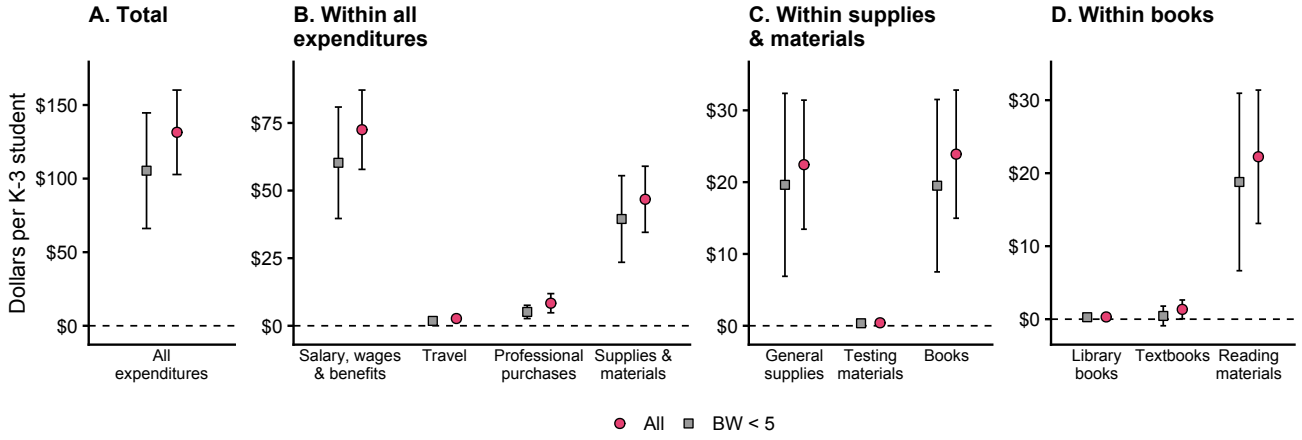
The money changed what was taught, not who taught it

How did districts spend their Reading First funding? During the funding years, eligible districts spent about $132 more per student each year from their Reading First grants (Figure 3). About $73 went to salaries, wages, and benefits and $47 to supplies and materials. Half of the materials spending was allocated to reading materials rather than textbooks or library books. This is the pattern one would expect if the districts bought materials to support phonics curricula.

Could it be that the districts used their Reading First funds to hire more teachers or attract teachers from better-performing districts? We do not find evidence either of teacher reallocation across districts or of an increase in teacher-to-student ratio. The money went instead toward paying existing teachers, consistent with the coaching and professional development the program required.

Was it really the phonics instruction that changed student performance? Since we do not know what teachers taught in the classroom or the content of the reading materials they used, we measure how students performed on different types of exam questions. We digitized every question on third-grade reading and math exams we could obtain and classified each question by its length in number of words, whether it tested overall comprehension, and what share of its words can be decoded with basic phonics skills (the comprehension classification is only relevant for reading exams). Reading First students gained the most exactly where reading skills matter the most: on longer questions, questions that test comprehension of a whole passage, and questions with more words that can be sounded out with basic phonics. Students performed better on math questions with longer verbal instructions, arguably because they were better at understanding those instructions. On math questions that involved pure computation (those whose answer choices are all numbers), there was no advantage at all for Reading First students, evidence of their improvement coming strictly through better reading skills, not because somehow Reading First improved fundamental math skills in addition to reading.

Figure 3: Reading First dollars went to spending on existing staff and reading materials, not on new textbooks



This figure shows how much more eligible districts spent per K–3 student from Reading First funds during 2004–2011 by category. Each panel breaks down the last bar of the panel to its left, so each panel has its own dollar scale. Circles refer to all districts; squares refer to districts within five percentage points of the 2003 state-wide fail rate of 13.6%. Bars are 95% confidence intervals.

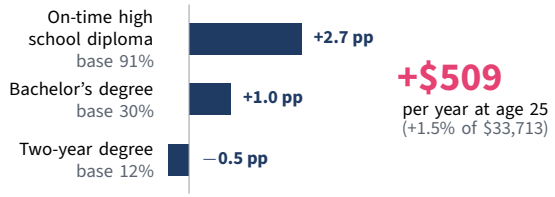
FINDING 4

Reading First students were more likely to finish high school on time, earn a bachelor's degree, and earned more money at age 25

The program's benefits extended into adulthood. We follow 5.9 million students who were in third grade between 1994 and 2014 as they grew up, 3.2 million of them into the labor market at age 25. Students in eligible districts were more likely to earn a standard high school diploma on time and obtain a bachelor's degree as shown in Figure 4. We also find a small decrease in degrees from two-year colleges for Reading First students. This is indicative of a substitution effect. Students that might have otherwise enrolled in a two-year college instead enrolled and completed a degree from a four-year institution as a result of their exposure to Reading First. These changes contributed to an earnings boost of $509 per year (a 1.5% increase relative to baseline earnings) at age 25, as illustrated in Figure 4.

