MARZANO RESEARCH

Evidence meets practice.

FY 25 Literacy Tools Approved Vendors

Effectiveness Review

June 2025

Prepared By
Christopher Wolfe, PhD
Abby Andres, PhD

FY 25 Literacy Tools Approved Vendors

Effectiveness Review

Prepared For

Idaho State Board of Education

By

Christopher Wolfe, PhD, and Abby Andres, PhD

Referencing This Report

This report should be cited as:
Wolfe, C., & Andres, A. (2025). *FY 25 literacy tools approved vendors effectiveness review*. Marzano Research.

Disclaimer

Marzano Research prepared this report to present results of an independent evaluation of the effectiveness of the vendors on the Idaho Literacy Tools Approved Vendors List. This report does not represent an endorsement of any product.

Contents

Tables and Figures

Overview

On July 1, 2020, Section 33-1616 of Idaho Code was amended to establish a Request for Proposal (RFP) process for identifying "adaptive learning technology" literacy products eligible for placement on the state's Literacy Tools Approved Vendors List. In this context, adaptive learning technology refers to products delivered exclusively via computer or web-based platforms. In 2021, this section was incorporated into the Idaho Literacy Achievement and Accountability Act and moved to Section 33-1807 of Idaho Code.

Sub-section (2)(b) of the statute mandates that Idaho school districts use products from the Approved Vendor List if they intend to use the product as a "comprehensive program" for literacy interventions. Sub-section (3) provides the criteria and review process for inclusion and continued placement on the Approved Vendor List:

Section 33-1807(3)

a. The state board of education shall select adaptive learning technology literacy intervention providers through a request for proposals process to provide adaptive learning technology literacy intervention tools for school districts and charter schools to use as part of their literacy intervention programs for students in kindergarten through grade 3 that:
    i. Include an academic program focused on building age-appropriate literacy skills that, at a minimum, include phonological awareness, phonics, fluency, comprehension, and vocabulary;
    ii. Use an evidence-based early intervention model;
    iii. Include a parental engagement and involvement component that allows parents to participate in their student's use of the tool at school or at home; and
    iv. Address early reading and literacy intervention through the use of an interactive and adaptive computer software program.

b. To remain on the Approved Vendor List after the first year of identification, programs must be evaluated each year to determine effectiveness by an independent external evaluator. The evaluation will be based on a full academic year of implementation of tools implemented with fidelity and will include, at a minimum, growth toward proficiency measures.

Background on the Review Process

In July 2020, the Idaho State Board of Education launched an RFP process aligned to this statute. The Office of the State Board of Education (OSBE) has provided guidance to support ongoing product submissions and revisions. Vendor responses are reviewed by trained volunteers recruited from Idaho's K–3 educator community. These in-the-field educators receive targeted training and use a standardized rubric to assess vendor proposals for completeness, alignment with Idaho Code, and readiness for use in literacy interventions.

To remain on the Approved Vendor List, vendors must submit annual effectiveness data to OSBE for review. This includes:

- A vendor-provided summary of aggregate student progress using their product
- A secure file listing K–3 Idaho students who used the product in the 2022–23 school year, which OSBE then matches with Idaho Reading Indicator (IRI) results from Fall 2023 and Spring 2024

Role of Marzano Research

In alignment with the statutory requirement for an independent evaluator (Section 33-1807[3][b]), the Idaho State Board of Education contracted with Marzano Research to conduct the FY2025 Effectiveness Review. Marzano Research was selected due to the organization's deep expertise in educational research, literacy evaluation, and rigorous program analysis. The organization has led similar statewide and regional evaluations and brings extensive experience in analyzing student outcome data in the context of implementation fidelity and program design.

Marzano Research received de-identified, matched student-level data and conducted the following analyses:

- Distribution of Fall and Spring 2023–24 IRI scores by grade level
- Change in proficiency rates (percentage of students at grade-level) from Fall to Spring, disaggregated by vendor usage group
- Average raw IRI scale score changes over the academic year

Where possible, vendor data was compared with statewide patterns. Based on these analyses, Marzano Research provided evidence-based recommendations to OSBE regarding the continued effectiveness of each product and its eligibility for inclusion on the Approved Vendor List for the 2025–26 school year. This included recommendations for:

- **Remain on List:** Both the vendor-submitted report and matched IRI results for participating students indicate that forward progress is being made at levels meeting or exceeding state average expectations
- **Reinstate to List - Probationary:** Vendor-submitted report indicates rationale for inclusion, but due to data limitations either because of data quality or student usage time, matched data cannot fully validate that forward progress is being made at levels meeting or exceeding state average growth expectations; if data issues and student usage time are insufficient in subsequent years, Marzano Research recommends excluding program from the list.
- **Remove from List - Insufficient Data:** Limited numbers of students and limited student usage time mean evidence for inclusion is not supported by the data; Marzano Research recommends re-evaluating when another year of data is available.

Data and Methodology

To assess the effectiveness of literacy interventions across Idaho, OSBE collaborated with approved vendors to obtain student-level usage data for each program. This included vendor-submitted reports detailing progress Idaho students made within each vendor's proprietary assessments, as well as student-level data detailing the amount of time students spent actively engaged in each program. This vendor-submitted student data was securely matched by OSBE staff to statewide records, including Fall and Spring Idaho Reading Indicator (IRI) scores. Where possible, deterministic matching using unique student identifiers was employed. In cases where IDs were unavailable, records were matched using a combination of student first and last names, grade level, local education agency (district), and school. Despite these efforts, not all records could be successfully matched due to data quality issues, including misspellings, use of initials, or inconsistent formatting, leading to some attrition in the final analytic samples.

In rigorous research settings, ideal match rates typically exceed 85%, particularly when unique student identifiers are available. However, in operational settings such as this—with variable data quality and reliance on indirect identifiers like names and district-grade combinations—match rates may reasonably range from 60% to 80%. Lower match rates should be interpreted with caution but can still be used to understand general trends.

Matched datasets were returned to Marzano Research for analysis. The primary goal was to examine the relationship between time spent in literacy interventions and growth on the IRI, leveraging research-informed thresholds for program fidelity. Marzano Research grouped students based on whether they met vendor-recommended time benchmarks, the state's benchmark thresholds of 30 and 60 cumulative hours, or neither. Even when students did not reach these benchmarks, Marzano Research conducted analyses using time-based buckets (e.g., 0–10 hours, 10–20 hours, etc.) to investigate whether increasing time in product was associated with greater literacy gains. This dose-response approach provides valuable insight into the potential for partial implementation to yield benefits and helps identify the point at which interventions begin to show measurable impact.

Average time students spent actively working in each product provide an insight into implementation fidelity and allow each product to be evaluated according to use. In cases where average time is low, it is likely that implementation protocol can be strengthened to return better data in future years.

Finally, student IRI outcomes were compared to average IRI growth trends across grade bands statewide to contextualize findings and assess the relative impact of each product. This methodologically rigorous approach allows for data-driven recommendations about which interventions are most effective in supporting early literacy development, and under what conditions they are most likely to succeed.

Vendors Reviewed

The following vendors submitted data for the FY2025 Effectiveness Review:

- Lexia Learning
- Imagine Learning
- Renaissance (Freckle ELA)
- Curriculum Associates (i-Ready)
- Savvas (SuccessMaker)
- MobyMax
- Reading Horizons
- Amira (Previously Istation)
- Waterford (Reading Academy)

This report presents the findings and recommendations from the FY2025 Effectiveness Review, using student outcome data from School Year 2023–24. Vendors remain on the Approved Vendor List while the review is in progress. The recommendations included here are intended to support final determinations regarding vendor eligibility for the 2025–26 school year.

Executive Summary

Data for all nine vendor products was analyzed across four categories to return recommendations in the following dashboard:

Vendor Evidence: Did the vendor-submitted report provide reasonable evidence of student growth within their own product? This is a yes/no decision.

Data Quality: Was the data the vendor-submitted able to be matched to Idaho Student data and IRI scores? Vendors were rated as follows:

- **Good:** Match rate of 85% or higher
- **Adequate:** Match rate of 60%–84.99%
- **Poor:** Match rate of less than 60%; data for vendors in this category should be evaluated with extreme caution.

Usage: Are students spending sufficient time in product to meet the intent of the Approved Vendor List? Because usage rates varied widely across vendors, and even within vendor products, a wider range of categories was used to evaluate results against average time for students who spent at least 30 minutes in a given vendor's product. Students spending less than 30 minutes were excluded from this calculation. Categories were created as follows:

- **Very High:** Average time spent in the vendor's product was at or above the 30-hour (1,800-minute) mark.
- **High:** Average time spent in the vendor's product was in the 20–30-hour range (1,200–1,799 minutes)
- **Moderate:** Average time spent in the vendor's product was in the 10–20-hour range (600–1,199 minutes)
- **Low:** Average time spent in the vendor's product was in the 5–10-hour range (1,300–599 minutes)
- **Very Low:** Average time spent in the vendor's product was less than 5 hours (30–299 minutes)

Impact: To what extent does IRI data suggest that the product is helping students meet or exceed growth in literacy skills over the course of the year? Categories were created as follows:

- **Very High:** No significant gaps seen when compared to statewide data; student gains all fall within expected ranges; gains in IRI scores are above state averages for all grades for which there is sufficient data
- **High:** No significant gaps are seen when compared to statewide data; student gains all fall within expected ranges; gains in IRI scores are above state averages for at least one grade for which there is sufficient data

- **Moderate:** No significant gaps are seen when compared to statewide data; student gains all fall within expected ranges
- **Low**: Minor gaps are seen when compared to statewide data; student gains all fall within expected ranges.
- **Very Low:** Significant gaps are seen when compared to statewide data or data suggests that time in product may be correlated with lower IRI scores.

Table 1: FY 2025 Evidence Based Recommendation Dashboard*

Vendor	Vendor Evidence	Data Quality	Usage	Impact	Recommendation
Lexia Learning	Yes	Good	High	High	Remain on List
Imagine Learning	Yes	Good	Moderate	Moderate	Remain on List
Renaissance (Freckle ELA)	Yes	Poor	Very Low	Moderate*	Reinstate to List - Probationary
Curriculum Associates (i-Ready)	Yes	Good	Moderate	High	Remain on List
Savvas (SuccessMaker)	Yes	Good	High	High	Remain on List
MobyMax	Yes	Poor	Very Low	Moderate*	Reinstate to List - Probationary
Reading Horizons	No	Good	Very Low	N/A	Remove from List - Insufficient Data
Amira (Previously Istation)	Yes	Good	Moderate	Moderate	Remain on List
Waterford	Yes	Good	Moderate	High	Remain on List

Note. Data comes from vendor-submitted data that could be matched by OSBE to IRI results.
*Because of data quality concerns, these results should be interpreted with caution.

The majority of vendors had excellent data quality that meet or exceed research best practices benchmarks. However, even within the data, usage rates were a concern for many vendors, with few students hitting vendor recommended time benchmarks and even fewer hitting state-recommended levels of 30-hours or 60-hours of intervention time.

