Early Literacy Initiative Final Evaluation Report

Office of Research, Planning, and Program Evaluation

January 2024

MARYLAND STATE DEPARTMENT OF EDUCATION

Carey M. Wright, Ed.D.
Interim State Superintendent of Schools

Dr. Deann Collins
Deputy State Superintendent
Office of Teaching and Learning

Geoff Sanderson
Deputy State Superintendent
Office of Accountability

Krishna Tallur
Deputy State Superintendent
Office of Finance and Operations

Wes Moore
Governor

MARYLAND STATE BOARD OF EDUCATION

Clarence C. Crawford
President, Maryland State Board of Education

Joshua L. Michael, Ph.D. (Vice President)

Shawn D. Bartley, Esq.

Chuen-Chin Bianca Chang, MSN, PNP, RN-BC

Susan J. Getty, Ed.D.

Dr. Monica Goldson

Nick Greer

Dr. Irma E. Johnson

Dr. Joan Mele-McCarthy, D.A., CCC-SLP

Rachel L. McCusker

Samir Paul, Esq.

Holly C. Wilcox, Ph.D.

Abisola Ayoola (Student Member)

Table of Contents

Executive Summary

The Early Literacy Initiative (ELI) grant was implemented from 2019 to 2022 to support struggling readers by assisting up to 50 qualifying Title I schools across the state of Maryland to implement an evidence-based literacy program in helping students to meet literacy proficiency targets by the end of grade 8.[1] This evaluation report provides evidence on ELI-funded activities. More specifically, the report has two objectives. First, in order to determine the extent to which ELI-funded activities adhered to the requirements of the grant, this report presents evidence from a process evaluation. Secondly, in order to determine the extent to which activities improved student-level academic outcomes, this report presents evidence from an impact evaluation. By presenting a broad range of evidence, the goal of the report is both to understand whether ELI grant-funded activities collectively worked as intended and, if so, to understand why.

In terms of process, the evaluation team carried out an analysis of student-level reports and data submitted by LEAs. The report finds:

- Schools implemented 16 separate interventions to 4,551 students in grades Pre-K through 5 over the four years of the grant, and the number and type of interventions varied by school and by grade.

- Interventions broadly aimed to improve early literacy skills such as letter-sound recognition, phonics, and reading comprehension.

- Most students received interventions that have been shown through rigorous research to have strong evidence of a positive impact on early literacy outcomes.

- Students selected to receive ELI-funded interventions had significantly lower academic proficiency as measured by their Kindergarten Readiness Assessment (KRA) Language and Literacy score, and to a lesser extent, higher prior ELA course failure rates, compared to their peers in the same grades and schools. ELI students were also significantly more likely to receive free and reduced-priced meals (FARMs).

In terms of impact, the evaluation team used longitudinal student-level data to estimate short-term (one-year) changes in grade repetition and ELA course failure for four cohorts of ELI students relative to their same-grade (and school) peers. It also examines changes in ELA MCAP scores from 2021-22 to 2022-23. The report finds:

- There is no compelling evidence to suggest ELI-funded interventions improved outcomes. Although there is evidence of a statistically significant reduction in one of three outcomes (grade repetition) for one of four cohorts of students (2019-20), no other differences are estimated as statistically significant and/or robust to the inclusion of alternative school-, grade-, and individual-level factors that could have driven outcome differences for students.

---

[1] COMAR 13A.060.0, the law establishing the grant, and Maryland House Bill 1415 stated that the program was to help students meet literacy proficiency targets by the end of grade 8. http://mdrules.elaws.us/comar/13a.06.09.01; https://mgaleg.maryland.gov/2018RS/bills/hb/hb1415e.pdf

- Restricting the sample of students to those who entered kindergarten with below-average literacy assessment scores, which more closely aligns to the criteria for selection to receive ELI-funded interventions, does not change the conclusions of the impact evaluation.

- There is no compelling evidence to suggest ELI-funded interventions improved outcomes for students who received them for two or more years, or for students who completed interventions and met the individual targets set for them.

- Effects of the pandemic and a number of data limitations likely have implications for the findings of the impact evaluation. This is discussed further in the Limitations, and Conclusion and Discussion sections of the report.

Introduction

Across the U.S., the growing prominence of the "Science of Reading" and similar efforts to improve evidence-based literacy instruction have focused policymakers' attention on the early grades, where gaps in educational achievement emerge as early as K-3 (Reardon et al., 2012). In Maryland, 4th grade reading scores on the National Assessment of Educational Progress (NAEP) stayed flat between 2011 and 2019. The state's assessment showed small improvements in 3rd grade ELA proficiency from 2015 to 2019, but with prominent gaps by race and socioeconomic group.

House Bill 1415, approved in May 2018, established the Maryland Early Literacy Initiative. In early 2019, the Maryland State Board of Education passed COMAR 13A.06.09 to create the Early Literacy Initiative. The purpose of the Early Literacy Initiative (ELI) grant was to support struggling readers by assisting up to 50 qualifying Title I schools to implement evidence-based literacy programs for students who were at risk of not meeting literacy proficiency targets by the end of grade 8. The ELI grant was based upon a school-level needs assessment and aligns with the local education agency (LEA) comprehensive literacy plan, as well as Maryland's 2020 State Comprehensive Literacy Plan, *Maryland's Keys to Comprehensive Literacy.*

Across the U.S., previous state-led efforts to improve reading outcomes have featured some combination of academic support, wraparound services for disadvantaged students, and/or mandated retention for students who do not meet pre-defined testing standards (Olson, 2023; Strunk et al., 2021; Ladd et al., 2015). Such efforts have shown mixed success (Westall and Cummings, 2023), but leave questions about the relative effectiveness of different components of programs and interventions. For example, some states have mandated grade retention for students who do not meet specific performance benchmarks on standardized tests while also reforming teacher training and requiring or encouraging alternative instructional materials in classrooms (Olson, 2023). The passage and/or adoption of different programs and initiatives at the same time makes it difficult to disentangle the effect of each one individually.

Building on the Early Literacy Initiative Grant Summary Report (January 2023), the goal of this report is to evaluate process and impact for students in schools that received an ELI grant and to contribute to the evidence base on narrowly targeted, state-led efforts to improve early grade literacy outcomes in the U.S.

PROGRAM DESCRIPTION AND CONTEXT

ELI provided $2.5 million over four years (2019-2022) to assist qualifying Title I schools to implement evidence-based early literacy programs for students deemed to be at risk of not meeting literacy proficiency targets by the end of grade 8. All LEAs were invited to apply for the Early Literacy grant on behalf of the Title I schools in their district, and eight LEAs (53 ELI grant participating schools) were selected. Schools were selected competitively through a grant application, where priority was given to schools with high concentrations of Pre-K to grade 3 students in poverty and those who proposed strong or moderate evidence-based strategies or interventions (see corresponding research question for more information on evidence base). Although only two out of eight LEAs explained how they chose the specific schools for the grant, both identified schools using literacy data.

Required elements for the applications were:

- Use of evidence-based strategies and interventions;

- Inclusion of a needs assessment based upon school-level data;

- Inclusion of early literacy intervention services for participating students;

- Provision of direct services to participating students at least twice a week;

- Inclusion of clear literacy targets at each grade level, pre-kindergarten through grade 8;

- Inclusion of a plan for benchmark assessment multiple times a year to identify students who need one-on-one interventions;

- Inclusion of a plan to collect data on student progress at least monthly;

- Inclusion of a plan for implementation and monitoring through the LEA or a non-profit organization; and

- Inclusion of the alignment with the LEA Comprehensive Literacy Plan.

Within each school, students were evaluated at the beginning of the year, and those identified as at risk of not meeting proficiency targets were selected to participate in the intervention. Students were selected based on literacy skills, but assessments and skills screened were not uniform across LEAs (See Appendix A Table 1). The program started in the 2018-2019 school year, and LEAs could request an extension into the 2021-2022 school year due to interruptions caused by the COVID-19 pandemic[2].

The ELI grant funds were administered during the COVID-19 pandemic, the associated school closures, and the subsequent switch to virtual instruction. For this reason, it is important to consider the ways in which ELI schools and students may have been impacted. Across Maryland, as in other states, students' academic proficiency on state assessments as well as scores on national assessments showed large declines from 2019 through 2022, with disproportionately negative impacts for children of color and economically disadvantaged students. Given that eligibility for the grant was restricted to Title I schools, and that students selected to receive ELI-funded interventions experienced other measures of disadvantage (see Table 2) it is reasonable to think the pandemic may have affected ELI schools and students disproportionately.

LEA reports submitted at the completion of grant activities reinforce the notion that schools' implementation of ELI programs was negatively impacted. Many schools reported challenges in making the switch from in-person instruction to facing school closures, followed by the subsequent switch to virtual instruction, and then transitioning again to a hybrid virtual and in-person model. In many cases, tutors and paraeducators providing ELI-funded services had to quickly learn how to juggle both types of instruction. One LEA struggled to establish contact with ELI students after school closures and had to work with schools to identify new students in need of services. Another LEA hired an external vendor to provide virtual instruction, and the delays in finalizing the contract resulted in delays in implementation. These examples highlight specific challenges in a context within which schools were already facing teacher shortages and broader challenges.

---

[2] Of the participating LEAs, only Washington County did not receive an extension for the 2021-2022 school year.

Evaluation Framework

This evaluation seeks to answer two broad questions in regard to ELI:

1. How did LEAs and schools implement ELI-funded interventions?

2. To what extent did the ELI improve academic outcomes for students?

To answer the first question, the evaluation team drew upon available implementation data to conduct a process evaluation. A process evaluation is a form of implementation research that is designed to understand the extent to which a program was implemented as initially intended (Hill et al., 2023). To answer the second question, the evaluation team conducted an impact evaluation, which is designed to answer the question of whether the program had an effect on student-level outcomes. When paired together, the ability to provide evidence in regards to both ELI's implementation and impact allows the evaluation team to answer a broad number of questions that are of interest to policymakers and practitioners in Maryland.

An important consideration in outlining the scope of an evaluation is to distinguish the ELI grant program from the separate actions and interventions that schools chose to implement in response to being selected for the grant. In this regard, it can be tempting to evaluate process and impact for each school's unique implementation of grant-funded activities. This evaluation thus aims to strike a balance between answering questions about the ELI grant program and answering questions about schools' separate implementation of grant-funded activities. To better address the broad research questions, this evaluation aims to answer the following evaluation questions (EQ):

- EQ 1: What are the characteristics of schools and students who participated in ELI?

- EQ 2: What did ELI programs look like, and which interventions did LEAs administer for students?

- EQ 3: Did interventions adhere to an evidence base?

- EQ 4: Did students complete interventions and did they achieve intervention goals?

- EQ 5: What was the impact of the grant?

- EQ 6: What was the impact of the grant by subgroup?

Data

This section describes the data sources and variables used to answer the process and impact research questions. To answer research questions about characteristics of ELI interventions and the process of implementing them, the evaluation team used final project reports submitted by LEAs. To answer research questions about student completion and success, the evaluation team used a spreadsheet compiled by MSDE's Division of Curriculum, Instructional Improvement and Professional Learning that listed state-assigned student IDs, schools, grades, and intervention years for all students who received ELI interventions. For each student and year, the team was provided with information on the intervention(s) that the student received, whether the student completed the intervention and reason for not completing the intervention(s), whether the student accomplished goals that were set for the intervention(s), and reason for not completing goals.

To examine questions about student selection and program impact, the evaluation team constructed a student-level dataset using data from MSDE's multi-year data warehouse. The construction of the dataset and explanation of outcome measures is described in detail in Appendix F. The intervention dataset contains 117,511 student-by-academic year observations (22,649 unique students), which includes 5,743 intervention observations (4,551 unique students). See Appendix A Table 3 for more information. For ease of analysis, students are divided into cohorts (2019 through 2022) based on when they first appear as an ELI student or first appear in the same school and grade as an ELI student. Outcome measures include grade repetition, ELA course failure, and ELA standardized assessment results for available years.