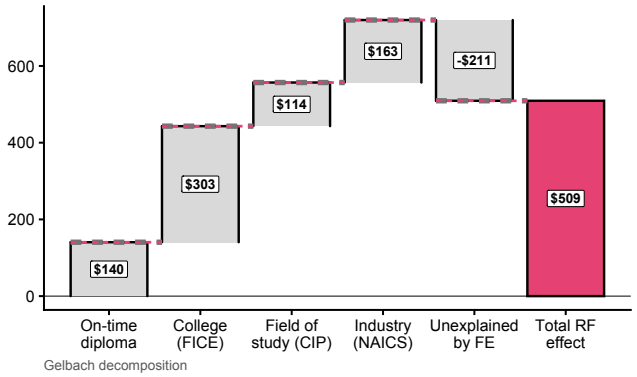
Figure 5 breaks down this $509 wage increase, attributing it to on-time high school graduation, bachelor's attainment, choice of college major, and choice of occupation. The main limitation in this approach is that we can only observe labor market outcomes for fully or partially treated students who were in third grade between 2004 and 2009, since our wage data end in 2025 (i.e. a student who attended third grade and was 8 years old in 2009, turned 25 in 2025). The literature suggests that age 25 is near the first year when the wage earnings premium from college attendance starts to materialize. That premium increases over time, so the improved labor market outcomes that Reading First students experienced will likely continue to grow.

Figure 4: Reading First raised high school completion, college graduation, and earnings at age 25



This figure shows the estimated effect of Reading First eligibility, with the pre-program baseline shown beneath each label, and the effect on annual earnings at age 25. Estimates compare eligible and ineligible districts before and after 2004 for third-grade cohorts 1994–2014 (1994–2011 for degrees, the last cohort observed to age 22).

Figure 5: Most of the earnings gain comes from more schooling and better jobs



This figure splits the $509 earnings effect into the parts accounted for by students finishing high school on time, which college they attended, their field of study, and their industry of employment at age 25. Together, these channels account for more than the total; the remainder is not explained by them.

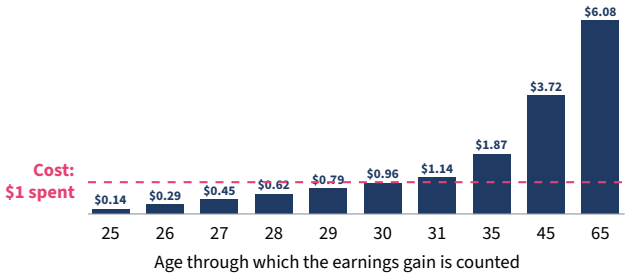
FINDING 5

The program paid for itself

Reading First was an exceptionally cost-effective policy. Eligible districts received about $132 per K–3 student per year while the funding lasted (2004–2011). That is $184 per student per year for each tenth of a standard deviation of reading gain, against the roughly $4,000 per student per year that research on general school spending finds is needed to buy the same gain. The $184 likely understates the full cost since the program relied on existing teacher time, administrative capacity, and professional development budgets. That said, even if we triple the program cost estimate, Reading First was far more cost-effective than other educational policies that achieve that same gains.

Students' earnings gains generated enough tax revenue to fully repay the cost of the program. If we count only the single year of earnings we can observe for the former students, their earnings at age 25, the government recovered 14 cents in taxes for every dollar spent. Assuming this earnings gain persisted as students aged, the program would have paid for itself by the time the former students reached age 31, when the added tax revenue reached $1.14 for every program dollar spent (Figure 6).

Figure 6: The net cost of Reading First (additional tax revenue earned per $1 spent in the program) became negative by the time the former students reached age 31



This figure shows the added tax revenue generated per dollar of Reading First spending when the age-25 earnings gain is assumed to have lasted through the stated age and then stopped. The first bar counts only the year of earnings that we observe. Amounts are discounted to age 7, the average age at which children were exposed, at 3%, and taxed at a combined tax-and-transfer rate of 19%.

Takeaways for policy

Phonics can be delivered at scale by a government program. More than 40 states have now legislated science of reading instruction. Reading First shows that a mandate backed by money for materials and coaching can raise reading scores and later earnings.

Early reading is a fork in the road. The gains persisted through eighth grade, into college, and into the labor market. Reading is an input into everything taught afterward, and a cheap one to improve.

What we cannot say. Reading First ran in nearly every state, and our evidence comes from only one of them. Moreover, we can't separately estimate the contributions of materials, coaching, and assessments within the program.

Want to learn more?

Read the paper **See the slides**

Contact: hollingsworth.126@osu.edu, alorte@iu.edu, nzirogia@iu.edu

All materials are freely available for use with citation. Supported by the Spencer Foundation (Grant #202400025). Results use restricted administrative data held by the Texas Education Research Center; conclusions are the authors' own.