In the future, as vendors work with districts and schools, attention will likely need to be paid to implementation to ensure that students are gaining the benefits from the literacy intervention programs they are being enrolled in.

Lexia, Imagine Learning, Curriculum Associates, and Amira all served a larger number of students. With no significant outliers in the data, we recommend they remain on the list.

Two vendors, Savvas and Waterford, had very small datasets, but very high-quality data that is showing solid usage numbers and positive impacts on student outcomes both within their own system assessments and as related to student IRI scores. We recommend they remain on the list.

In the cases of two vendors, Renaissance, and Moby Max, poor data match and very low usage rates make it difficult to say whether the positive impacts being seen are related to the product or other variables we cannot see in the large amount of missing data. We recommend these two vendors be reinstated on the list with probationary status, with the expectation that the results of their FY 26 Effectiveness Review must be improved for them to remain on the list.

One additional vendor, Reading Horizons, admittedly struggled with implementation this year, and neither their data nor the matched data is sufficient to make the recommendation that they remain on the list.

Table 2 provides some of the summary data used to help make these decisions. More complete data follows for each vendor in vendor specific analysis sections.

Table 2: FY2025 Summary Data

Vendor	Match Rate % (n-size)	% Reported as Meeting Vendor Recommended Usage Targets	% Met State-Recommended 30 hours	Average Time in Intervention (minutes)	Average Change in IRI Score (Fall to Spring)
Lexia Learning	99.5% (10,011)	55.08%	40.61%	1,691	69.2
Imagine Learning	92.6% (5,115)	13.69%	6.41%	609	62.7
Renaissance (Freckle ELA)	53.2% (1,616)	6.72%	0%	170	71.4
Curriculum Associates (i-Ready)	99.9% (13,442)	42.38%	6.91%	946	67.0
Savvas (SuccessMaker)	100% (188)	87.23%	9.57%	1,324	63.3
MobyMax	51.51% (770)	0%	0%	128	66.8
Reading Horizons	98.4% (123)	0%	0%	105	71.5
Amira (Previously Istation)	99.2% (87,638)	0%	7.89%	903	70.1
Waterford	92.6% (1,259)	20.92%	12.77%	893	70.8

Note. Data comes from vendor-submitted data that could be matched by OSBE to IRI results.

Overview of the Idaho Literacy Intervention Program

The Idaho State Literacy Program seeks to provide technology-based, adaptive literacy tools for K–3 students across the state, with an emphasis on building foundational skills such as phonological awareness, phonics, fluency, vocabulary, and comprehension. The 2020 RFP process invited vendors to propose tools to support differentiated instruction both in the classroom and at home. Selected programs are intended to serve a wide range of learners—including English Language Learners (ELLs), students with disabilities, and students reading below grade level—while offering professional development and data-driven tools to help teachers personalize instruction.

Common themes across vendor applications include:

- Focus on foundational reading skills, personalized/adaptive learning pathways, and use of embedded or ongoing assessments
- Daily use recommendations ranging from 15 to 60 minutes, usually 3–5 days per week
- Availability for both classroom and at-home learning environments
- Strong support for ELLs and students with disabilities
- Pricing structures typically range from $20 to $60 per student per year, with additional costs for professional development and implementation services

Table 3 provides a summary of program offerings by vendor.

Table 3. Idaho Literacy Program Vendor Programming Details

Vendor	Grades Served	Key Skills Included	Recommended Usage	Cost Per Student	Special Population Research Focus
Lexia Learning	PreK–5 (Core5); 6–8 (PowerUp)	Phonological awareness, phonics, fluency, vocabulary, comprehension	20 min/day; 5 days/week	$40–$60	ELLs; Special Education
Imagine Learning	K–3 (up to grade 6)	Language acquisition, phonics, comprehension, vocabulary	20–40 min/day	$25–$35	ELLs
Renaissance (Freckle ELA)	K–3 (up to 12)	Phonics, fluency, vocabulary, comprehension	15–30 min/day	$25–$45	Adaptive
Curriculum Associates (i-Ready)	K–3 (up to 8)	Phonics, comprehension, vocabulary, fluency	15–30 min/session; 3–5x/week	$30–$50	ELLs; Special Education

Vendor	Grades Served	Key Skills Included	Recommended Usage	Cost Per Student	Special Population Research Focus
Savvas (SuccessMaker)	K–8	Phonics, comprehension, fluency	20–30min/day	$50–$60	Adaptive
MobyMax	K–8	Full literacy suite (phonics, comprehension, vocabulary, fluency)	20 min/day	$5–$16	ELLs; Special Education
Reading Horizons	K–3 (Discovery); 3+ (Elevate)	Phonics (Orton-Gillingham), fluency, comprehension	20–30 min/day	$40–$60	Dyslexia, English Learners
Amira (Previously Istation)	K–3 (up to 8)	Phonemic awareness, phonics, fluency, comprehension, vocabulary	30 min/day; 3–5x/week	$29–$60	ELLs; Special Education
Waterford	PreK–2	Oral language, phonics, comprehension, vocabulary	15–20 min/day	$100+	ELLs; Low Income

Note. Data comes from vendor applications and annual reports submitted by each vendor.

The Idaho State Board of Education's literacy initiative showcases a strategic investment in adaptive, evidence-based reading tools aimed at improving outcomes for early learners statewide. The 2023–24 evaluation cycle included a comparative analysis of vendor platforms approved for K–3 literacy instruction. All programs evaluated are aligned to Idaho's educational priorities: targeting foundational literacy skills, supporting students across a range of ability levels, and providing educators with real-time insights through embedded assessments. Usage recommendations generally ranged from 15 to 30 minutes per day, and while cost per student varied significantly across vendors, all included core access and reporting features, with optional add-ons for professional development and expanded services. Many programs offer dual delivery modes, allowing for flexible use both at school and at home. Importantly, multiple vendors provided data disaggregated by student subgroups (e.g., ELLs, Special Education, below-grade readers), illustrating differential impact and highlighting tools best suited for particular populations. The detailed vendor comparisons and effectiveness summaries underscore the importance of both program quality and implementation fidelity in achieving measurable literacy growth.

Professional Development, Research, and Implementation Fidelity

Beyond content and pricing, program success often hinges on strong support structures, credible research, and practical implementation models. Most vendors offer tailored professional development options, from embedded teacher coaching to webinars and district-specific rollouts.

Several programs, including i-Ready, Imagine Learning, and Lexia Core5, are backed by peer-reviewed efficacy studies or ESSA-aligned research that demonstrates statistically significant student gains. Implementation plans vary widely; some emphasize turnkey onboarding and family engagement (e.g., Waterford), while others stress structured success management and regular fidelity monitoring (e.g., Savvas, Amira). Nearly all programs reinforce the correlation between consistent student usage and learning gains—typically citing thresholds of 15–30 minutes per day, 3–5 days per week—as critical to impact. Notably, products like Amira and Imagine Learning applied to be included in the program and provided evidence of clear, tiered growth across usage levels and demographic subgroups, affirming their capacity to serve both struggling and high-achieving readers. As Idaho continues scaling this initiative, these insights will inform decisions about vendor renewal, targeted support, and equitable student access at local levels.

Table 4. Idaho Literacy Program Vendor Research and Implementation Overview

Vendor	Professional Development Approach	Research Backing Provided	Implementation Model
Lexia Learning	Customer Success Manager, on-site and remote PD	Peer-reviewed studies, efficacy reports	Phased rollout with CSM guidance
Imagine Learning	Dedicated Customer Success Manager, customized onsite and virtual PD	State and district reports, longitudinal data	District-specific plans and data consults
Renaissance (Freckle ELA)	Webinars, training tools, account managers	Case studies, correlational studies	Quick start onboarding and dashboards
Curriculum Associates (i-Ready)	Comprehensive PD, coaching, support plans	ESSA Tier II and III studies, internal data	Structured plans and fidelity monitoring
Savvas (SuccessMaker)	Scheduled training, implementation specialists	District case studies, EISP evaluations	Admin-driven fidelity plans and data reviews
MobyMax	Basic training included, extended PD optional	ESSA Tier I study, internal and external research	Self-paced or supported, scalable
Reading Horizons	Embedded coaching, customized PD plans	ESSA-aligned, external evaluations	Facilitated coaching, fidelity emphasis
Amira (Previously Istation)	Live and recorded PD, tiered support plans	Johns Hopkins study, district results	Onboarding, usage tracking, success plans
Waterford (Reading Academy)	Blended PD, family engagement training	Multiple state and external evaluations	Turnkey, parent-centered, remote-first

Note. Data comes from vendor applications and annual reports submitted by each vendor.

2023–2024 Vendor Outcomes

The following section provides a vendor-by-vendor analysis of Idaho's K–3 literacy interventions. Each subsection begins by presenting the outcomes reported by the vendor, typically based on their internal assessments or benchmark tools embedded within their platforms. These vendor-reported outcomes are then compared with student growth on the Idaho Reading Indicator (IRI), a statewide standardized assessment of early literacy. This two-pronged approach allows for both validation of vendor claims and an independent analysis of effectiveness based on a common outcome measure. By comparing reported and observed results, we aim to determine the degree to which each program contributes to literacy growth when implemented in Idaho schools.

Each section begins with the high-level recommendation, key findings, and rationale for the recommendation, and is followed by both vendor-reported efficacy data using vendor-submitted reports about their own assessments and data, and matched IRI data to help validate vendor measures of success. Only records for students who could both be matched to state data and who tracked time in each vendor's systems are included in IRI score gain and time in product analyses.

For all vendors, the following data was analyzed, as possible:

- Fall and Spring IRI score distribution, all students, by grade
- Change in percentage of students proficient in Fall and Spring and the change in proficiency rate between the two (in percentage points) for students who have both scores, by usage group.
- Average raw scale score changes from Fall 2023 to Spring 2024 for students who have both scores, by usage group, both overall and as compared to state averages
- Average time spent in each product/program by grade as compared to vendor expectations for efficacy and state recommendations of 30 and 60 hours for students beginning the year below grade level on the IRI

Data Considerations

Three issues quickly became apparent in reviewing vendors' 2023–24 data for effectiveness:

1. **Inability to Directly Correlate Students' Progress to the Vendors' Programs**

   In order to appropriately identify the impact of vendors' programs, research would need to be able to control for potential student demographic characteristics (socioeconomics, ELL status, Special Education status, etc.), as well as the impacts of a school's educators (level of experience, professional development received, etc.), core instruction (curriculum, time spent in core instruction, etc.), and implementation fidelity of the vendor program.

   Unfortunately, this is not feasible as it would require an in-depth research plan integrating detailed data from districts and individual educators that is not tracked, nor would it be feasible to track this statewide without creating an undue burden at both the local school and OSBE level.