Process Evaluation

This section provides evidence for whether or not LEAs and schools adhered to the requirements of the grant, and whether they were able to report successfully implemented interventions and met individual goals. These questions can be thought of as touching on implementation and context, two of the four key areas that often guide implementation research (Hill et al., 2023). Although data collection efforts for this evaluation were not designed to yield a full and comprehensive picture of program implementation, evidence in this section can be useful in informing how ELI interventions were implemented and is also useful in further illuminating findings from the impact study.

EQ 1: WHAT ARE THE CHARACTERISTICS OF SCHOOLS AND STUDENTS WHO PARTICIPATED IN ELI?

The ELI grant was designed to target Title I schools and students who experienced reading and other academic challenges in the early elementary grades. To better understand the extent to which interventions were successfully targeted toward schools, Table 1 presents characteristics of the 53 ELI schools. Because LEAs were responsible for selecting schools, the table compares ELI school characteristics to 292 other elementary grade schools in the same LEAs as those that implemented ELI interventions, for the 2017-18 academic year (one year prior to the initial administration of ELI funds). The table shows that ELI schools had larger shares of Black/African American students (55% compared to 41%) and students receiving free and reduced meals (65% compared to 44%), while they had lower shares of White students (20% compared to 31%).

To better understand the extent to which interventions were targeted toward students, Table 2 shows characteristics of ELI students for each year. Because students could have received ELI interventions in multiple years, we assigned each student to a cohort based on the year in which they first received interventions. To understand differences, the table shows "non-ELI" students as those who were in the same school and grade as ELI students. The table shows there are important differences within cohorts between ELI and non-ELI students, as well as between cohorts. Perhaps the most striking difference is that ELI students had lower reading domain KRA scores, suggesting the program was effective in targeting interventions toward students experiencing academic difficulties. It is also notable that for the 2020 through 2022 cohorts, ELI students were at least 10 percentage points more likely to receive free and reduced meals, although the difference is not as large for the 2019 cohort. Lastly, although ELI schools were more likely to have higher shares of Black/African American students than other schools in their respective LEAs (Table 1), within ELI schools, there were not particularly large differences in race/ethnicity between ELI and non-ELI students. For example, there were slightly larger percentages of Black/African American and Hispanic/Latino students among ELI participants compared to non-ELI students in the 2021 cohort (6-10 percentage points), but otherwise the race/ethnicity of students remains fairly similar (2-5 percentage points) for ELI and non-ELI students across cohorts.

Table 1 - Characteristics of ELI and Comparison Schools/Student Bodies, 2017-18 (%)

Characteristic	Non-ELI Schools in Same LEAs as ELI Schools	ELI Schools	Difference
Female	48.1	48.1	0.0
Male	51.9	51.9	0.0
Asian	6.9	2.0	-4.9*
Black/African American	40.5	54.9	14.4*
Hispanic/Latino	17.2	19.4	2.2
White	30.9	19.7	-11.3*
Other Race/Eth.	4.5	4.1	-0.5
FARMs	43.5	64.7	21.2*
English learners	11.7	13.8	2.1
Students with Disabilities	13.0	13.4	0.4

Note: Table shows average characteristics (in percentages) of September K-12 enrollment for the 2017-18 school year. ELI schools (N = 53) are those that received ELI interventions at any point between the 2018-19 and 2021-22 school years. Comparison schools (N = 292) are other elementary grade schools in the same LEAs as ELI schools. *Difference is statistically significant at the 5% level after applying a multiple hypothesis test procedure.

Summary:
ELI schools had higher shares of students eligible for FARMs, higher shares of Black/African American students, and lower shares of White students, compared to schools in the same LEAs. Within ELI schools, students selected to receive ELI-funded interventions were significantly more likely to participate in FARMS. Based on available data, ELI students also had significantly lower academic proficiency as measured by their KRA Language and Literacy score, and to a lesser extent, higher prior ELA course failure rates.

Table 2 - Characteristics of ELI and non-ELI Students Within Schools that Implemented ELI-funded Interventions

	Cohort							
	2019		**2020**		**2021**		**2022**	
	N.E.	E.	N.E.	E.	N.E.	E.	N.E.	E.
Panel A: Characteristics of Students								
Female (%)	48.9	52.5	47.5	46.7	48.7	46.0	48.3	47.8
Male (%)	51.1	47.5	52.5	53.3	51.3	54.0	51.7	52.2
Asian (%)	0.9	0.3	3.8	2.1*	2.0	1.8	2.5	0.9*
Black/African American (%)	73.0	78.5*	47.2	49.6	52.8	45.9*	42.5	44.9
Hispanic/Latino (%)	17.0	14.4	34.7	33.3	19.9	30.2*	31.3	28.2*
White (%)	7.4	5.6	11.7	10.9	19.5	17.0	18.5	19.2
Other Race/Eth. (%)	1.8	1.2	2.6	4.1*	5.8	5.0	5.2	6.8*
FARMs (%)	65.9	69.8*	64.1	74.5*	59.7	71.0*	59.7	70.2*
Students with Disabilities (%)	13.7	4.1*	11.8	8.3*	10.8	9.6	10.5	8.6
English learners (%)	15.8	10.9*	29.8	29.4	11.9	19.3*	20.7	17.8*
Repeated grade (%)	2.6	5.3*	1.5	1.7	0.7	1.0	1.2	1.4
KRA (Std. Score)	0.114	-0.177*	0.090	-0.211*	0.112	-0.130*	-0.059	-0.243*
Failed ELA Course (%)	12.0	15.7*	5.7	11.7*	4.2	2.8	8.7	10.0
Panel B: Number of Student Observations								
Overall	4171	755	4624	1128	5005	1197	4298	1471
KRA	2189	430	1670	537	1871	557	1856	627
Failed ELA Course	3481	618	3200	718	3481	917	2252	789

Note: Table shows characteristics of Non-ELI (N.E.) and ELI (E.) students by academic year cohort. Demographic and service data are from the September enrollment data collection, for the fall of the school year in which students received ELI interventions. Course failure information is for the school year prior to which students first received ELI interventions. KRA results are standardized to have a mean of 0 and standard deviation of 1 within year, for fall of the school year in which students received ELI interventions or before. Non-intervention students are those who were in the same school and grade as students that received ELI interventions. *Difference (with non-ELI, value not shown) is statistically significant at the 5% level after applying a multiple hypothesis test procedure.

EQ 2: WHAT DID ELI PROGRAMS LOOK LIKE, AND WHICH INTERVENTIONS DID LEAS ADMINISTER FOR STUDENTS?

A review of reports submitted by LEAs at the conclusion of ELI-funded activities suggests that programs looked quite different depending on the LEA. Even within an LEA, schools may have administered a combination of different interventions in different years. Different student groups and grade levels were prioritized to receive programs, even when they were identified as at risk of not meeting proficiency targets. The aforementioned challenges brought about by the pandemic, school closures and changes to instruction, had major implications for how ELI-funded activities were implemented.

To understand, in a general sense, how LEAs designed programs using ELI funding, the following are three examples with information drawn from LEA reports:

- Worcester County identified students at risk of not meeting proficiency targets using DIBELS data. Teachers, administrators, reading interventionists and literacy coaches then developed a "plan for success" for each student, and met regularly to discuss progress throughout the year. Students received interventions at least 2-3 times per week. [3]

- Baltimore City used ELI funds to implement tutoring interventions and to purchase supplemental literacy materials to support instruction. Literacy Labs were tutoring interventions supporting low income, EL, homeless, and/or special education students in pre-K. Reading Partners used community volunteers trained by LEA staff and AmeriCorps members to serve students in elementary grades who were reading below grade level, who were fluent in English and that did not have an IEP designation. Experience Corps tutors provided small group instruction at least two times per week.

- Montgomery County implemented Stepping Stones for students in Pre-K and Sound Partners for students in grades K-3. Stepping Stones, an intervention designed to supplement classroom instruction, was administered to all Pre-K students in year 1 of the grant and then students were prioritized in the 2nd and 3rd years based on academic need. These lessons included some Spanish language prompts because the majority of ELI students in pre-K were Emergent Multilingual Learners. Students receiving Sound Partners were entirely prioritized based on need.

A review of these reports suggests that understanding how ELI programs were implemented and how they may have improved outcomes for students requires an appreciation that no two programs were the same. Although the goal of this evaluation is to summarize for the purposes of describing implementation and impact, care will be taken to qualify and disaggregate information when available.

As part of ELI-funded activities, schools and LEAs were required to choose evidence-based strategies and interventions for improving student literacy outcomes. To better understand which interventions LEAs administered and for how many students, Table 3 summarizes student-level data provided by LEAs to the evaluation team. The table shows some commonalities; for example, at least two LEAs administered Wilson Fundations, Sound Partners, and Stepping Stones. The table also suggests that some LEAs used ELI funding to tailor interventions at a highly individualized level; for example, Charles County reported administering 10 separate interventions, while Kent and Somerset each implemented only 1.

---

[3] More detailed intervention descriptions available in Appendix A Table 2.

Summary:

LEAs implemented 16 separate interventions to students in grades Pre-K through 5 over the four years of the grant. Some LEAs chose a diversity of interventions, while other LEAs only focused on a single or small group of strategies. Recognizing that no two ELI-funded programs were the same, the information in this section suggests that interventions broadly aimed to improve early literacy skills such as letter-sound recognition, phonics, and comprehension.

Table 3 - Intervention Summary by LEA

LEA	Intervention	Year(s)	Num. Schools	Grade(s)	Num. Students
Baltimore City	Reading Partners	2019 - 2021	14	K - 4	976
	Tutoring (Literacy Labs and Experience Corps)	2019 - 2022	11	Pre-K - 3	580
Carroll	Being a Reader/Writer	2020 - 2022	3	K - 4	97
	Wilson Fundations	2020 - 2022	2	K - 3	90
	Lexia	2021 - 2022	3	K - 3	90
	SIPPS	2022	1	K - 1	12
Charles*	Leveled Literacy Intervention	2021 - 2022	8	Pre-K - 5	505
	Wilson Fundations	2021 - 2022	7	K - 5	268
	Amplify	2022	8	2 - 5	118
	Heggerty	2021 - 2022	3	K - 3	19
	Tutoring	2022	1	K - 3	18
Kent	Sound Partners	2020 - 2022	3	Pre-K - 5	158
Montgomery	Sound Partners	2020 - 2022	10	K - 5	846
	Stepping Stones	2020 - 2022	6	Pre-K	362
Somerset	Leveled Literacy Intervention	2021 - 2022	3	K - 5	167
Washington	Stepping Stones	2019 - 2020	1	Pre-K - 4	44
	Sound Partners	2019 - 2020	1	K - 1	14
Worcester	American Reading Co.	2020 - 2022	3	Pre-K - 3	460
	Wilson Fundations	2022	3	K - 3	214
	Amplify	2021 - 2022	1	Pre-K - K	14

Note: Tabled data are based on data submitted by LEAs. Students may have received multiple interventions and may have received interventions for more than one year. Counts by grade level are as recorded in MSDE's data warehouse as of September 30 enrollment. * Charles County reported fewer than 10 students receiving iReady, Sound Partners, Words Their Way, Reading Recovery, and/or System 44.

EQ 3: DID INTERVENTIONS ADHERE TO AN EVIDENCE BASE?

A key requirement for ELI grantees was to demonstrate their chosen interventions were evidence-based. Although LEAs were provided flexibility in choosing interventions, many adhered to MSDE guidance that was aligned to the definition of evidence-based interventions included in the 2019 Ready to Read Act.[4] The following is a review of research and evidence on the effectiveness of the most common interventions chosen by grantees, covering over 95% of children that received ELI-funded interventions. The evaluation team used a combination of the What Works Clearinghouse (WWC), Evidence for ESSA, Google Scholar, and product website searches.