Thus, while this analysis examines vendor products, it does not attempt to do a comprehensive analysis of the correlation between use of vendors' programs and student results. To mitigate the inability to isolate the impact of vendors' products, Marzano Research staff used comparisons to state data whenever possible, as we can reasonably expect the product to be at least as successful as state averages, and for state averages to fall within one standard deviation of average student scores for students using any given vendor product.

2. Vendors' Difficulty in Providing Adequate Student Information for Matching IRI Results

Though both state statute and the request for proposals (RFP) for the Approved Vendor List specify the requirement for the annual Effectiveness Review, there are no specifics regarding the data that is required or whether a certain sample size is necessary for inclusion in the review. In communicating with vendors regarding the process, it has become clear that most vendors do not require students' EDU IDs to be used as the student identification number (or even be entered) in their systems. Gathering EDU IDs from their client districts after the fact has proven difficult for some vendors. As a result, the match rate varied between vendors. This is noted in the analysis for each vendor. In some cases, the lack of matched data causes the sample to fall below the 60% threshold considered viable for operational data such as this program data. In these cases, recommendations are presented with reservations, as there is less confidence that the results are representative of student experiences within the vendor platforms.

OSBE staff is continuing to encourage vendors to identify processes to ensure that EDU IDs are available for all students on their roster for a given school year.

3. Potential Implementation Concerns

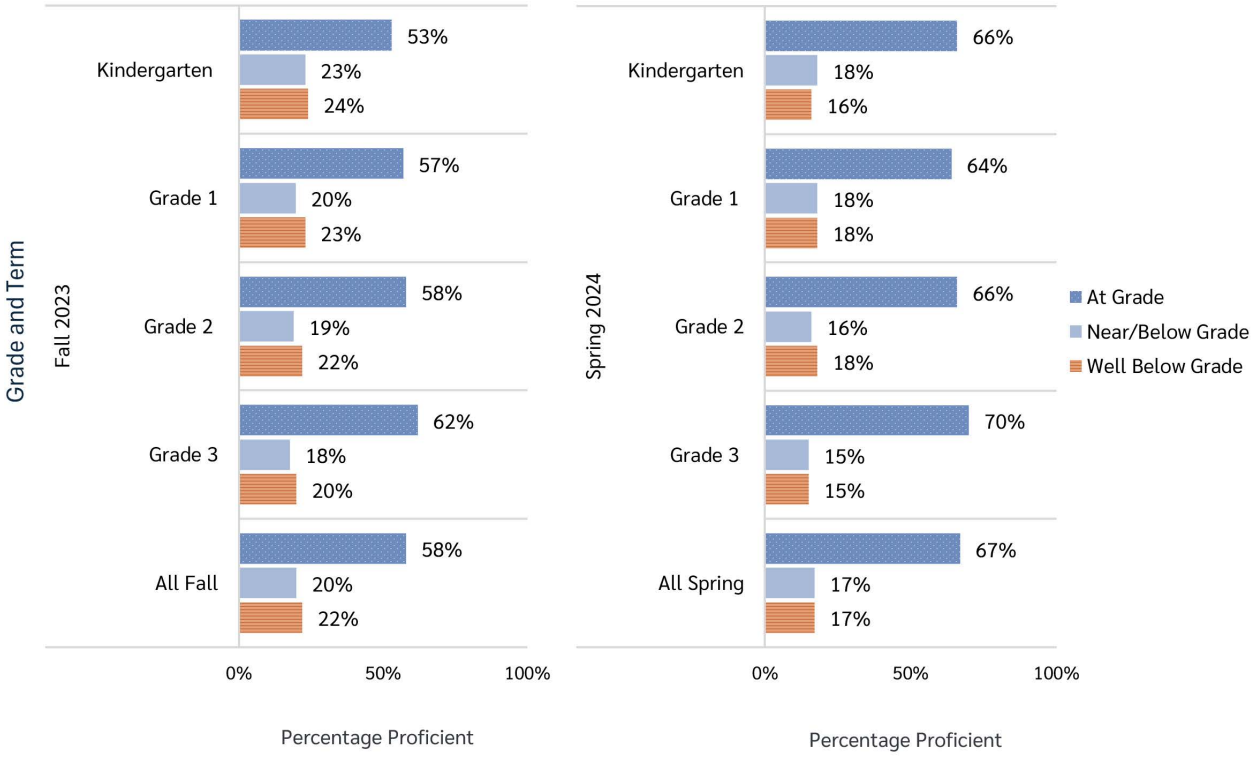
In many cases, students do not come close to meeting even vendor recommendations for time needed in these programs to achieve results, and most vendor recommendations fall below Idaho recommended 30 and 60 hours of intervention time. This makes it difficult to identify the extent to which these programs are impacting students. To the extent possible, IRI increases have been mapped to vendor recommended times as well as to the 30-hour and 60-hour marks to determine how dosage in the programs is impacting student outcomes.

In cases where this is not possible due to limited student usage time, dose-response analyses were conducted using buckets based on available vendor data.

Idaho Statewide Results

Figure 1 shows the statewide Idaho Reading Indicator (IRI) results for the 2023–24 school year, as provided in the 2023–24 Student Achievement Report, completed by the Idaho State Department of Education in cooperation with the State Board of Education's Accountability Oversight Committee (AOC) as a part of the AOC FY25 Recommendations Report.

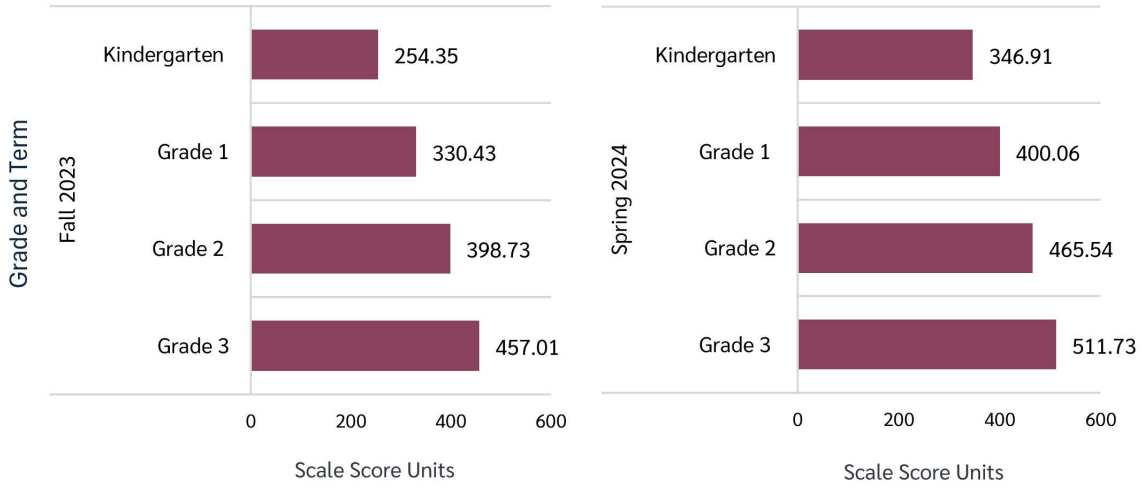
Figure 1. IRI Results for 2023–24



The data is provided for all students, indicating that 58% of students scored At Grade Level on the IRI in Fall 2023 and 67% of students scored At Grade Level in Spring 2024. While the proficiency rates and gains Fall to Spring vary by grade, on average (K–3) the state had a 9 percentage-point gain in the proficiency rate from Fall to Spring in the 2023–24 school year.

Even within a proficiency level, students make increases and show growth over the course of the year. Figure 2 details statewide scale scores by grade in Fall 2023 and Spring 2024.

Figure 2. Statewide Scale Scores, Fall 2023 and Spring 2024



On average, kindergarteners' scale scores increased 92.56 units between Fall and Spring, first graders increased 69.63 units, second graders increased 66.81 units, and third graders increased 54.72 units. We would expect these averages to be close to averages for each vendor and well within a standard deviation of students IRI gains for each vendor. Data that differs drastically from these numbers should be examined closely.

Lexia Learning (Core 5)

Recommendation

Remain on List

Summary Highlights

- High-quality data with a >99% student match rate across grades K–3
- Strong alignment with state trends in IRI proficiency growth
- Above-state-average gains in grade 3 and solid growth across all grades for students meeting usage targets
- Substantial implementation, with over half of matched students meeting Lexia's recommended usage thresholds

Vendor-Reported Outcomes

Lexia reported that 71% of K–3 students in Idaho advanced at least one grade level in Core5 during the 2023–24 school year. Among students who met Lexia's recommended usage (defined as using the program for ≥20 weeks and meeting weekly usage goals at least 50% of those weeks), 88% demonstrated grade-level advancement. Students meeting usage targets completed an average of 22 skills over the year—double the number completed by non-target users and five times that of partial-year users. These data were drawn from Lexia's internal analytics system, which tracks student progression across 21 levels covering PreK through grade 5.

Lexia's data file included 10,057 students from 47 schools in 12 Idaho districts.

Data Matching and Evaluation Scope

Lexia submitted individual usage data for 10,057 K–3 students. The overall match rate for Lexia Learning was excellent. Across all grades, the match rate of student data was greater than 99%. Of the 10,057 records submitted, 10,011 were able to be matched, and 9,975 of these matched students had at least 30 minutes of time in product.

Implementation and Usage Patterns

Among matched students with sufficient usage:

- 55.08% met Lexia's recommended usage thresholds
- 40.61% reached at least 30 hours in the program
- 4.78% reached 60 hours or more

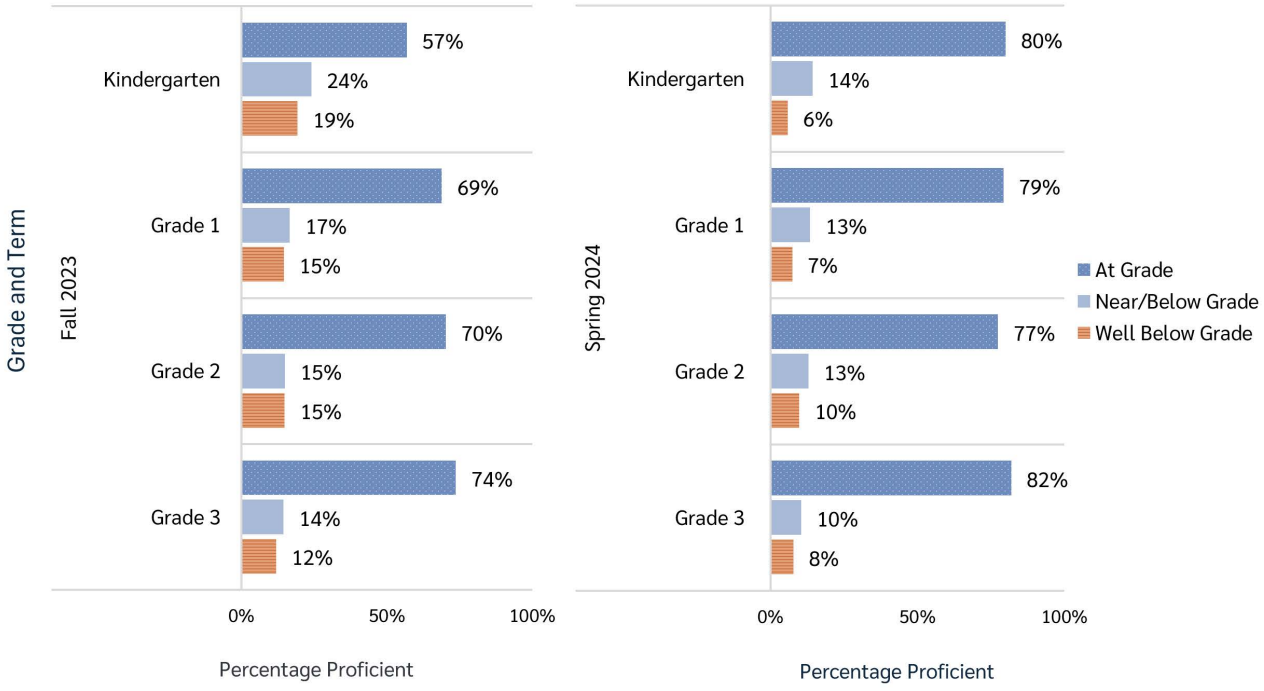
Average total time ranged from 1,441 minutes (grade 3) to 1,923 minutes (grade 1), reflecting Lexia's adaptive weekly time prescriptions based on student needs. Lexia reported that students meeting usage targets completed 9–10 units per week in their program on average.

IRI Outcomes and State Comparison

Across all grades, matched Lexia users demonstrated proficiency gains from Fall to Spring that mirrored or exceeded state averages, with notably high gains in kindergarten.

As seen in Figure 3 below, for all grades, the percentage of students performing significantly below grade level also declined at rates similar to or above state rates.

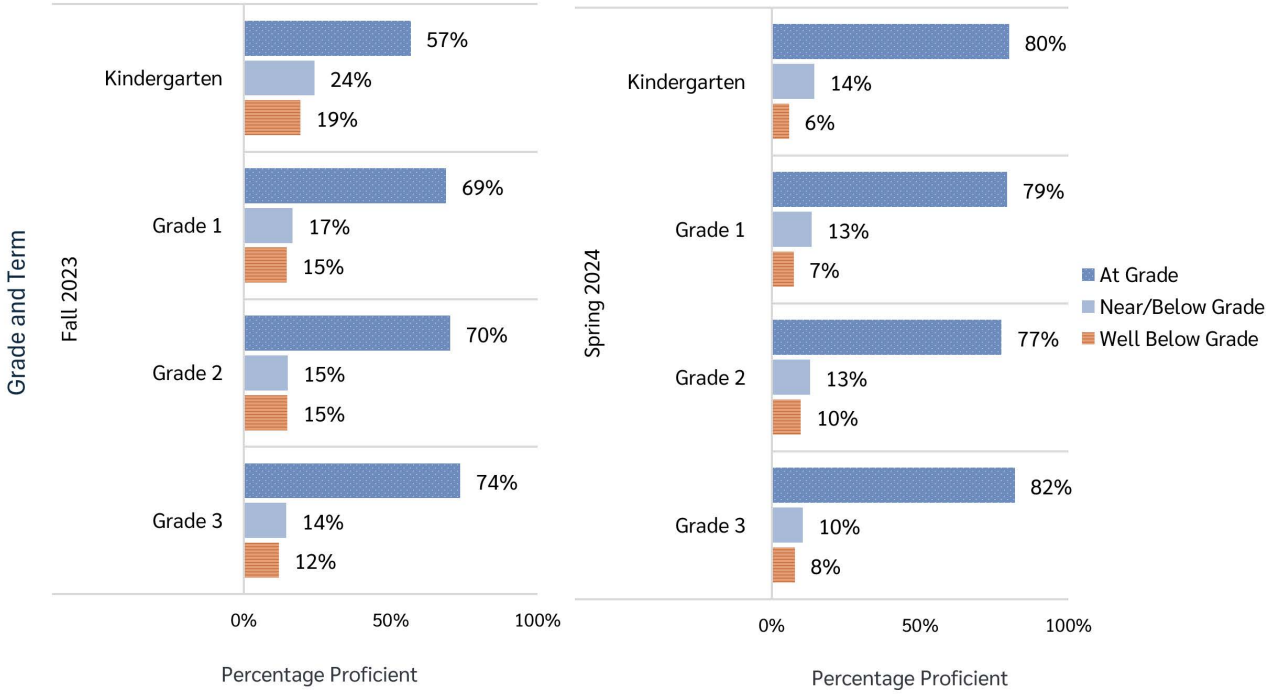
Figure 3. Lexia - IRI Proficiency Outcomes for Students With At Least 30 Minutes of Time in Product



Note. Percentage of all matched students by category.

Students who met Lexia's usage thresholds saw even stronger results: Spring proficiency exceeded 77% across all grades. Results for students meeting vendor recommended usage thresholds can be seen in Figure 4.

Figure 4. Lexia - IRI Proficiency Outcomes for Students Who Met Vendor Recommended Usage Targets



Note. Percentage of all matched students by category.

While proficiency gains are one important indicator of progress, gains in scores are also important indicators of student growth when using intervention products. There are large score bands within each proficiency level. Students, particularly those who are starting with a lower score, may need to make up more ground to be able to cross to the next proficiency level. Looking only at proficiency scores can hide the growth that students are making, and can mask lack of growth for students performing at higher levels. To help evaluate this, we examined both the proportion of students who may be losing ground over the course of the year, as well as the relationship of average student score changes between Fall and Spring to state averages. Because student scores can vary widely, we looked not only at the average student score for each grade, but also at whether the expected score range, based on standard deviation, fell within the state average (Table 5).

Table 5. Lexia - Summary Outcomes for Students with Matched Scores and at Least 30 Minutes Usage

Grade Level	Average IRI Score Increase	Increase Compared to State Average	Is Average within Expected Growth Range	Average Time in Product (mins in SY)	% Students Losing Ground (Fall to Spring)
Kindergarten	89.7	Below	Yes	1,571	N/A
Grade 1	68.7	Below	Yes	1,923	N/A
Grade 2	63.5	Below	Yes	1,803	N/A
Grade 3	57.5	Above	Yes	1,441	N/A
TOTAL	69.2	N/A	Yes	1,691	2.29%

Note. Includes percentages of students with matched scores who had at least 30 minutes in product.

Grade 3 score increases were above the state average growth for third graders, and while grades K–2 results were under state averages, they were all within the expected growth range. A total of 2.29% of students lost ground between Fall and Spring IRI assessments, which is within the expected 2%–4% range.

Dosage and Differential Impacts

At the heart of the effectiveness question is whether spending more time using an intervention helps students become more proficient in their literacy skillset. To understand this, we examined dosage and its relationship to student scores. In general, dosage recommendations mirrored Lexia's own analysis of the data. Students who were meeting the recommended product usage grew more (77.6 units on IRI) than students who had not met recommended usage requirements (58.9 points).

Lexia usage demonstrated a clear dosage effect. Students with more time in Core5 showed higher proficiency rates by Spring as seen below.

Table 6. IRI Proficiency by Time in Lexia

Student Time in Product	n	Fall Level Proficient (n)	Spring Level Proficient (n)	Fall % Proficient	Spring % Proficient	% Change
<10 hours	1,407	622	670	44.2%	47.6%	3.4%
10–19.99 hours	2,175	1,096	1,329	50.4%	61.1%	10.7%
20–29.99 hours	2,419	1,381	1,671	57.1%	69.1%	12.0%
30–39.99 hours	2,007	1,115	1,418	55.6%	70.7%	15.1%
40+ hours	2,049	1,169	1,468	57.1%	71.6%	14.6%

Note. Includes percentages of students with matched scores.

Conclusion

Lexia Learning's Core5 program continues to deliver consistent and scalable results for Idaho K–3 students, with:

- High fidelity of implementation
- Clear impact on literacy growth and proficiency
- Alignment with state trends and benchmarks
- Strong data quality supporting robust evaluation

We recommend Lexia Learning remain on the Approved Vendor List.

Imagine Learning (Language and Literacy)

Recommendation

Remain on List

Summary Highlights

- Good data quality (92.6% match rate)
- Evidence of impact where time-on-task thresholds are met
- Score gains within expected ranges, though below state averages
- Moderate usage, with only a small percentage meeting vendor time guidelines

Vendor-Reported Outcomes

Imagine Learning reported strong literacy and vocabulary gains among Idaho K–3 students during the 2023–24 school year. Of the 4,336 students who logged at least one hour in the program, 2,368 completed both beginning-of-year (BOY) and end-of-year (EOY) benchmark assessments.

- First graders showed the highest average literacy benchmark gain (+323 points), followed by kindergarteners (+258), second graders (+235), and third graders (+176)
- Internal vocabulary assessments followed a similar pattern
- The percentage of students performing at or above grade level increased from 68% in the Fall to 85% in the Spring

Regression analysis confirmed a statistically significant, positive relationship between time spent in Imagine Language and Literacy and student growth on benchmark assessments.

The submitted data file included 5,522 K–3 students across 110 schools and 41 Idaho districts.

Data Matching and Evaluation Scope

Of the 5,522 student records submitted, 5,115 were successfully matched to state data, yielding a 92.63% match rate. Among matched students, 79.24% (4,374 students) had at least 30 minutes of product usage and were included in the analysis sample.