The What Works Clearinghouse[5] has produced intervention reports for Sound Partners, Leveled Literacy Intervention, Wilson Reading System (Fundations), Lexia Reading and Stepping Stones to Literacy. Each of these reports cites multiple studies that both meet WWC design standards and have produced positive effects on early literacy outcomes, including measures of phonics and comprehension. Additionally, the WWC cites one individual study of Reading Partners that meets its design standards and has produced positive impacts on early literacy outcomes. Evidence for ESSA, a more recent website that has different criteria for inclusion of studies, also includes a rating for different interventions based on the strength of evidence. This information is summarized in Table 4.

Summary:
Based on a review of evidence aggregation websites, Google Scholar and product website references, all of which focus on experimental or quasi-experimental research, most of the students that were targeted by ELI funding received interventions that have shown strong evidence of a positive impact on literacy outcomes for students in the early grades.

---

[4]The 2019 Ready to Read Act defines supplemental reading instruction as "evidence-based, sequential, systematic, explicit, and cumulative instruction or intervention to mastery of foundational reading skills including phonological or phonemic awareness and processing, phonics, and vocabulary to support development of decoding, spelling, fluency, and reading comprehension skills to meet grade level curriculum"; retrieved from https://marylandpublicschools.org/programs/Documents/ELA/ReadingDifficulties/SummaryTierIISupplementalPrograms.pdf.

[5] Intervention reports are retrieved from https://ies.ed.gov/ncee/wwc/.

Table 4 - Summary of Evidence on ELI Interventions

Intervention	What Works Clearinghouse	Evidence for ESSA	Other Evidence
	Description of evidence*	ESSA Rating	
Sound Partners	Medium to large for alphabetics, fluency, and comprehension; small for general reading achievement	Strong	
Leveled Literacy Intervention	Medium to large for general reading achievement and small for reading fluency and alphabetics	Strong	
Wilson Reading System	Small for alphabetics, fluency, and comprehension	Strong	
Lexia Reading	Small for alphabetics, fluency, comprehension, and general reading achievement	Promising	
Stepping Stones to Literacy	Small for alphabetics	-	
Reading Partners	Statistically significant improvements in comprehension, fluency, and sight word efficiency	Strong	
100 Book Challenge	-	-	Yes
Tutoring	-	-	Yes

Note: * that met WWC evidence standards, with or without reservations.

EQ 4: DID STUDENTS COMPLETE INTERVENTIONS AND DID THEY ACHIEVE INTERVENTION GOALS?

According to the Early Literacy Initiative Grant Summary Report (January 2023) report, "an average of 91% of students who participated in ELI met the LEAs' goal(s) set for the intervention. This is self-reported by the LEAs and cases that were missing data or had incomplete data were excluded." To better understand the extent to which LEAs met their goals for ELI interventions, this section draws from LEA reports and from the student-level summative data submitted by LEAs to MSDE. According to LEA reports, goals were designed to encompass a variety of literacy skills aligned with research that adheres to a Science of Reading framework, such as reading level, phonological awareness, phonics, high frequency words, and decoding, depending on grade. For instance, Baltimore City reported that 46% of K-5 students attained or surpassed benchmark levels on DIBELS at the end of the year. Kent County reported mixed results varying by literacy program, school, and year, ranging from 20% to 100% of students attaining goals. Montgomery County reported mixed results varying by school and year, but 93% or more Pre-K students met goals while 12% to 37% met goals in K-3 depending on grade and year.

When turning to student-level data (Table 5), information generally corroborated what was seen in the reports. Upon surveying the data, 82% of the records indicated that students completed the intervention, which means that they were judged to have participated fully for the year. Completion rates differed by LEA; when examining by LEA, Kent and Worcester had the lowest completion rates, where less than two-thirds completed the intervention. In contrast, records for all students in Somerset and Washington indicated completion. However, the percentage dropped to less than half (42%), for records where students completed the intervention and met the target. When examining by LEA, students in Baltimore City, Kent, and Worcester had the lowest completion-and-met-target rates of under one-third while records for all students in Somerset and Washington indicated completion and meeting of targets.

Table 5 - Intervention Data Summary of Intervention Completion

LEA	Number of students	Students completed* (range across schools)	Students completed and met target (range across schools)
Baltimore City	2,073	84% (41-96%)	31% (8-62%)
Carroll	163	83% (74-95%)	58% (50-62%)
Charles	990	92% (80-100%)	79% (27-100%)
Kent	193	59% (50-79%)	0% (0-0%)
Montgomery	1,430	89% (77-96%)	37% (2-53%)
Somerset	167	100% (100-100%)	100% (100-100%)
Washington	63	100% (100-100%)	100% (100-100%)
Worcester	663	46% (42-50%)	22% (18-29%)
Overall	5,742	82% (41-100%)	42% (0-100%)

Note: Table shows the % of students who completed their respective intervention(s) and the % of students who both completed and met the target set for them. The range shows the high and low % of students across schools for that LEA. * most common reasons for not completing include disruption to schedule, exited program and/or transferred out, absenteeism, and tutor ended service.

Summary:
 Although most (82%) students completed interventions, there was wider variation across LEAs in terms of the percentage of students who met goals/targets for ELI. This percentage ranged from 0% to 100% across LEAs, with an average across students of 42%. Similarly, LEAs reported varying success in implementing interventions and in meeting goals.

Impact Evaluation

This section details findings for the impact evaluation. Measuring the causal impact of ELI on student outcomes is an important part of understanding whether the grant program achieved its intended purpose. The following points summarize the primary considerations that guided the analysis in this section, and Appendix F provides additional detail about the research design, including sample selection, choice of comparison group, and empirical approach.

1. The empirical strategy follows a "difference-in-differences" design, comparing short-term (1-year) trends in outcomes for ELI students to those of their same-grade (and school) peers.[6] Because students were selected to receive ELI-funded interventions based on their risk of not meeting literacy targets, the main analysis in this section is supplemented with an analysis of trends in outcomes for students with below-median Language and Literacy domain KRA scores in the following section. While this considerably changes (and restricts) the available sample of students, the combination of these two analyses represents the strongest approach for understanding the causal impact of ELI-funded interventions.

2. It is important to note that interpreting estimates as causal rests on an assumption that trends in outcomes for ELI and non-ELI students would have been parallel if ELI had never been administered. This is a strong assumption, and so caution is urged in interpreting estimates as causal. This is discussed further in the Limitations section.

3. For each outcome, the analysis shows trends in raw outcomes side by side with regression-adjusted estimates of impact, which attempts to account for other factors that may drive outcome differences between groups of students.[7] This transparency allows the reader to understand how the evaluation team arrived at all preferred estimates.

4. Due to the differences in which schools administered ELI interventions across years, as well as the effects of the COVID-19 pandemic and school closures, separate estimates are provided for each academic year in which ELI was administered to students. Because increasing the number of estimates for each outcome increases the risk of finding a false significant effect, the evaluation team highlights findings that remain statistically significant after applying a multiple hypothesis test correction.[8]

EQ 5: WHAT WAS THE IMPACT OF THE GRANT?

ELA Assessment Scores

This section draws on MCAP assessment data to examine changes in ELA proficiency. It divides the sample of students into four cohorts (2018-19 through 2021-22) based on the year in which they first received ELI-funded interventions (ELI students) or first appeared in the same school and grade as a student receiving ELI-funded interventions (non-ELI students). For each of the four cohorts, the change in assessment outcomes from 2021-22 to 2022-23 is shown.

---

[6] See Appendix B for an intuitive explanation of this strategy.

[7] See Appendix C for more information on regression-based approach.

[8] See Appendix E for more information on multiple hypothesis test corrections.

Panel A in Figure 1 shows that for each of the four cohorts, ELA scores were higher for non-ELI students than for ELI students in the 2021-22 academic year. Despite this, the trends from 2022 to 2023 looked reasonably parallel for three of the four cohorts, indicating scores for all students improved at an approximately similar rate. Only for the 2020 cohort does it appear that ELA outcomes improved faster for ELI students. Panel B compares the unadjusted difference-in-differences value, shown with green triangles and calculated using the raw trends in Panel A, to a regression-adjusted estimate shown as circles with vertical bars. The value of using regression in this case is two-fold; one, to provide a summary estimate of impact while accounting for individual and school-level factors (such as grade level) that could drive differences in assessment outcomes, and two, to estimate the degree of uncertainty around impact estimates in the form of standard errors. The figure shows that despite the slight improvement for the 2020 cohort, none of the four estimates are statistically significantly different from zero. Overall, this analysis does not find evidence that receiving ELI-funded interventions led to an improvement in ELA assessment outcomes between 2021-22 and 2022-23.

Figure 1 - Impact on ELA Assessment Scores (2021-22 to 2022-23)
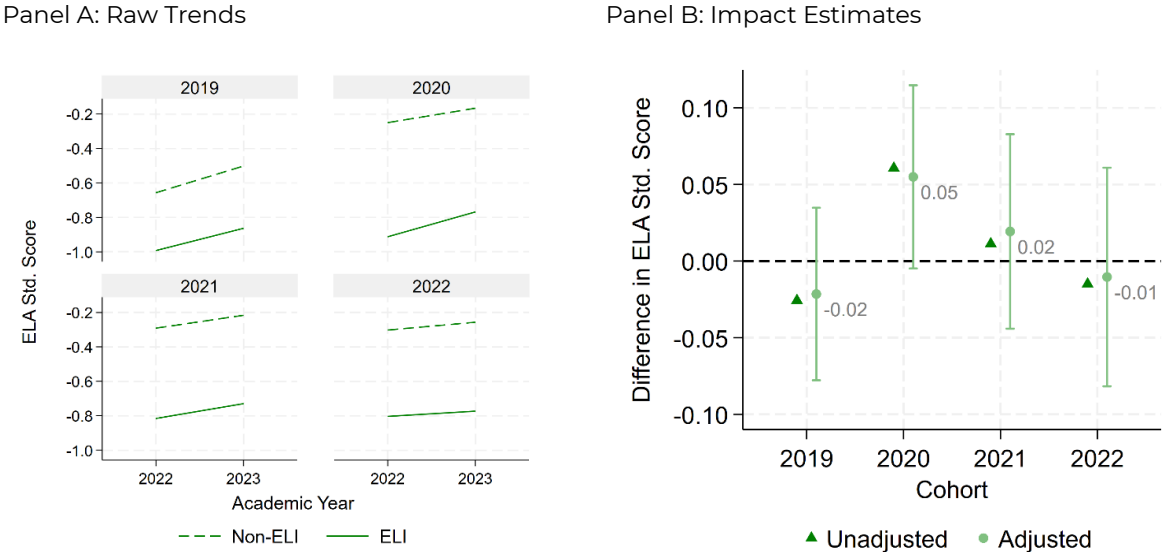
Note: * p<.05 ** p<.01, *** p<.001. Panel A presents the trend in ELA assessment scores from 2021-22 to 2022-23 for the 2019 (N=6,586), 2020 (N=6,216), 2021 (N=4,590), and 2022 (N=2,912) cohorts. Panel B compares the unadjusted difference-in-differences value calculated using the raw trends from Panel A to a regression-adjusted point estimate along with 95% confidence intervals. See Appendix C for additional details on regression analysis.