Implementation and Usage Patterns

Among matched students with sufficient usage:

- 13.69% met Imagine Learning's recommended usage threshold (≥20 hours/year)
- 6.41% met the 30-hour state guideline
- 0.25% met or exceeded 60 hours of usage

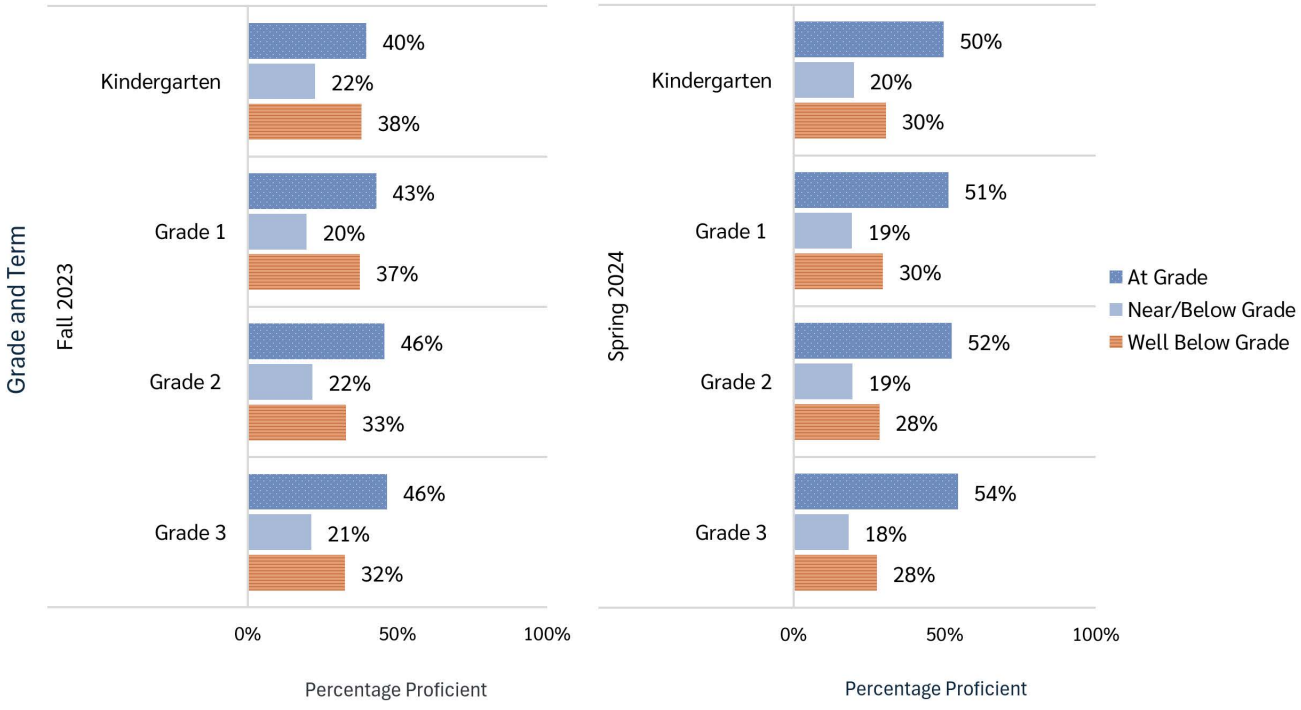
Average time in product ranged from 449 minutes (grade 3) to 751 minutes (Kindergarten), or approximately 7.5 to 12.5 hours per year. These averages were below Imagine Learning's recommended threshold of 20 hours and also lower than what the vendor-reported for the broader user base (15.8 hours average).

This discrepancy suggests possible implementation gaps among the matched student population compared to the full user set. Nonetheless, vendor-reported analyses show that students who exceeded 20 hours demonstrated significantly greater growth ($p = .001$ for literacy, $p < .001$ for vocabulary) than students with lower usage.

IRI Outcomes and State Comparison

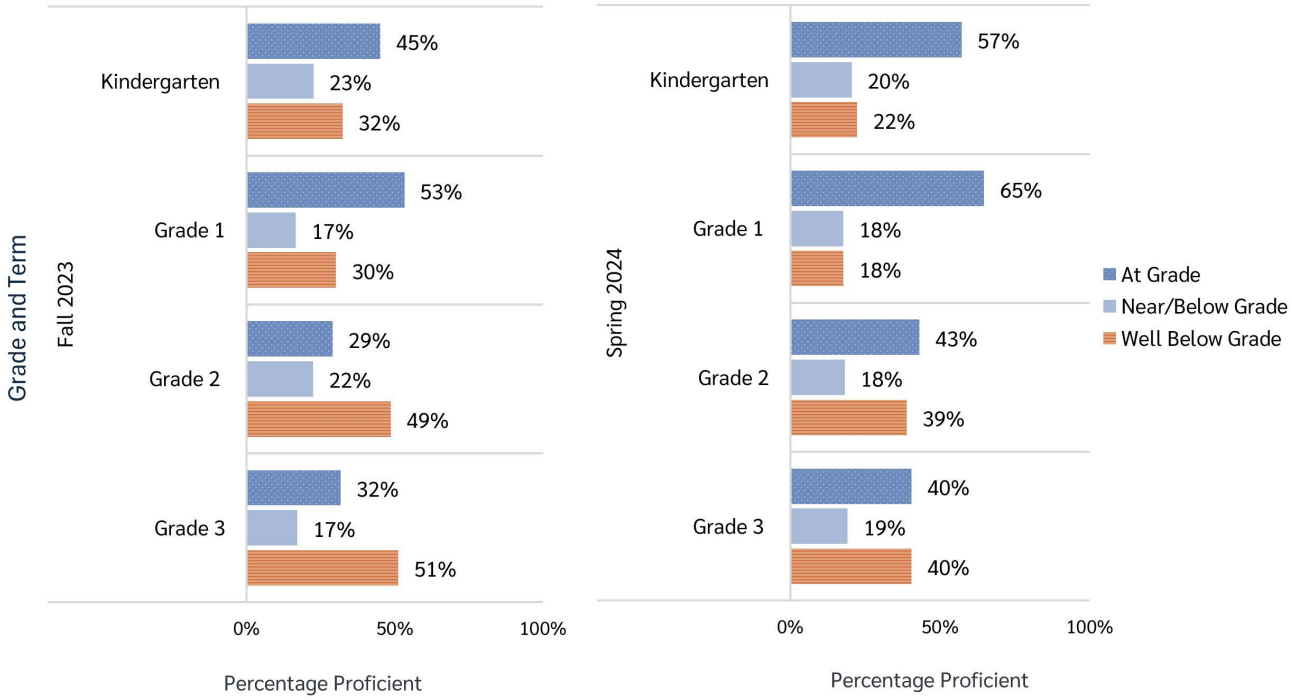
Matched students demonstrated IRI proficiency gains from Fall to Spring that largely mirrored statewide trends. Although their starting proficiency rates were lower than the state average, students showed strong reductions in the proportion performing significantly below grade level.

Figure 5. Imagine Learning - IRI Proficiency Outcomes for Students With At Least 30 Minutes of Time in Product



Note. Percentage of all matched students by category.

Figure 6. Imagine Learning - IRI Proficiency Outcomes for Students Who Met Vendor Recommended Usage Targets



Note. Percentage of all matched students by category.

Growth Beyond Proficiency: Score Analysis

While proficiency level changes are critical, average score increases provide a fuller picture of growth. Students performing below benchmark may make substantial gains without necessarily crossing into the next level. We examined both the magnitude of score increases and the percentage of students who lost ground over the year (Table 7).

Table 7. Imagine Learning - Summary Outcomes for Students with Matched Scores and at Least 30 Minutes Usage

Grade Level	Average IRI Score Increase	Increase Compared to State Average	Is Average within Expected Growth Range	Average Time in Product (mins in SY)	% Students Losing Ground (Fall to Spring)
Kindergarten	79.9	Below	Yes	751	N/A
Grade 1	65.7	Below	Yes	694	N/A
Grade 2	59.2	Below	Yes	571	N/A
Grade 3	50	Below	Yes	449	N/A
TOTAL	62.7	N/A	N/A	609	2.27%

Note. Includes percentages of students with matched scores who had at least 30 minutes in product.

While all average score increases were under state averages, they fell within the expected growth range. A total of 2.27% students lost ground between Fall and Spring IRI assessments, which is within the expected 2%–4% range.

Dosage and Differential Impacts

Program effectiveness increased with usage. Students who met the 20-hour benchmark grew an average of 73.5 IRI score units, compared to 60.7 units for those who did not.

These findings validate vendor-reported findings that highlight the added benefit of higher engagement, particularly for ELLs, who constituted a large portion of the Idaho sample. Further subgroup analysis (e.g., by ELL status or baseline IRI tier) may provide more granular insights into which populations benefit most from sustained program use.

Table 8. Imagine Learning - Average IRI Proficiency Rates by Usage Category

Usage Bucket	n	Fall Level Proficient (n)	Spring Level Proficient (n)	Fall % Proficient	Spring % Proficient	% Change
<5 hours	3,020	1,156	1,423	38.3%	47.1%	8.8%
10–19.99 hours	915	316	413	34.5%	45.1%	10.6%
5–9.99 hours	831	282	368	33.9%	44.3%	10.3%
20+ hours	756	290	385	38.4%	50.9%	12.6%

Note. Includes percentages of students with matched scores.

This dosage pattern aligns with Imagine Learning's internal analysis and demonstrates that even incremental increases in usage yield stronger outcomes.

Conclusion

Imagine Learning (Language and Literacy) delivers consistent, positive outcomes for Idaho students, particularly when implemented with fidelity. While average growth fell below state benchmarks, all results remained within expected growth ranges, and students who met usage recommendations showed substantially stronger gains.

We recommend Imagine Learning remain on the Approved Vendor List.

Renaissance (Freckle ELA)

Recommendation

Reinstate to List - Probationary — Rationale Present, But Data Limitations Require Monitoring

Summary Highlights

- Renaissance was previously on the list but removed after the FY 24 review for having two years of inadequate data; the data submitted this year was provided with a request for consideration for reinstatement
- Low data quality: 52% match rate is well below acceptable thresholds; improvement to ≥85% is strongly recommended for next year
- Weak implementation fidelity: Very few students met vendor usage guidelines; none met state benchmarks for time in product
- Promising results in grades 2 and 3, particularly for students who met vendor targets
- Students who met usage thresholds significantly outperformed peers, but sample size is too limited to confirm effectiveness

Vendor-Reported Outcomes

Renaissance reported that Idaho K–3 students using Freckle ELA at recommended levels demonstrated stronger literacy growth than non-users. These findings were based on internal Star Reading and Star Early Literacy benchmark assessments, which are adaptive tools measuring growth from Fall to Spring. Key findings include:

- Students meeting Renaissance's recommended usage levels were significantly more likely to achieve "above typical" growth (SGP > 50)
- 89% of students who met the usage threshold reached the end-of-year benchmark (40th percentile or higher), compared to 70% of students with lower usage and 68% of matched non-users
- Students meeting benchmarks also had higher engagement: more sessions, more time per session, and higher in-program accuracy

Renaissance used data on 361 matched users and constructed comparison groups using rigorous propensity score matching techniques. Matching criteria included district, grade, ethnicity, test type, and baseline performance. Effect size checks confirmed that matched groups were statistically equivalent at baseline.

Data Matching and Evaluation Scope

Renaissance was previously on the list but removed after the FY 24 review for having two years of inadequate data. The data submitted this year was provided with a request for consideration for reinstatement to the list.

In this year's request, a total of 3,036 student records from 46 schools across 26 districts were submitted. Only 1,616 of these could be matched to state rosters—an overall match rate of just 52.23%. Of those, only 848 students had at least 30 minutes of product use and were included in the analysis sample. This is both a low match rate and a low fidelity rate. As a result, insights should be viewed with caution. Our recommendation is to reinstate Renaissance on a probationary basis. If match rates and usage fidelity do not improve substantially next year, the evidence base will remain insufficient to justify keeping Renaissance on the list.