Summary:
No evidence of an effect of ELI on ELA assessment outcomes was found.
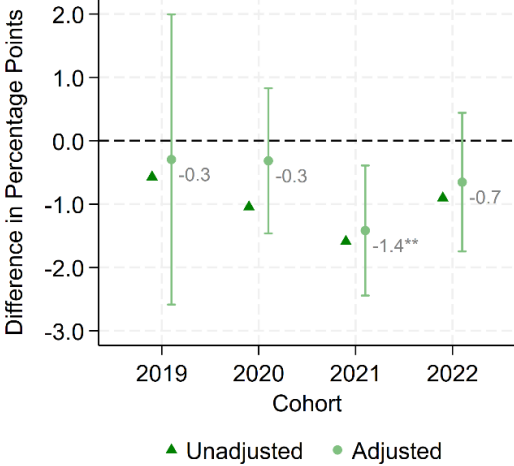
Grade Repetition

In contrast to the analysis of ELA assessment scores which are available only for the 2021-22 and 2022-23 school years, the analysis of grade repetition and ELA course failures draws on the availability of data following the same students for each of the school years in which the ELI grant funds were administered (2018-19 through the 2021-22 school years). In other words, as shown in Figure 2 Panel A, the data allow for an analysis of a change in grade repetition from the fall of the initial year in which students received ELI-funded interventions (year 1) to the fall of the subsequent year (year 2). Panel A shows that for each cohort, grade repetition declined for ELI students, while it either declined to a lesser extent (2018-19), stayed the same (2019-20), or increased slightly (2020-21 and 2021-22) for the other cohorts. It is also interesting to note that while the rate of grade repetition was much higher for ELI students in the fall of 2018-19 compared to their peers in the same grade, it began as approximately the same for each of the 2019-20 through 2021-22 cohorts before diverging in the following year. Panel B compares the unadjusted difference-in-differences values calculated using the raw trends in Panel A to regression-adjusted estimates with 95% confidence intervals. Despite the improvement in grade repetition, only one of the four estimates (2020-21) is statistically significantly different from zero (at the .01 level) after using regression to adjust for other factors that could have driven differences.

Figure 2 - Impact on Grade Repetition in the Short Term

Panel A: Raw Trends
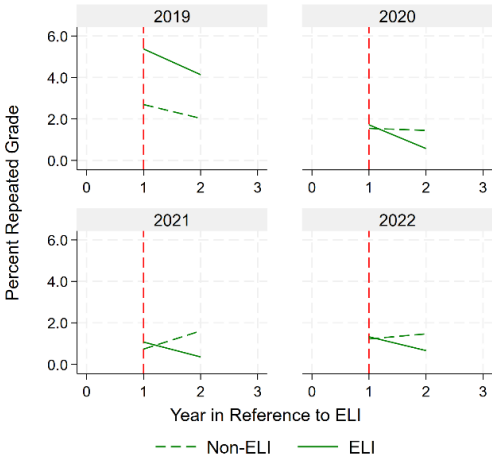
Panel B: Impact Estimates

Note: * p<.05, ** p<.01, *** p<.001. Panel A presents raw trends in grade repetition for the 2019 (N=4,624), 2020 (N=5,130), 2021 (N=5,462), and 2022 (N=5,102) cohorts. Panel B compares the unadjusted difference-in-differences value calculated using the raw trends from Panel A to a regression-adjusted point estimate along with 95% confidence intervals. See Appendix C for additional details on regression analysis.

Summary:

There is evidence for a statistically significant reduction in grade repetition for the 2021 cohort. The analysis shows grade repetition was reduced almost to zero (by 1.4 percentage points) for ELI students. This estimate remains significant at the 5% level after applying a multiple hypothesis correction (see Appendix E). There is no evidence that ELI reduced the rate of grade repetition for students in any of the other three cohorts.

ELA Course Failures

Similar to grade repetition, the analysis of ELA course failures draws on the availability of data following the same students for each of the school years in which the ELI grant funds were administered. In the analysis that follows, the change in the rate of ELA course failure is shown from the end of the school year prior to students receiving ELI-funded interventions (year 0) to the end of the school year in which students received ELI (year 1). Figure 3 Panel A shows that the rate of ELA course failure increased slightly for ELI students in the 2018-19 cohort while it decreased slightly for non-ELI students. For the other cohorts, the rate of course failure for ELI students decreased faster (2019-20), increased more slowly (2020-21), or decreased at the same rate (2021-22) as for non-ELI students. Panel B converts the raw trends into difference-in-differences estimates, comparing the unadjusted value to a regression-adjusted estimate with 95% confidence intervals. The figure shows for the 2021-22 cohort there was a regression-adjusted 2.8 percentage point reduction in ELA course failures for ELI students, which is statistically significant at the .05 level. The analysis suggests that while the rate of course failure increased for ELI students, it did not increase as much as it did for non-ELI students. Despite this, there is no evidence for a statistically significant reduction in course failures for any of the other cohorts.

Figure 3 - Impact on ELA Course Failures in the Short-Term

Panel A: Raw Trends

Panel B: Impact Estimates
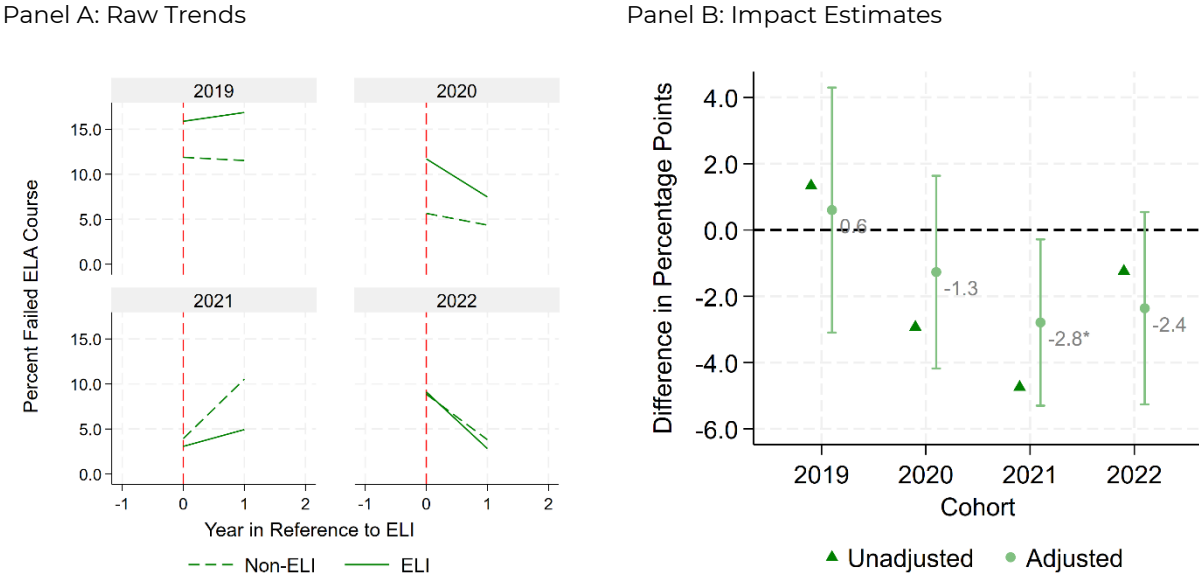
Note: * p<.05, ** p<.01, *** p<.001. Panel A presents raw trends in course failures for the 2019 (N = 4,011), 2020 (N = 3,859), 2021 (N = 3,474), and 2022 (N = 1,042) cohorts. Panel B compares the unadjusted difference-in-differences value calculated using the raw trends from Panel A to a regression-adjusted point estimate along with 95% confidence intervals. See Appendix C for additional details on regression analysis.

Summary: **The analysis above suggests there was a statistically significant (2.8 percentage point) reduction in ELA course failures for the 2021 cohort of ELI students, relative to non-ELI students. However, this estimate does not remain significant after correcting for multiple hypothesis testing (see Appendix E). Additionally, there is no evidence that course failures were reduced for other cohorts. Overall, the analysis does not find evidence that ELI-funded interventions reduced the rate of ELA course failure.**

## EQ 6: WHAT WAS THE IMPACT OF THE GRANT BY SUBGROUP?

This section presents an analysis of the impact of ELI for different groups of students. There are important reasons to think the effect of ELI-funded interventions may have differed depending on characteristics of students or schools. At the same time, there are numerous empirical and statistical problems with identifying too many groups and over-testing hypotheses (Wasserstein and Lazar, 2016; Gelman and Loken, 2013). For this reason, the analysis in this section limits the tested hypotheses to those with strong theoretical and practical rationale. What follows is an overview of the rationale for testing impacts for the particular groups of students in this section.

- As a reminder, because students were selected to receive ELI-funded interventions based on their risk of not meeting literacy proficiency targets, one section below analyzes outcomes for the subgroup of non-Pre-K students with *below-median Language and Literacy domain KRA scores*.

- Many students *received ELI-funded interventions for more than one school year*. It may be that students who received ELI-funded interventions for more than one school year experienced more intensive literacy supports compared to those who only received interventions for one year. At the same time, students who received ELI-funded interventions for more than one school year may have faced fewer disruptions and school closures, which could suggest they benefited from other advantages that would lead to better outcomes even without ELI. Despite the theoretical ambiguity, another section below analyzes the effects of ELI for the group of students who received interventions for more than one school year.

- Students who *completed ELI-funded interventions and met the targets* set for them may have seen improved outcomes for a number of reasons, including that they faced fewer disruptions at an individual or household level, that their schools were able to provide continuous programming throughout the year, or their likelihood of having improved outcomes could be correlated with differences in how schools define completion and set targets. Nevertheless, and again despite the theoretical ambiguity, because of the importance of this group of students, a third section below analyzes their outcomes.

- There may be differences in student outcomes by the *types of ELI interventions that schools administered*. Although the program was not designed in a way so as to be able to test this hypothesis convincingly, there is strong general interest in understanding whether there are any lessons to be learned in this regard. The fourth section below analyzes differences in average outcomes for students across schools, grouped by the types of ELI-funded interventions they administered, with the strong caveat that strong conclusions on the basis of this analysis should be avoided.

Group: Students with Below Median KRA Scores

This section presents estimates of impact for the group of students with below median KRA scores (who did not also receive ELI interventions in pre-K). The analysis focuses on regression-adjusted difference-in-differences estimates. Figure 4 shows impact estimates on ELA scores. Similar to the main effects, there is a positive but imprecisely (not statistically significant) estimated effect for the 2020 cohort. As can be seen, standard errors are wide enough that they would not be able to rule out an effect as large as 0.10 standard deviations. Overall, the analysis finds no evidence for an improvement in ELA scores for students with low prior literacy achievement.

Figure 4 - Impact Estimates on ELA Assessment Scores for Students with Below Median KRA Scores (2021-22 to 2022-23)
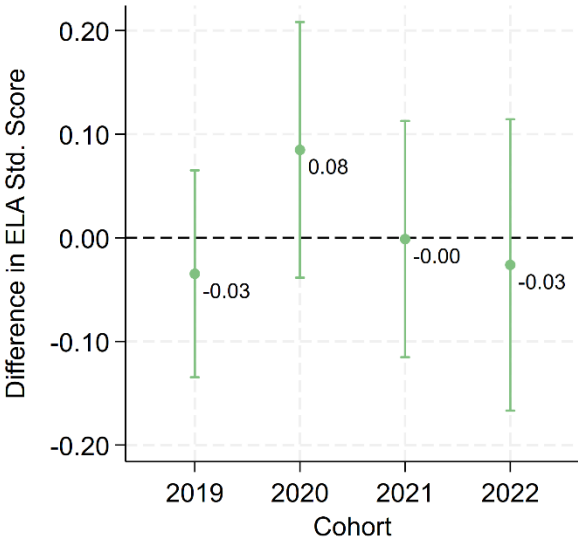
Note: * p<.05, ** p<.01, *** p<.001. Figure presents point estimates and 95% confidence intervals from regression-adjusted difference-in-differences estimates. See Appendix C for additional details on regression analysis.

In **Figure 5** Panel A shows regression-adjusted difference-in-differences estimates for grade repetition, and Panel B shows the same for ELA course failures. The reduction in grade repetition for the 2021 cohort is not replicated for the group of students with low prior literacy achievement, as proxied by the KRA domain scores; in fact, this estimate is close to zero. And although there is suggestive evidence for reductions in both grade repetition and course failures for the 2022 and (to a lesser extent) the 2021 cohorts, estimates are imprecisely estimated and not statistically significant. This analysis does not find evidence for an effect of ELI on academic outcomes for the group of students with low initial literacy assessment outcomes.

Figure 5 - Impact Estimates on Grade Repetition and ELA Course Failure for Students with Below Median KRA Scores, Short-Term
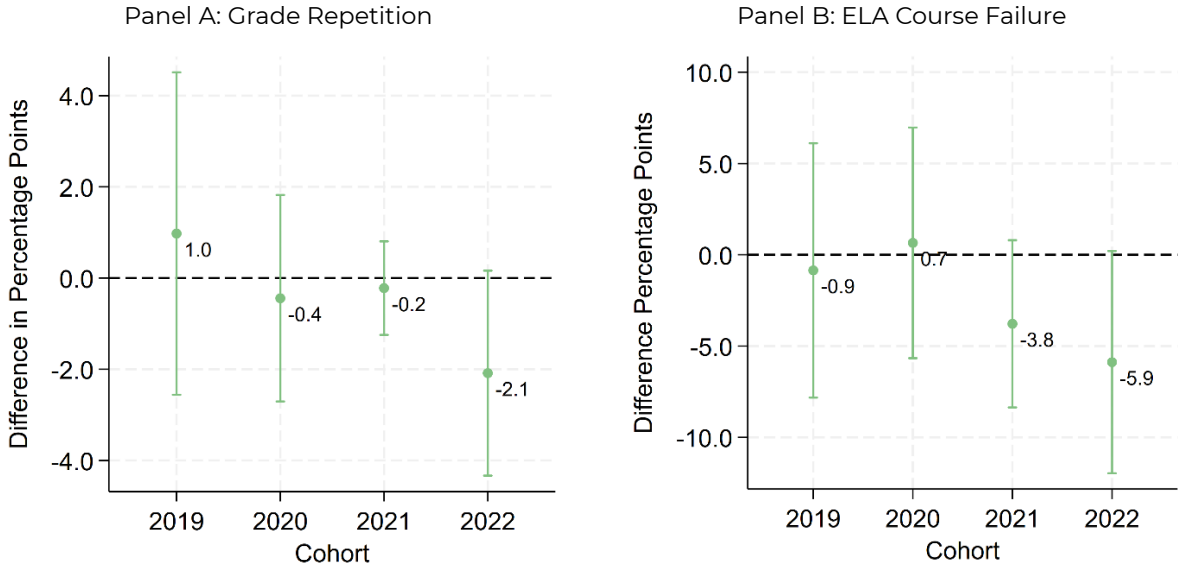
Panel A: Grade Repetition

Panel B: ELA Course Failure

Note: * p<.05, ** p<.01, *** p<.001. Panels A and B present point estimates and 95% confidence intervals from regression-adjusted difference-in-differences estimates. See Appendix C for additional details on regression analysis.

Summary:

Comparing short-term trends for ELI and non-ELI students among the group who entered kindergarten with below-median Language and Literacy domain KRA scores does not strengthen the interpretation of evaluation findings. The effect on grade repetition for the 2021 cohort is not replicated, and although there is some suggestive evidence of improvements for later cohorts, all estimates are too imprecise to draw conclusions. The analysis in this section does not find evidence for an effect of ELI.

Group: Students who Received ELI Interventions for Two or More Years

Because many ELI students received interventions for two or more years, there is an opportunity to examine whether their pattern of outcomes was similar to the full group of ELI students. In other words, if impacts are stronger for students who received ELI for two or more years, this could represent additional evidence to support its effectiveness. Of course, as stated previously, it could also represent evidence that students who had the opportunity to receive ELI-funded interventions for two or more years were particularly advantaged relative to those who only received interventions for one year. To analyze impact for the subgroup of students who received ELI interventions for 2 or more years, it is only possible to look at impacts for the 2019, 2020, and 2021 cohorts; for the 2022 cohort, the evaluation team would need to be able to draw on 2024 data which is not yet available.

Figure 6 presents difference-in-differences estimates on ELA scores by cohort, for the sub-sample of students who received ELI interventions for at least two years. The figure is consistent with the main results-a positive but not statistically significant coefficient for the 2020 cohort and estimates close to zero for the other cohorts. This analysis does not support the hypothesis that students who received ELI interventions for a longer time period saw improved outcomes.

Figure 6 - Impact Estimates on ELA Assessment Scores for Students Who Received ELI Interventions for Two or More Years, 2021-22 to 2022-23
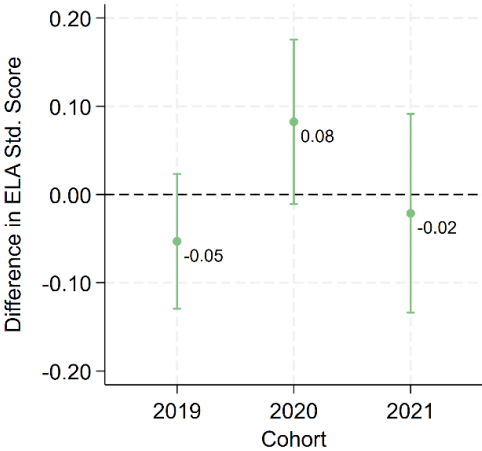
Note: * p<.05, ** p<.01, *** p<.001. Figure presents point estimates and 95% confidence intervals from regression-adjusted difference-in-differences estimates. See Appendix C for additional details on regression analysis.

In Figure 7, Panel A shows short-term, regression-adjusted impacts at time 3 (which reflects 2 years of ELI exposure) on grade repetition for the three cohorts of students with available data. In contrast to findings for the main results, the figure suggests there was a positive and statistically significant impact (grade repetition increased) for the 2020 cohort, with no differences for other cohorts. Panel B shows short-term, regression-adjusted impacts at time 2 (which reflects 2 years of ELI exposure) on ELA course failures. These estimates are close to zero and not statistically significant.

Figure 7 - Impact Estimates on Grade Repetition and ELA Course Failure for Students Who Received ELI Interventions for Two or More Years, Short-Term
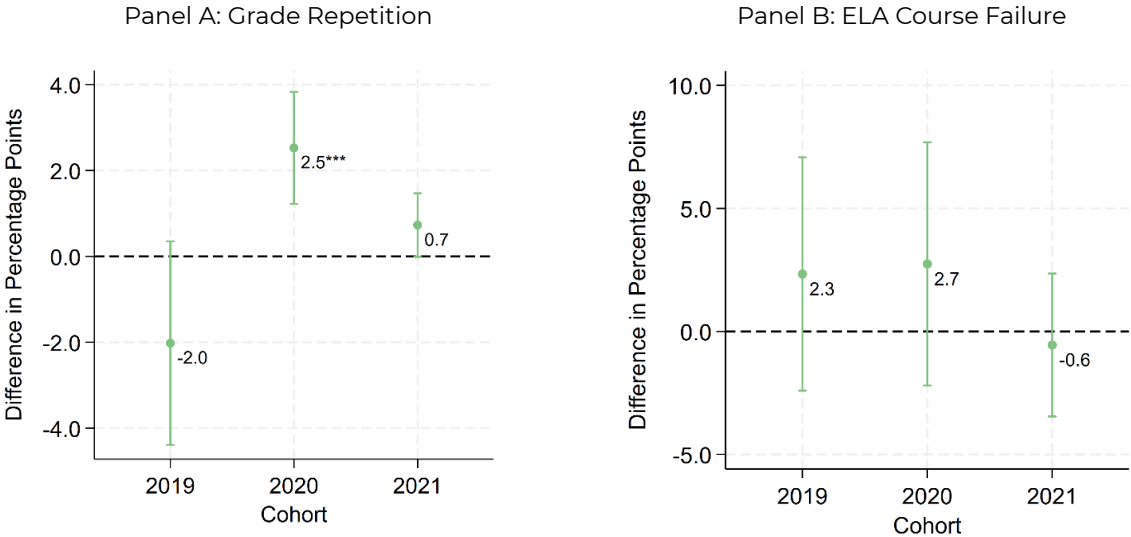
Note: * p<.05, ** p<.01, *** p<.001. Panels A and B present point estimates and 95% confidence intervals from regression-adjusted difference-in-differences estimates. See Appendix C for additional details on regression analysis.

Summary:

This section examines outcomes for ELI students who received interventions for two or more years. Perhaps puzzling, the analysis suggests that ELI actually increased grade repetition after initially reducing it (as shown for the full group of students), for the 2020 cohort, a finding that remains significant at the 5% level after a multiple hypothesis test correction (see Appendix E). This finding is discussed further in the limitations and conclusion section. Most importantly, the evidence in this section does not support the hypothesis that students who received two or more years of ELI-funded interventions saw improved academic outcomes.

Group: Students who Completed and Met Goals

This section analyzes results for the sub-group of students who were recorded as having completed and met the targets/goals set for that intervention, which may be considered as a measure of the extent to which ELI interventions were implemented completely. The comparison group remains students in the same grade and school as ELI students. Figure 8 shows ELA assessment score impacts for each cohort. There is no evidence for a positive effect of ELI on students who completed and met goals.

Figure 8 - Impact Estimates on ELA Assessment Scores for Students Who Completed and Met Goals for ELI Interventions, 2021-22 to 2022-23
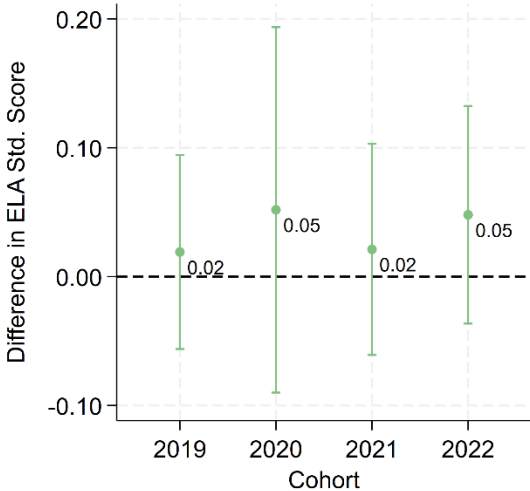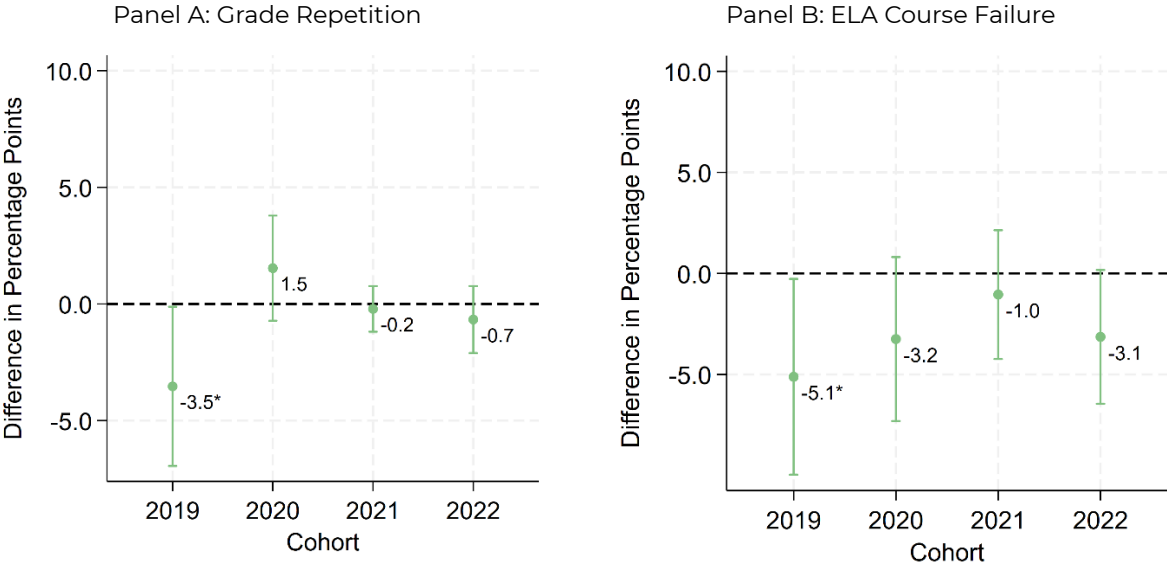

Note: * p<.05, ** p<.01, *** p<.001. Figure presents point estimates and 95% confidence intervals from regression-adjusted difference-in-differences estimates. See Appendix C for additional details on regression analysis.