Implementation and Usage Patterns

Among matched students with ≥30 minutes of usage:

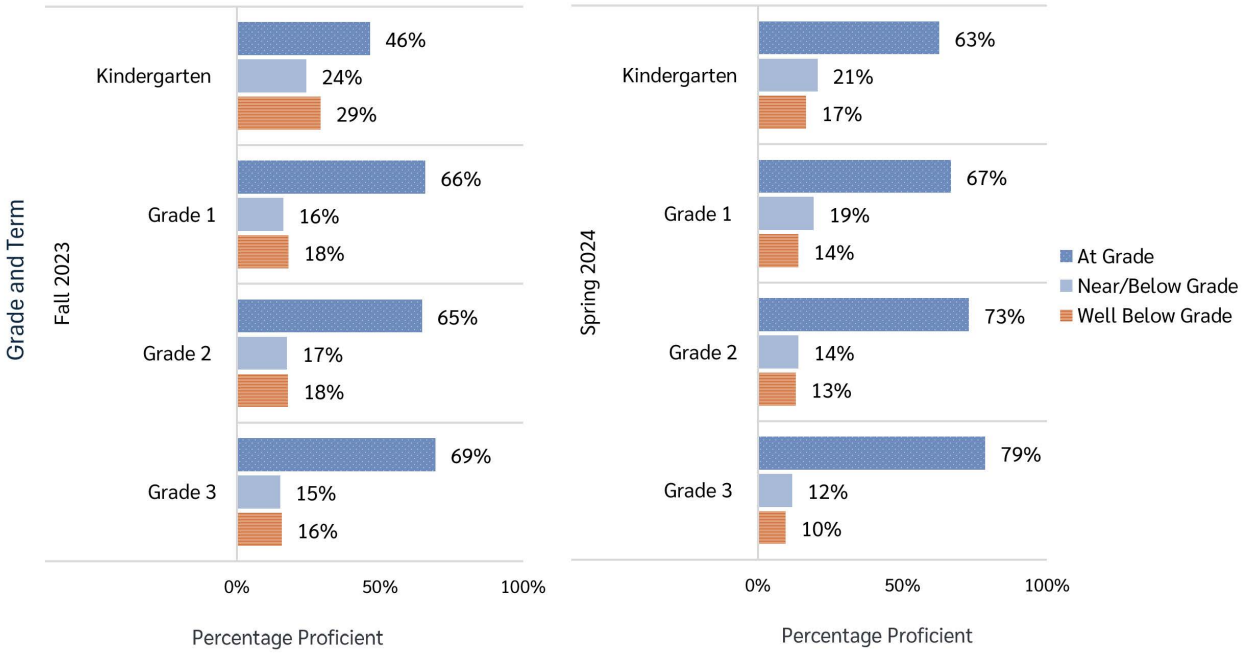
- Only 6.72% met Renaissance's recommended usage thresholds (20 minutes/week for grades K–2, 40 minutes/week for grade 3)
- No students met the 30- or 60-hour state time in product guidelines

Average usage ranged from 113 minutes (grade 2) to 203 minutes (Kindergarten)—approximately 2–3.5 hours per year, far below recommended levels. This is not only lower than Renaissance's own thresholds but also lower than the usage levels reported in their national efficacy study. These discrepancies likely indicate major implementation gaps.

IRI Outcomes and State Comparison

IRI proficiency gains as seen below for matched Renaissance users tracked closely with state averages. Importantly, the percentage of students performing significantly below grade level decreased across all grades.

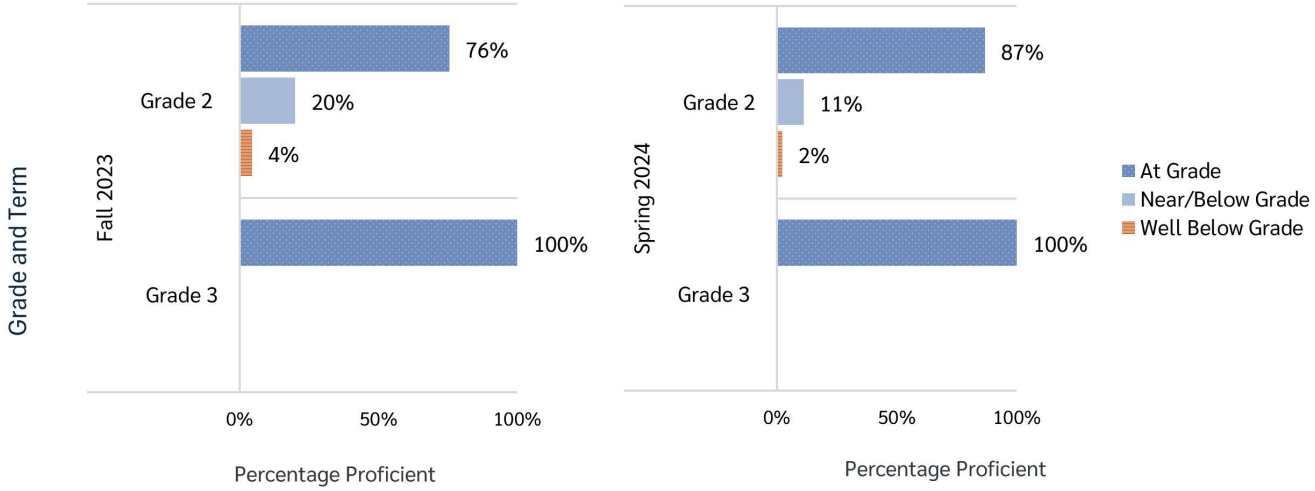
Figure 7. Renaissance - IRI Proficiency Outcomes for Students With at Least 30 Minutes of Time in Product



Note. Percentage of all matched students by category.

Only grades 2 and 3 had enough students meeting usage thresholds to analyze gains by grade. In both, Spring proficiency rates improved significantly, with grade 2 showing a >10 percentage point increase.

Figure 8. Renaissance - IRI Proficiency Outcomes for Students Who Met Vendor Recommended Usage Targets



Note. Percentage of all matched students by category.

Beyond proficiency bands, we examined average IRI score gains and the proportion of students losing ground. Average gains for grades 2 and 3 were above the state average, though the limited sample and poor match rate temper interpretation.

Table 10. Renaissance - Summary Outcomes for Students with Matched Scores and at Least 30 Minutes Usage

Grade Level	Average IRI Score Increase	Increase Compared to State Average	Is Average within Expected Growth Range	Average Time in Product (mins in SY)	% Students Losing Ground (Fall to Spring)
Kindergarten	89	Below	Yes	203	N/A
Grade 1	63.7	Below	Yes	165	N/A
Grade 2	73.9	Above	Yes	113	N/A
Grade 3	67.83	Above	Yes	117	N/A
TOTAL	71.4	N/A	N/A	170	2.00%

Note. Includes percentages of students with matched scores who had at least 30 minutes in product.

All grade-level averages were within expected growth ranges, and only 2% of students lost ground—a rate consistent with expectations.

Dosage and Differential Impacts

Students who met usage guidelines showed greater growth (85.4 points) than peers with lower usage (70.6 points). However, due to extremely low numbers of students in higher usage brackets, dosage effects could not be reliably confirmed.

Table 11. Renaissance - Average IRI Proficiency Rates by Usage Category

Usage Bucket	n	Fall Level Proficient (n)	Spring Level Proficient (n)	Fall % Proficient	Spring % Proficient	% Change
<5 hours	2,749	918	1,048	33.4%	38.1%	4.7%
10–14.99 hours	44	17	19	38.6%	43.2%	4.5%
5–9.99 hours	230	91	99	39.6%	43.0%	3.5%
15+ hours	Insufficient Size	N/A	N/A	N/A	N/A	N/A

Note. Includes percentages of students with matched scores.

Limited sample sizes preclude firm conclusions, but early trends mirror those observed in Renaissance internal benchmark analysis.

Conclusion

Despite substantial data limitations, Renaissance (Freckle ELA) demonstrated promising outcomes in grades 2 and 3 and strong results for the small subset of students meeting usage recommendations. These findings are not yet generalizable but are sufficient to justify continued inclusion on a provisional basis.

We recommend Renaissance be reinstated to the Approved Vendor List for one additional year on a probationary basis. Including Renaissance beyond 2024–25 will be contingent on the vendor and districts addressing:

- Data quality issues – match rate should be ≥85%
- Implementation fidelity – A significantly higher percentage of students must meet usage thresholds

Failure to demonstrate improved usage and match fidelity should result in removal from the list in the next review cycle.

Curriculum Associates (i-Ready)

Recommendation

Remain on List

Summary Highlights

- Excellent data quality with a 99.9% match rate
- Strong implementation fidelity, with over 42% of users meeting vendor-recommended usage—among the highest rates across vendors
- Growth outcomes within expected state ranges at all grade levels, though slightly below average
- Clear association between increased usage and improved outcomes
- Usage trends show higher implementation fidelity in upper grades, where support is often needed

Vendor-Reported Outcomes

Curriculum Associates reported that Idaho students using i-Ready Personalized Instruction demonstrated strong academic growth and performance during the 2023–24 school year. Students in grades K–3, on average, exceeded 100% of their Typical Growth targets, with many also achieving Stretch Growth, an ambitious benchmark aimed at closing achievement gaps and advancing proficiency. Key highlights from vendor-submitted data:

- Idaho students outperformed both pre-pandemic (2018–19) and 2023–24 national samples across reading domains, including phonological awareness, phonics, vocabulary, and comprehension
- Among students who met 30 minutes/week usage with a 70%+ pass rate, growth exceeded expectations, reaching 153% of Typical Growth and surpassing Stretch Growth targets
- Usage data and in-platform performance allowed for targeted teacher intervention and instructional planning
- Growth was evident across all student performance levels, with below-grade-level students making notable gains in foundational skills, and on/above-grade-level students progressing in higher-order domains

Data Matching and Evaluation Scope

Curriculum Associates submitted 13,452 student records for evaluation, with an exceptionally high match rate of 99.9% (13,442 students were successfully matched to state data). Of those, 13,234 students (98.4%) had at least 30 minutes of time in the i-Ready product and were included in the analysis sample. The data represented 84 schools across 22 Idaho districts, providing a robust and representative dataset for assessing implementation and student outcomes.