In the figure below, Panel A shows short-term, regression-adjusted impacts on grade repetition and Panel B shows the same for ELA course failures. Grade repetition reduced by 3.5 percentage points, and ELA course failures reduced by 5.1 percentage points, for the 2019 cohort of students who completed and met goals, estimates which are significant at the .05 level. However, these estimates do not remain significant after adjusting p-values for multiple hypothesis testing (see Appendix E), and there is no evidence that ELI affected academic outcomes for any of the other cohorts.

Summary:

For the group of students who completed ELI interventions and met goals, there is evidence of a statistically significant (at the .05 level) reduction in grade repetition and course failures for the 2019 cohort, with no evidence of a change for other cohorts. However, after adjusting p-values for multiple hypothesis testing (see Appendix E), these estimates do not remain significant. Overall, this analysis does not support the hypothesis that ELI improved academic outcomes for students who completed and met intervention goals.

Figure 9 - Impact Estimates on Grade Repetition and ELA Course Failure for Students Who Completed and Met Goals for ELI Interventions, Short-Term

Note: * p<.05, ** p<.01, *** p<.001. Panels A and B present point estimates and 95% confidence intervals from regression-adjusted difference-in-differences estimates. See Appendix C for additional details on regression analysis.

Group: Schools Grouped by Intervention/Strategy

Another important question is in understanding whether there were differential effects of ELI based on the interventions that schools adopted. In carrying out this analysis, there were too many separate interventions to estimate an average treatment effect using the typical subgroup analysis carried out as above. For example, even if restricting to the five most common interventions, this would increase the number of hypothesis tests substantially and greatly increase the risk of Type I error. This analysis would also be statistically underpowered to detect effects, and therefore, to produce confidence intervals.

Instead, the following approach was adopted. ELI schools were ranked in terms of the following:

- Highest (descriptive) average difference in differences value for change in MCAP scores among intervention sample students from 2022 to 2023.

- Greatest reduction in average grade repetition following the start of ELI interventions.

- Greatest average reduction in ELA course failures following the start of ELI interventions.

For each of the above separately, schools were ranked from 1 to 53. The average rank across schools in terms of the strategy/intervention adopted was calculated. Taking into account the number of students (i.e., not including interventions that were administered to a comparatively small number of students), the following patterns are highlighted:

Schools implementing Sound Partners, Lexia, Being a Reader/Writer and Stepping Stones saw the largest average increase in MCAP scores for ELI students compared to their peers.

Schools implementing Stepping Stones, Tutoring/Personalized Instruction, and Sound Partners saw the largest average reduction in grade repetition for ELI students compared to their peers.

Schools implementing Tutoring/Personalized Instruction, Being a Reader/Writer, and Lexia saw the largest average reduction in rate of ELA course failure for ELI students compared to their peers.

With the above patterns highlighted, it is important to note that this analysis is not intended to suggest that the interventions themselves are what improved subsequent outcomes for students, only that among our sample of students and schools, these were the interventions adopted by schools that saw the largest average subsequent improvement in key outcomes.

Summary: Schools adopting Sound Partners, Stepping Stones, tutoring/personalized instruction, Being a Reader/Writer, and Lexia had students who experienced the largest average improvement in outcomes, although particular caution is encouraged in ascribing causality in this case.

Limitations

The following are some key limitations with regard to the above analysis:

There is a lack of standardization to the data used to select students to receive ELI-funded interventions – Another important limitation is that there is no comparable data on the selection of students to receive ELI-funded interventions. This means that a group of students who are similar in terms of their literacy proficiency but did not receive interventions cannot be identified for the purposes of evaluating impact. This evaluation opts to compare outcomes for ELI students and all students in the same grade (and school), but to assume that ELI caused these differences requires assuming that outcomes for the two groups of students would have moved in parallel (see Appendix B), regardless of whether ELI had been administered. This should be considered a relatively strong assumption.

Inability to fully account for the effects of the pandemic and related disruption to schooling – This evaluation takes two steps to account for disruptions caused by the pandemic. For one, it focuses on a within-school analysis for the impact evaluation, attempting to limit any between-school differences in school closures, instructional modality and other disruptions. Secondly, it analyzes impacts separately for each school year cohort, with the common understanding that material disruptions caused by the pandemic were most severe during the 2019-20 and 2020-21 school years. Nevertheless, there is no way to fully account for the effects of the pandemic and so this should be considered an important limitation.

ELA assessment data is only available from 2021-22 to 2022-23 - An important limitation is with the evaluation of ELA assessment outcomes, which are only available for the 2021-22 and 2022-23 school years. Because ELI funding was administered for each of the 2018-19 through 2021-22 school years, each cohort of students in the evaluation sample varied with respect to the time lag to outcome measurement. Consider the 2018-19 cohort as an example. There is no evidence that their scores improved relative to their peers from 2021-22 to 2022-23. Is that because ELI-funded interventions did not improve their ELA proficiency, or is it because the four-year time lag means interventions had an effect but "faded out"? There is no way to know this based on available data, and so the time lag between receiving ELI-funded interventions and the availability of ELA assessment data is considered a limitation.

The sample of students with available data changes based on outcome - It is best practice to identify a single group for whom to analyze multiple outcomes, but due to the data availability and the imperative to measure outcomes for all students, this evaluation was unable to follow this. Any further analysis of this program may consider drawing on future data availability to identify a single group of students (for example, all ELI students by the end of grade 3) for whom to measure multiple outcomes.

Examining robustness to alternative data and modeling choices can offer additional insights – The evaluation team has described in detail the rationale for choices and assumptions necessary for the present analysis, and tested those assumptions where possible. Future analyses of this data may consider ways to further probe and examine these.

Conclusion and Discussion

This evaluation provides evidence for impact and for process implementation of ELI-funded interventions.

In terms of process, the evaluation finds:

- Schools implemented 16 separate interventions to students in grades Pre-K through 5 over the four years of the grant, and the number and type of interventions varied both across schools and within schools by grade.

- Interventions broadly aimed to improve early literacy skills such as letter-sound recognition, phonics, and reading comprehension.

Based on a review of evidence aggregation websites (What Works Clearinghouse and Evidence for ESSA), most students received interventions that have shown through experimental and quasi-experimental research to have strong evidence of a positive impact on early literacy outcomes.

Students selected to receive ELI-funded interventions had significantly lower academic proficiency as measured by their KRA Language and Literacy domain score, and to a lesser extent, higher prior ELA course failure rates, compared to their peers in the same grades and schools. ELI students were also significantly more likely to receive FARMS.

In terms of impact, the evaluation finds:

- There is no compelling evidence to suggest ELI-funded interventions improved rates of grade repetition, ELA course failures, or ELA assessment outcomes for students, relative to their same-grade (and school) peers. Although grade repetition improved for one cohort (2019-20), there were not statistically significant differences for any of the other three cohorts after accounting for other school-, grade-, and individual-level factors.

- Restricting the sample of students to those who entered Kindergarten with below-average literacy assessment scores, which more closely aligns to the criteria for selection to receive ELI-funded interventions, does not change the conclusions of the impact evaluation.

- There is no compelling evidence to suggest ELI-funded interventions improved outcomes for students who received them for two or more years, or for students who completed interventions and met the individual targets set for them.

- Although there is some descriptive evidence to suggest ELI students improved relative to their peers, in general, estimates are not statistically significant and/or robust to the inclusion of alternative school-, grade-, and individual-level factors that could have driven outcome differences for students.

What explains the lack of an impact for these grant funded programs, when the same kinds of programs have shown evidence of success in the past? The most immediate explanation is that the COVID-19 pandemic and associated disruptions prevented schools from implementing programs in full. It is also important to recognize that in many previous studies, a single intervention has been implemented by one partner, whereas with ELI it is often a package of interventions which are together funded by a grant program. Put another way, differences in implementation and the scale at which interventions are administered may be related to differences in program effectiveness. It's also important to note this evaluation draws on relatively coarse outcome data; what exactly is being measured by course failures for 1st and 2nd graders? Impact evaluations of interventions cited often looked at within-year outcomes and were more closely aligned to the skills targeted by interventions. Lastly, technical aspects (see Limitations) section, perhaps it is statistically under-powered as evidenced by relatively wide confidence intervals. For example, this evaluation could not rule out test score effect sizes as large as 0.05 – 0.08 standard deviations, which would be considered a modest effect of an early literacy intervention.

Recommendations

Even in the midst of the pandemic, LEAs and schools should be commended for deploying evidence-based literacy interventions and effective targeting of funds toward schools and students in need of support. The following is a list of recommendations based on the conclusions of this evaluation.

The following are recommendations to aid future _monitoring_ efforts by MSDE in support of LEAs' and schools' efforts to evaluate their individual program goals and outcomes:

Improve data analysis at the school-level: Schools and districts should consider ways to more rigorously assess whether their own programs are having an impact. Across LEAs, it was common for reports to suggest that programs worked because ELI students improved in specific literacy skills over a particular time period. While improvement is positive, this information is insufficient for understanding whether ELI in fact improved student outcomes; to understand impact, improvement must be measured in reference to a comparison group of students, ideally those who share similarities such as prior literacy proficiency with intervention students. One exception for ELI was Montgomery County, who used MAP data to assess the progress of ELI students during 2021-22 in comparison to a group of non-ELI students. This kind of comparison can be the basis for a more rigorous understanding of whether ELI interventions worked.

More guidance for implementing early literacy programming: Some LEA reports requested more guidance and oversight in implementing grant programs and activities. For example, some LEAs requested additional help in using data to identify students for participation in program activities, and one LEA suggested it would have been helpful to have a template spreadsheet that would have helped them to track students in a uniform manner. Other LEAs stated greater availability of trainings in implementing program activities would have been beneficial. It was also observed by the evaluation team that there may have been differences across LEAs in terms of how students were defined as having completed an intervention, and/or whether students met targets. In this regard, it is promising that MSDE has been convening "literacy labs" to work on these kinds of issues in support of LEAs' implementation of programs associated with the Ready to Read Act.

- The following are recommendations for future _evaluation_ work led by MSDE in support of district- and school-level partners:

Decide on outcomes and hypotheses prior to program implementation: In an ideal scenario, program evaluations are designed with strong a priori knowledge of the design and implementation of interventions, so as to be able to most rigorously and convincingly answer research questions. In reality, the difficulties in coordinating between state, district, and school-level staff over the length of time required to complete an intervention, realities concerning resources and budget, and limitations in available data sources present a large number of challenges to the proper execution of a program evaluation. Future evaluation efforts should aim to devote more effort to designing evaluations from the outset of program implementation, including deciding which outcomes to analyze and which hypotheses to test. This kind of informal "pre-registration" can improve transparency, facilitate greater buy-in and improve shared trust among implementers, families, and district- and state administrators.