Implementation and Usage Patterns

Among matched students with sufficient usage:

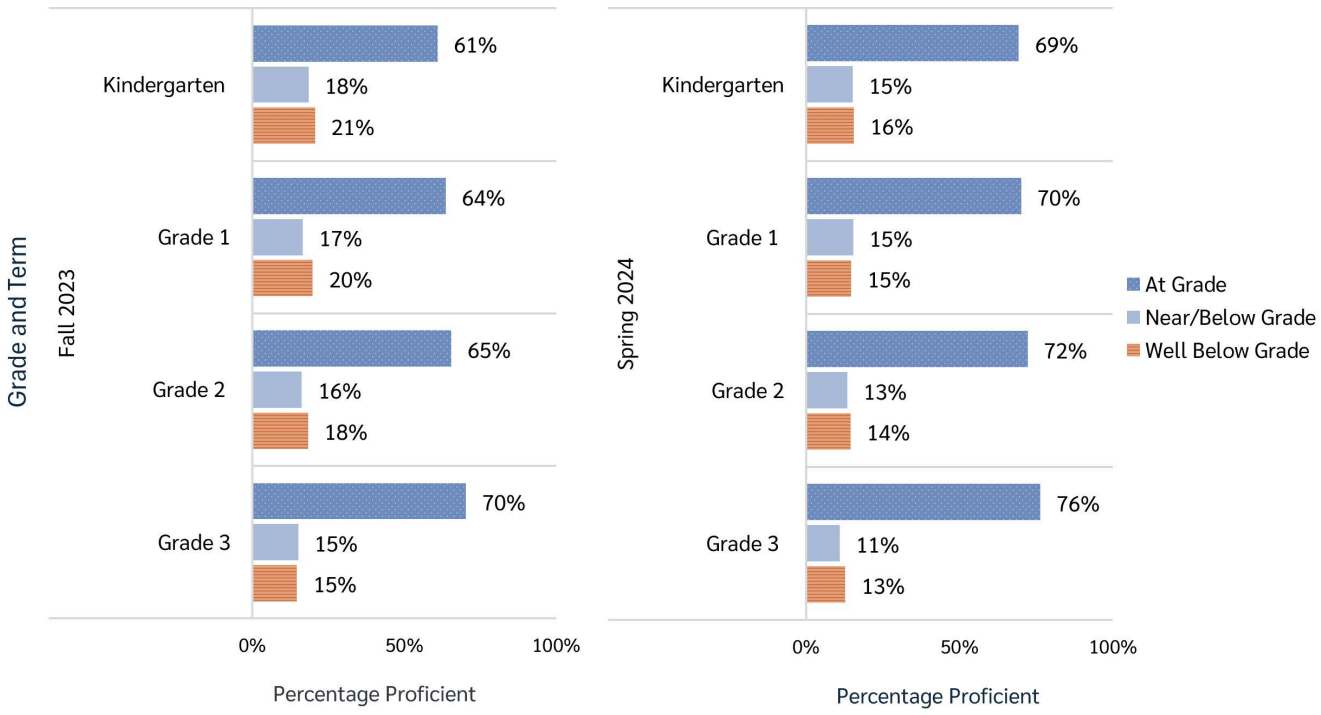
- 42.38% met vendor-recommended usage thresholds
- 6.91% met the 30-hour state guideline
- 0.16% met or exceeded 60 hours of usage

Average time in product ranged from 845 minutes (Kindergarten) to 1,029 minutes (grade 3). i-Ready demonstrated a unique trend among vendors on the list: greater implementation fidelity at higher grade levels, where Tier 2 and Tier 3 interventions are often concentrated.

IRI Outcomes and State Comparison

Proficiency gains among i-Ready users closely mirrored statewide trends. While average gains were slightly below those of the state overall, they remained within the expected growth range for all grades. In general, i-Ready students started with slightly higher Fall proficiency rates, potentially limiting the amount of improvement possible by Spring.

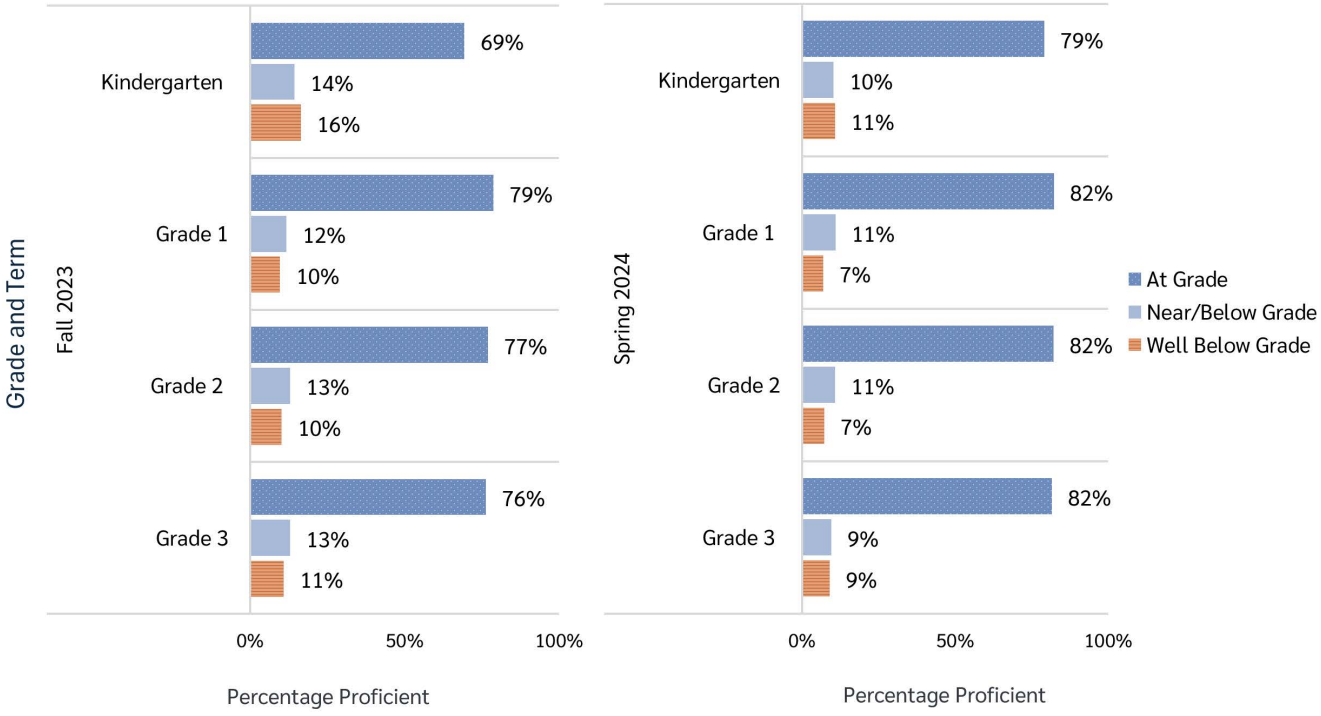
Figure 9. Curriculum Associates - IRI Proficiency Outcomes



Note. Percentage of all matched students by category.

Students who met vendor usage recommendations significantly outperformed the state average, consistent with vendor-submitted findings.

Figure 10. Curriculum Associates - IRI Proficiency Outcomes for Students Who Met Vendor Proficiency Targets



Note. Percentage of all matched students by category.

Growth Beyond Proficiency: Score Analysis

While no grade-level cohort exceeded the statewide average for IRI score gains, all remained within the expected range, and students demonstrated steady improvement across grade bands.

Table 12. Curriculum Associates - Summary Outcomes for Students with Matched Scores and at Least 30 Minutes Usage

Grade Level	Average IRI Score Increase	Increase Compared to State Average	Is Average within Expected Growth Range	Average Time in Product (mins in SY)	% Students Losing Ground (Fall to Spring)
Kindergarten	85.4	Below	Yes	845	N/A
Grade 1	68.9	Below	Yes	907	N/A
Grade 2	65.6	Below	Yes	982	N/A
Grade 3	52.2	Below	Yes	1029	N/A
TOTAL	67	N/A	N/A	946	2.27%

Note. Includes percentages of students with matched scores who had at least 30 minutes in product.

Only 2.27% of students lost ground between Fall and Spring, well within the 2%–4% expected range.

Dosage and Differential Impacts

Students who met vendor usage guidelines demonstrated slightly greater IRI score gains (67.4 points) than those who did not (66.8 points), indicating a modest dosage effect. While the difference is small, it is consistent with vendor findings that students adhering to usage targets are more likely to exceed growth expectations. Overall dosage impact was inconclusive across the larger sample, and it might be beneficial to do additional analysis to understand which participating schools are using both core and supplemental curriculum to determine if there is a fatigue effect in intervention time that could be contributing to the inconclusive results.

Table 13. Curriculum Associates - Average IRI Proficiency Rates by Usage Category

Usage Bucket	n	Fall Level Proficient (n)	Spring Level Proficient (n)	Fall % Proficient	Spring % Proficient	% Change
<10 hours	4,170	1,785	2,264	42.8%	54.3%	11.5%
10–19.99 hours	5,333	3,473	4,032	65.1%	75.6%	10.5%
20–29.99 hours	3,032	2,299	2,534	75.8%	83.6%	7.8%
30–39.99 hours	660	540	576	81.8%	87.3%	5.5%
40+ hours	257	203	220	79.0%	85.6%	6.6%

Note. Includes percentages of students with matched scores.

Conclusion

Curriculum Associates' i-Ready Personalized Instruction is widely used across Idaho with high data quality and evidence of impact. While score gains were modestly below state averages, they remained within expected ranges, and implementation fidelity was higher than most other vendors. While gains slightly tapered at the highest usage levels, the consistent upward trend in performance with time-in-product supports the effectiveness of sustained usage.

We recommend Curriculum Associates remain on the Approved Vendor List. Continued monitoring is warranted, particularly of subgroup outcomes and fidelity at early grades. Overall, i-Ready remains a reliable and well-implemented tool for personalized literacy instruction and growth measurement.

Savvas (SuccessMaker)

Recommendation

Remain on List

Summary Highlights

- High-fidelity implementation, with over 87% of students meeting vendor usage targets, the highest across all vendors
- High-quality data from one pilot district, with 100% match and fidelity rates
- Strong evidence of effectiveness in grade 3, including gains exceeding expected state growth averages
- Clear dosage effect, with greater reading growth observed at higher usage levels

Vendor-Reported Outcomes

Savvas Learning Company reported that students using SuccessMaker at Riverview Elementary in Shelley Joint School District 60 demonstrated clear, usage-aligned gains in reading performance during the 2023–24 school year. Among third graders, those who used SuccessMaker for more than 30 hours over the year gained an average of 1.4 grade levels, compared to just 0.5 grade levels among students with under 10 hours of usage. These gains followed a clear dose-response pattern, with each increasing time tier associated with stronger achievement outcomes. Fourth graders demonstrated nearly identical trends, reinforcing the scalability and consistency of the product's impact.

SuccessMaker delivers adaptive, individualized reading instruction, placing students at their skill level and advancing them through scaffolded, standards-aligned pathways. For Grade 3:

- Students with <10 hours gained 0.49 grade levels
- 10–19.9 hours: 0.58 grade levels
- 20–24.9 hours: 0.88 grade levels
- 25–29.9 hours: 1.11 grade levels
- 30+ hours: 1.40 grade levels

These outcomes, submitted by Savvas, also showed increases in average skill mastery, from 59% in the lowest usage group to 76% among those with the highest engagement. The design of SuccessMaker—which incorporates strategic struggle, immediate feedback, and cumulative review—likely supports this sustained progress. Notably, students who began the year with below-grade-level placements also achieved substantial gains, indicating the tool's suitability for remediation and intervention contexts.

Data Matching and Evaluation Scope

Savvas submitted data for 188 third graders in their pilot implementation. All records were matched to state IRI scores, and all students had at least 30 minutes of usage, yielding 100% match and fidelity rates. This high-quality dataset, though limited to one district, provides strong, reliable evidence of impact.