Consider greater standardization in approaches to screening: For the purposes of a comprehensive evaluation, some approach to standardization of screening tools across LEAs and schools would be beneficial. This could involve requiring all schools to use the same screening tool, but it could also involve working with schools to collect screening data and examine (ex-ante) some feasible approach

to standardizing data. The value of this approach is in having a well-defined comparison group for intervention students so as to provide a summative measure of impact.

REFERENCES

Abell Foundation. (2021). Literacy Tutoring for Baltimore: What we know, where we are, and how to move forward. Abell Foundation: Baltimore, MD. Retrieved from https://abell.org/publications-library/.

Angrist, J.D., and Pischke, J.S. (2009). Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press.

Benjamini, Y., Krieger, A., and Yekutieli, D. (2006), "Adaptive Linear Step-Up Procedures That Control the False Discovery Rate," Biometrika, 93, 491–507.

Gelman, A., and Loken, E. (2013). The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time. Available at: http://www.stat.columbia.edu/~gelman/research/unpublished/p_hacking.pdf

Hill, C.J., Scher, L., Haimson, J., & Granito, K. (2023). Conducting implementation research in impact studies of education interventions: A guide for researchers. (NCEE 2023-005). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Ladd, H.F., Muschkin, C.G., & Dodge, K.A. (2014). From Birth to School: Early Childhood Initiatives and Third-Grade Outcomes in North Carolina. *Journal of Policy Analysis and Management, 33*(1).

Olson, L. (2023). The Reading Revolution: How States are Scaling Literacy Reform. FutureEd: Georgetown University. Available at: https://www.future-ed.org/teaching-children-to-read-one-state-at-a-time/

Reardon, S.F., Valentino, R.A., & Shores, K.A. (2012). Patterns of Literacy among U.S. Students. *The Future of Children*, *22*(2).

Schochet, P.Z. (2008). Technical Methods Report: Guidelines for Multiple Testing in Impact Evaluations. IES: Washington, DC.

Strunk, K.O., Wright, T.S., Kilbride, T., Zhu, Q., Cummings, A., West, J., Turner, M., De Voto, C. (2021). Michigan's Read by Grade Three Law: Year One Report. Education Policy Innovation Collaborative, Michigan State University.

Wasserstein, R.L., and Lazar, N.A. (2016). The ASA Statement on p-Values: Context, Process, and Purpose. *The American Statistician, 70*(2).

Westall, J., and Cummings, A. (2023). The effects of early literacy policies on student achievement. Education Policy Innovation Collaborative, Michigan State University.

Appendix

APPENDIX A

Supplementary Tables and Figures

Appendix A Table 1 - Skills and Assessments Used in Screening

LEA	Criteria	Assessments
Baltimore City		DIBELS Text, Reading, Comprehension Wilson Fundations
Carroll	Reading below grade level	
Charles	Reading below grade level	
Kent	Lowest 20th percentile	NWEA MAP
Montgomery	Early Warning Indicators Below grade level benchmarks Data Meeting with grade level teams, school literacy team/admin, ELI Point of Contact ESOL Prior participation in literacy intervention program Teacher Recommendation Letter Identification	District assessments
Somerset	Title I students not meeting expectations on multiple measures	ELA KRA MCACP State and local NWEA MAP
Washington	Not attaining 4 or higher on Concepts of Print Not meeting MOY IGDIs Not meeting MOY DIBELS benchmarks	Concepts of Print (PreK) MyIGDIs (PreK) DIBELS (K/1st)
Worcester	Initial placement Early warning indicators	

Appendix A Table 2 - Intervention Description

Early Literacy Intervention	Description
Sound Partners	An intervention that is designed for paraprofessionals to use with students to acquaint or re-teach foundational reading skills. Lesson sets are designed to produce improvement in alphabetics and letter and word identification, using an explicit and balanced phonics approach resulting in scaffolded storybook passages that demonstrate understanding.
Reading Partners	Reading Partners is a one-to-one tutoring program delivered by community volunteers who are recruited, trained, and supported by Reading Partners staff and AmeriCorps members at a reading center within a school. The program serves students in elementary grades who are reading one to twenty five months below grade level, are fluent in English, and do not have IEPs. Students attend 45-minute individualized sessions twice each week. Each tutoring session follows a consistent structure. The tutor reads to the student, stopping to ask comprehension questions. The tutor then uses specific materials to introduce specific reading skills. Finally, the student reads aloud to the tutor and practices the new skill. The tutor provides feedback and coaching following prompts built into the curriculum. Reading Partners requires partner schools to have a dedicated space for a reading center, usually a classroom. Reading Partners then provides a full-time site coordinator to manage the program and recruits community volunteers as tutors.
Leveled Literacy Intervention (LLI)	A short-term, supplementary, small-group literacy intervention designed to help struggling readers achieve grade-level competency. The intervention provides explicit instruction in phonological awareness, phonics, fluency, vocabulary, reading comprehension, oral language skills, and writing. LLI helps teachers match students with texts of progressing difficulty and deliver systematic lessons targeted to a student's reading ability±.
Tutoring	Three LEAs reported providing early literacy tutoring and/or personalized instruction, either during the school day or after school.

| Early Literacy Intervention	Description
Wilson Fundations	Fundations is a prevention and early-intervention program designed to help reduce reading and spelling failure. The program is aimed at students in grades K–3 and involves daily 30-minute lessons which focus on carefully-sequenced skills that include print knowledge, alphabet awareness, phonological awareness, phonemic awareness, decoding, spelling, and vocabulary development. Fundations is designed to complement existing literature-based reading programs in general education classes, but can also be used in small groups of low-achieving or learning disabled students for 40–60 minutes each day. Students rotate through different targeted interactive activities. The program is based on the principles of the Wilson Reading System±
American Reading Co.	The "100 Book Challenge Program" is a motivational and instructional support program designed to help students improve their level of proficiency in reading. An integral part of the program is the stipulation that students read a minimum of 100 books during a designated period of time. Both students and parents are asked to verify that the books have actually been read. Teachers who participate in the Program are provided with an array of books at various difficulty levels for their classroomΔ.
Stepping Stones	An intervention composed of 25 lessons emphasizing listening conventions, phonological awareness, phonemic awareness, and serial processing (processing information in a left to right format to promote automaticity). Each lesson promotes visual, oral, and written language integrating letters and sounds repeatedly.
Amplify	Provides tutoring services designed to complement the efforts of K-6 students, educators, and families in building reading proficiency and confidence. Students participate in engaging and adaptive lessons with tutors while accelerating their proficiency in foundational learning skillsΔ.
Being a Reader/Writer	Being a Reader follows a continuum of reading development to meet each student at their instructional point of need and take them to their next level of literacy. Each grade level includes either seven or eight instructional units organized around two instructional strands: Reading and Word Study. The strands work together to develop comprehension, fluency, decoding strategies, word analysis, spelling, vocabulary, and independent reading.

Early Literacy Intervention	Description
Lexia	Lexia Core 5 is an adaptive blended learning program that is grounded in technology with a face to face instructional component provided by a paraeducator for students who are in the process of learning to read. Based on student performance in the online program, Lexia recommends and provides customized student resources which focus on instructional areas most necessary for continued literacy progress. The ELI paraeducator extends this effort by working on these areas in a face to face small group environment (3–4 students) to explain and discuss how the skills fit into the process of reading, and check for understanding through demonstration.

Note: Table presents intervention summaries for those identified as having been administered to 20 or more students as part of ELI grant-funded activities. ±This description is taken from the WWC Intervention summary. Δ This description is adapted from the product website.

Appendix A Table 3 - ELI Count of Students by Grade and Year

Year	Pre-K	K	1	2	3	4	5	Total
	Panel A: All students							
2019	41	170	209	172	105	58	0	755
2020	193	281	396	329	224	51	1	1,475
2021	96	265	370	370	317	122	54	1,594
2022	90	289	487	416	399	141	97	1,919
Total	420	1,005	1,462	1,287	1,045	372	152	5,743
	Panel B: First-time ELI students							
2019	41	170	209	172	105	58	0	755
2020	193	246	296	213	142	37	1	1,128
2021	96	206	295	255	193	98	54	1,197
2022	90	266	411	280	243	111	70	1,471
Total	420	888	1,211	920	683	304	125	4,551

Appendix A Figure 1 - Distribution of KRA Scores by Cohort, Grade, and ELI Status
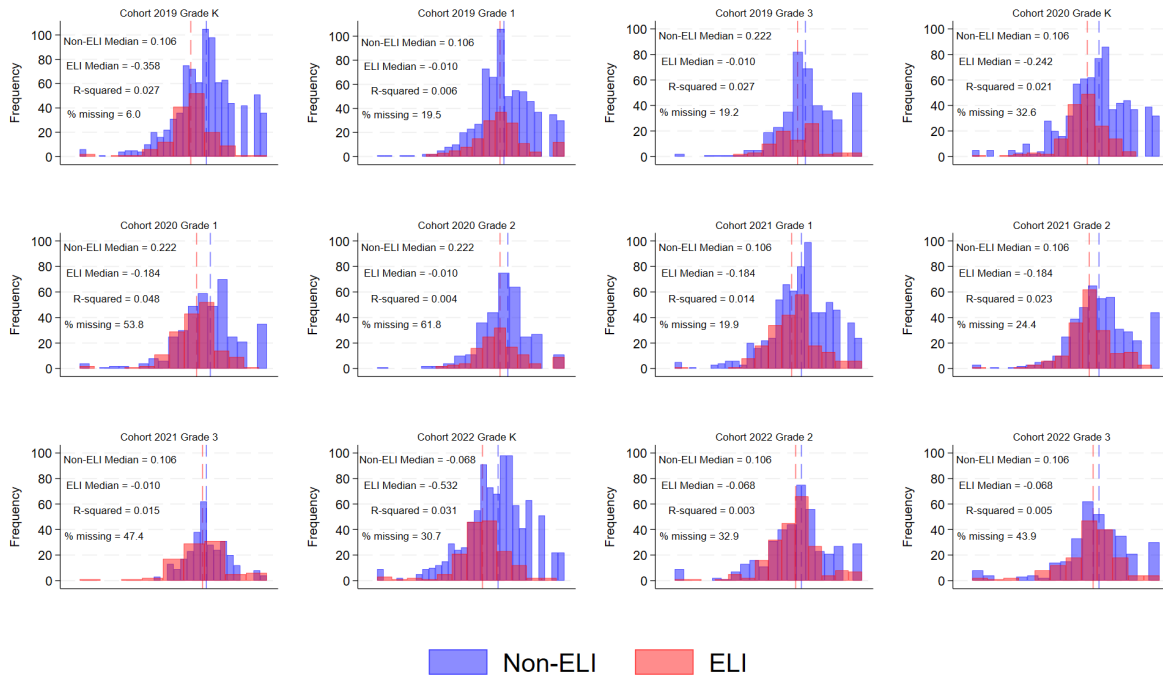
Note: Figure shows distribution of KRA Reading and Literacy domain scores (standardized within year) by cohort and grade. Data for cohorts and grades with greater than 90% missing observations are suppressed.

Appendix A Table 4 - Expanded Summary of Evidence on ELI Interventions

Intervention	What Works Clearinghouse				Evidence for ESSA			Other
	Report Year	# Studies	Sample(s)	Description of evidence*	# Studies	Sample	ESSA Rating	
Sound Partners	2010	7	442 K-1 students in urban schools in Pacific Northwest and Midwest	Medium to large for alphabetics, fluency, and comprehension; small for general reading achievement	2	488	Strong	
Leveled Literacy Intervention	2017	2	747 K-2 students in 3 school districts across 3 states	Medium to large for general reading achievement and small for reading fluency and alphabetics	2	566	Strong	
Wilson Reading System	2007	1	70 3rd grade students in Pennsylvania	Small for alphabetics, fluency, and comprehension	1	158	Strong	

Intervention	What Works Clearinghouse				Evidence for ESSA			Other
	Report Year	# Studies	Sample(s)	Description of evidence*	# Studies	Sample	ESSA Rating	
Lexia Reading	2009	3	314 K-1 students in 2 states	Small for alphabetics, fluency, comprehension, and general reading achievement	3	3971	Promising	
Stepping Stones to Literacy	2007	2	120 K students in 17 elementary schools in the Midwest	Small for alphabetics	-	-	-	
Reading Partners	2016	1	1151 2-5 students in 19 schools in California, New York, and Washington DC	Statistically significant improvements in comprehension, fluency, and sight word efficiency	1	1166	Strong	
100 Book Challenge	-	-	-	-	-	-	-	Yes
Tutoring	-	-	-	-	-	-	-	Yes

Note: * that met WWC evidence standards, with or without reservations

APPENDIX B

Primer on Difference-in-Differences Framework

One quasi-experimental approach to measuring causal effects in program evaluation is using a difference-in-differences framework (Angrist and Lavy, 2009). In short, a difference-in-differences framework attributes deviations in outcome trends between a treated and comparison group to the causal effect of the policy or program. The figure below demonstrates two stylized examples of the method.

In both examples, program schools receive a policy intervention in the 2021-22 academic year. Outcomes for program schools are plotted in dark blue, and for non-program schools in dark red. Under the difference-in-differences framework, a "counterfactual" estimate for program schools (light blue dotted line) is estimated based on the trend for non-program schools.

Panel A shows that the average attendance rate for program schools increased from 92% to 95% between the 21-22 and 22-23 academic years, while it stayed at 90% for non-program schools. The difference-in-differences impact estimate (3-0 = 3 percentage points) can be calculated as the over-time difference for program schools (95-92 = 3) minus the over-time difference for non-program schools (90-90 = 0). Another way to see this is to compare the actual and counterfactual difference in attendance rates in 22-23, which also equals 3 percentage points.

Panel B shows average test scores for the same group of schools. As shown, test scores increased from 520 to 540 (20 score points) for program schools and from 480 to 520 (40 score points) for non-program schools, between 2021-22 and 2022-23. Even though test scores increased for program schools, they did not increase by as much as comparison schools, so the impact estimate in this case is negative (20-40 = -20 points).

Appendix B Figure 2 - Stylized Examples of a Difference-in-Differences Analysis

Panel A: Example 1
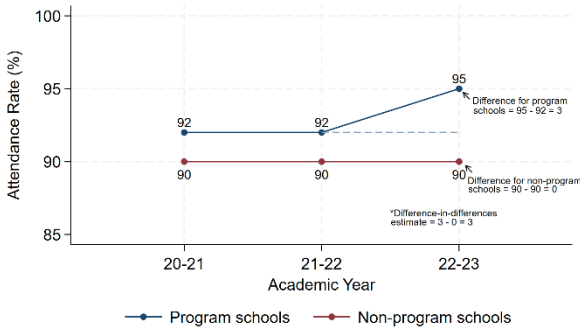
Panel B: Example 2
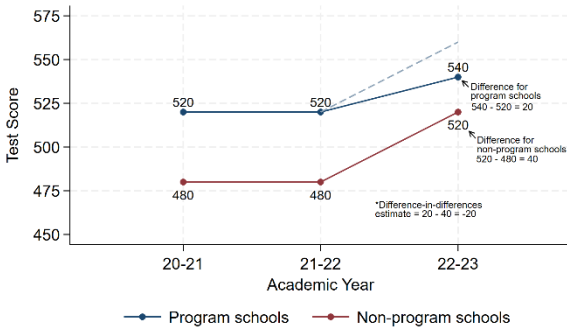
Note: Figures present two stylized examples of the difference-in-differences method.

In this framework, as with any for estimating causal effects, it is important to assess the plausibility of the comparison group. In this case, how likely is it that program schools would have followed the same trends as non-program schools if the intervention had never happened? Detailed knowledge about program and non-program schools, such as other funding/programs that were administered at the

same time, can help in this understanding. Two empirical ways of improving plausibility is by using regression to control for other factors known or suspected to cause differences between groups, and to examine pre-trends. In the above example, the outcome trends between program and non-program schools are parallel in the period prior to the intervention, which lends some plausibility to the assumption that they would have continued to be parallel if there had been no intervention.

APPENDIX C

Information on Regression-Adjusted Impact Estimates

The impact analysis uses the following ordinary least squares (OLS) regression equation to estimate the impact of ELI, for each cohort separately, on the change in ELA scores from 2021-22 to 2022-23:

$$y_{ist} = \alpha + ELI_{is} + \beta([1 = YEAR_{2023}] * ELI_{is}) + \beta X'_i + \gamma_s + \epsilon_{ist}$$

In this equation, the standardized ELA MCAP score for student i in school s at time t is regressed on a dummy variable for whether the student ever received ELI interacted with a dummy variable for the 2023 school year. The fully interacted model is estimated with a vector X of individual-level characteristics, including dummies for gender, FARMS status, race, ethnicity, grade at ELI assignment, as well as standardized KRA reading domain score. Lastly, it includes school (of ELI assignment) fixed effects. Standard errors are clustered at the student level.

The identifying assumption for estimating a causal impact of ELI is quite strong; conditional on individual and school-level characteristics accounted for in the model, no other within-cohort factors are assumed to have driven differences in MCAP scores for students. To improve the plausibility of this assumption, the impact analysis includes estimates of the same equation for additional sub-groups, including students with below-median KRA scores. Still, future analysis should take advantage of additional years of data for ELI students and schools in order to improve the identification strategy.

To estimate the impact of ELI on grade repetition and course failure, the above equation replaces the 2023 year dummy with "event time" dummies which estimate the change over time in outcomes in reference to time 1 (fall of the cohort academic year) for grade repetition and time 0 (spring of the prior academic year) for failure. The short-term estimates are captured in the form of a single β_c for each cohort.

APPENDIX D

Notes on Dataset Construction

To examine questions about student selection and program impact, the evaluation team followed a three step process for constructing a student-level panel dataset using data from MSDE's multi-year data warehouse. First, the school year, school and state-assigned IDs are merged to the EDW data to capture enrollment and attendance histories of ELI students. Next, students in the same grade and year as students who received ELI interventions were identified. Third, the educational histories of intervention and comparison students were obtained back to their pre-K year.

Outcome measures include grade repetition, ELA course failure, and ELA standardized assessment results. Grade repetition is measured using the September enrollment file, and is coded as 1 for students who repeated a grade and 0 for students who did not repeat a grade. Students who entered Maryland schools for the first time have values of grade repetition that are coded as missing. Course information is collected at the end of the school year. A student's value is coded as 0 if they take and complete a full- or school-year term courses with the subject "English Language and Literature", and 1 if they complete but receive a failing status. In addition to the above, we examine differences in reading/ELA assessment outcomes. MCAP ELA outcomes from grade 3 and above are assessed for the 2021-22 and 2022-23 school years, and scale scores are standardized to have a mean of 0 and standard deviation of 1 within grade, for each school year. We also draw on KRA Language and Literacy sub-domain scores for students with available data, standardized within year.

APPENDIX E

Multiple Hypothesis Corrections

When carrying out statistical inference, it is well known that as the number of hypothesis tests for the same or similar populations increases so does the chance of Type I error, or the chance that a finding will be estimated as statistically significant when there is no true effect (Schochet, 2008). Researchers have a variety of multiple hypothesis adjustments they can choose from in order to reduce this risk. To control for the false discovery rate, or the expected proportion of null hypotheses that will be incorrectly rejected, researchers can apply a procedure described in Benjamini, Krieger, and Yekutieli[9] (2001). In doing so, it is assumed that the three outcomes examined in this evaluation (assessment, grade repetition, and ELA course failure) fall under three outcome domains and within each domain it is necessary to account for 15 separate significance tests (four for each year for the main and two of the three subgroups, and 3 for the third subgroup). The results of this analysis suggest that only the two significant findings for grade repetition (the main negative effect for the 2021 cohort, and the positive effect at time 2 for the 2020 cohort) remain after adjusting for the chance of Type I error.

---

[9] This multiple hypothesis adjustment is carried out using the "multproc" command in Stata.

APPENDIX F

Notes on Research Design

Sample Selection and Choice of Comparison Group

The evaluation team draws on data for all students who were selected to receive ELI-funded interventions. The simplest approach is to examine outcomes for this group of students. However, there are two primary reasons for why analyzing impacts by school year is preferable:

There were large differences across school years in terms of the LEAs, schools, and grades that received interventions. The difference in schools is reflected in part by the differences in student characteristics across cohorts (Table 2).

ELI funding was administered during the course of the COVID-19 pandemic, associated school closures, and the switch to virtual instruction. While it is impossible to account for these factors in the data, it is conceivable that these effects were very different depending on the school year.

While additional choices for grouping students, such as by grade level, by LEA, and by type of intervention, could all be reasonable, the number of groups could be quite large and so the evaluation team opts for simplicity in focusing on cohort/academic year.

Having decided on the sample of students and how their information should be presented, the simplest approach to examining impacts on academic outcomes is to compare each cohort's change in outcomes from the year prior to receiving ELI to their outcomes in the year after receiving ELI. For example, if the average rate of grade repetition for the 2019 cohort reduced from 10% in 2019 to 5% in 2020, this would be a good thing, but what assurances would we have that this reduction was caused by ELI? One important step, and indeed a fundamental component of any impact evaluation, is devoting special care and consideration to the choice of a comparison group. The purpose of choosing a comparison group for ELI students is in constructing a "counterfactual"; in other words, how would ELI students' outcomes have changed if the grant funding had never been administered at all? As stated previously, ELI students were chosen based on their perceived risk of not being able to meet literacy proficiency targets by the end of grade 8. Therefore, the simplest comparison group for ELI students would be a group who were also at risk of not meeting proficiency targets, but who did not in fact receive ELI-funded interventions.

Appendix A Figure 1 shows how using data to identify a comparable group of students who were also at risk of not meeting proficiency targets, but who did not receive ELI-funded interventions, would be challenging. The figure shows the distribution of KRA Language and Literacy domain scores for students based on grade and cohort, for the four years of the grant. The figure shows that while ELI students do in all cases have lower scores, on average, than their peer students in the same grade, there is still a large amount of overlap in scores. Further, within each grade and school year, there is a large amount of missing data. Both of these factors suggest that it would be challenging to use data to identify a group of students who were at a similar risk of not meeting proficiency targets as ELI students. Therefore, the following analysis keeps the choice of a comparison group simple, comparing trends in outcomes for ELI students with those of their peers in the same school and grade. In addition to this simple comparison, further analysis will compare the change in outcomes for ELI students to the group of students with below median KRA scores.

Empirical Approach

Having decided on the intervention and comparison groups of students, and made some basic decisions on how the data will be presented, the next step is to decide on an empirical approach for estimating the impact of ELI. As stated, the simplest approach would be to identify a group of students who were at similar risk of not meeting proficiency targets, and compare their post-period outcomes to ELI students. However, given the aforementioned data challenges combined with the fact that we can draw on an extensive database of students in Maryland Public Schools, we opt to compare _trends_ in outcomes for the two groups of students; ELI students and their peers in the same schools and grades. This approach is what is known as a difference-in-differences strategy, and further explanation along with examples is provided in Appendix B. In short, under a difference-in-differences framework, a deviation in outcome trends between the two groups (ELI students and their peers that didn't receive ELI) is assumed to reflect the causal impact of the grant. This is admittedly a strong assumption, although it is argued here to be an improvement over a single group pre-post comparison, as well as a comparison between two groups at a single time point. In addition, we can draw on available data to estimate a difference-in-differences impact using a regression-based framework. Using regression allows us to do two things: 1) control for other factors that are known or suspected to drive differences in outcomes, and 2) estimate uncertainty around the impact estimate; in other words, how confident can we be that the effect is different from zero?