Implementation and Usage Patterns

Among students using SuccessMaker:

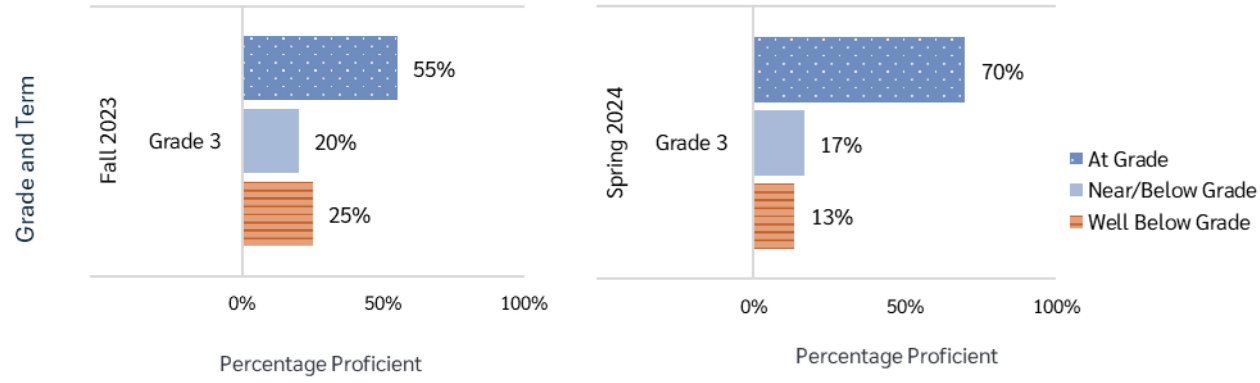
- 87.23% met vendor-recommended usage thresholds
- 9.57% met the 30-hour state guideline
- No students reached the 60-hour mark
- Average time in product: 1,324 minutes (~22 hours)

This level of usage fidelity is unmatched among vendors and may offer a model for future implementations.

IRI Outcomes and State Comparison

Third graders using SuccessMaker demonstrated proficiency gains from Fall to Spring that mirrored state averages, while the percentage of students performing well below grade level declined at a rate greater than the state average.

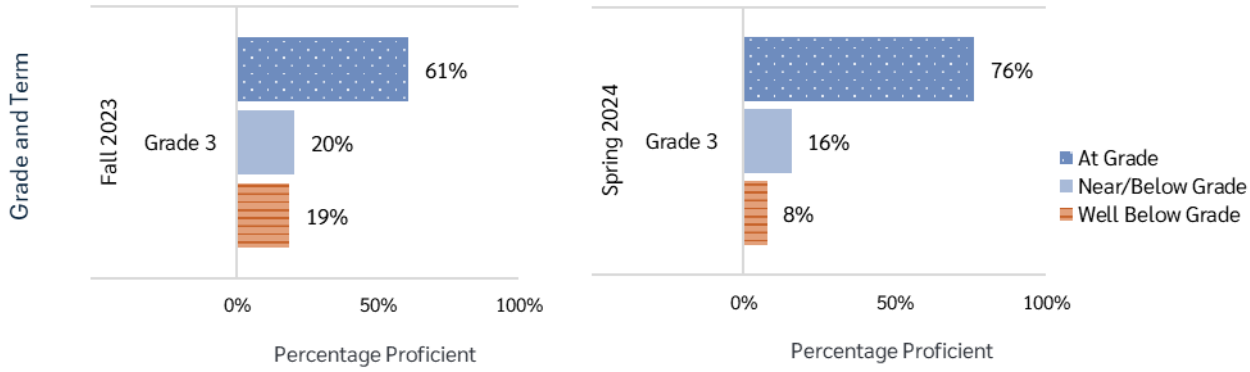
Figure 11. Savvas - IRI Proficiency Outcomes



Note. Percentage of all matched students by category.

Among students meeting vendor usage benchmarks, proficiency gains were even more pronounced, with a 15% increase in students scoring at or above grade level—nearly double the average statewide increase.

Figure 12. Savvas - IRI Proficiency Outcomes for Students Who Met Vendor Proficiency Targets



Note. Percentage of all matched students by category.

Growth Beyond Proficiency: Score Analysis

The average IRI scale score increase for third graders using SuccessMaker was 63.3 points, which exceeded the state average and fell within the expected growth range.

Table 14. Savvas - Summary Outcomes for Students with Matched Scores and at Least 30 Minutes Usage

Grade Level	Average IRI Score Increase	Increase Compared to State Average	Is Average within Expected Growth Range	Average Time in Product (mins in SY)	% Students Losing Ground (Fall to Spring)
Grade 3	63.3	Above	Yes	1,324	3.19%

Note. Includes percentages of students with matched scores who had at least 30 minutes in product.

These results suggest that even within a small sample, SuccessMaker students achieved gains above average expectations, with loss rates consistent with state benchmarks.

Dosage and Differential Impacts

Students who met the vendor-recommended usage threshold outperformed their peers, with an average IRI score gain of 64.8 points compared to 45.7 points for those who did not. This reflects a meaningful dosage effect, even within the relatively small sample.

Table 15. Savvas - Average IRI Proficiency Rates by Usage Category

Usage Bucket	n	Fall Level Proficient (n)	Spring Level Proficient (n)	Fall % Proficient	Spring % Proficient	% Change
<12 hours	25	2	8	8.0%	32.0%	24.0%
12–19.99 hours	37	13	20	35.1%	54.1%	18.9%
20–29.99 hours	108	71	86	65.7%	79.6%	13.9%
30+ hours	18	13	16	72.2%	88.9%	16.7%

Note. Includes percentages of students with matched scores.

This tiered performance reinforces the relationship between consistent use and reliable academic benefit, as documented in the Savvas vendor report.

Conclusion

While limited to a single district pilot, the Savvas implementation of SuccessMaker in Shelley SD 60 offers high-quality, high-fidelity evidence of effectiveness. Strong usage patterns, above-average gains, and clear alignment with the vendor's internal findings support keeping SuccessMaker on the list.

We recommend Savvas remain on the Approved Vendor List. Future expansions should focus on broader implementation to confirm scalability and effectiveness across a variety of Idaho contexts.

MobyMax

Recommendation

Reinstate to List - Probationary — Improvement in Implementation and Data Quality Needed

Summary Highlights

- MobyMax was previously on the list but removed after the FY 24 review for having two years of inadequate data; the data submitted this year was provided with a request for consideration for reinstatement
- Very low data match rate (51.5%) and inconsistent file formatting significantly limited evaluable data
- No students met vendor or state usage benchmarks, making it difficult to assess true program impact
- Promising gains observed in grades 2 and 3 among students who used the program for at least 5 hours
- Clear evidence of dosage effect in small high-usage samples supports potential with better implementation

Vendor-Reported Outcomes

MobyMax reported that Idaho students using its core adaptive ELA modules—including Language, Foundational Reading, Reading Skills: Informational, and Reading Skills: Literature—demonstrated measurable reading gains when time in product exceeded five hours. For instance, students who used the Foundational Reading module for 5–9.9 hours gained an average of 1.89 grade levels, compared to 0.18 grade levels for those using the product for less than 5 hours. However, average growth across all users, many of whom spent only a few minutes in product, ranged from 0.08 to 0.19 grade levels, underscoring the importance of consistent and sustained use.

MobyMax recommends approximately 20 hours of annual use (30 minutes per week) to support one full grade level of reading growth. In Idaho, very few students approached this threshold. For example:

- In the Language module, 483 students used the program for <5 hours (average growth: 0.11), while 20 students with 5–9.9 hours of use achieved 0.71 grade levels of growth
- In Reading Skills: Informational, 386 students used the module for <5 hours (average growth: 0.05), while 8 students who exceeded 5 hours gained up to 2.5 grade levels

Despite the limited high-fidelity user base, these findings suggest that MobyMax can yield significant outcomes when used consistently. However, its potential was not fully realized in the 2023–24 implementation due to low engagement across most Idaho schools.

Data Matching and Evaluation Scope

MobyMax was previously on the list but removed after the FY 24 review for having two years of inadequate data. The data submitted this year was provided with a request for consideration for reinstatement.

MobyMax submitted usage data for 1,495 students, but only 770 students (51.51%) could be matched to IRI data. Of those, 422 students had at least 30 minutes of usage, resulting in only 28.23% of submitted records being eligible for analysis. Data integration was further complicated by module-specific submissions requiring extensive manual reconciliation. Inconsistencies in student IDs and formatting may have led to slight underreporting of time in product, despite best-effort matching procedures. The dataset spanned 25 schools across 17 districts.

Implementation and Usage Patterns

No matched students met either the vendor-recommended (20 hours) or state-recommended (30 or 60 hours) usage thresholds. Only grades 2 and 3 had enough students to support meaningful analysis. Average usage was low:

- Grade 2: 115 minutes
- Grade 3: 124 minutes
- Overall: 128 minutes (approx. 2 hours/year)

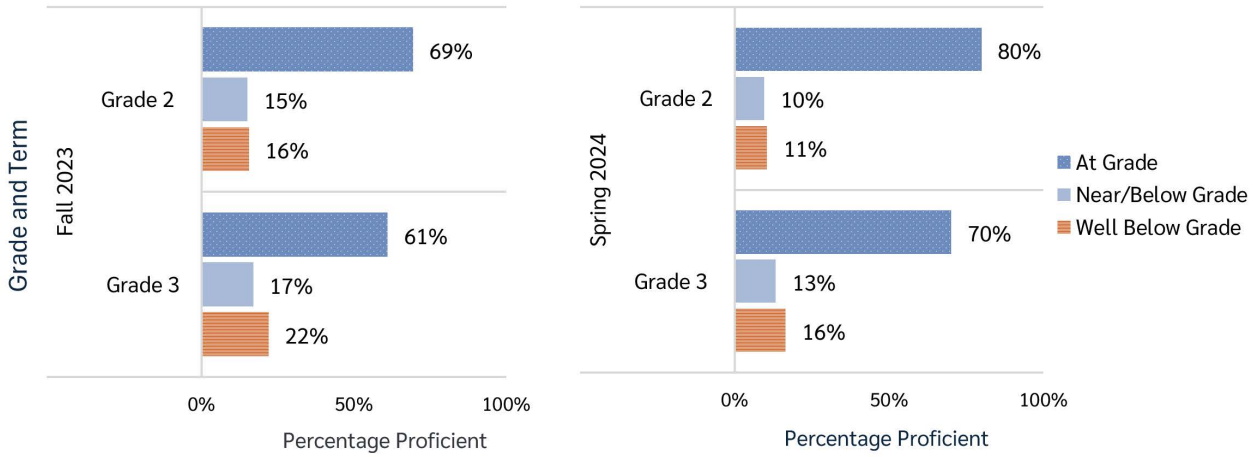
These time-on-task levels are far below recommended thresholds, limiting the ability to observe dosage impacts at scale.

IRI Outcomes and State Comparison

While limited in scope, IRI data for grades 2 and 3 showed promising trends:

- Increased proficiency rates from Fall to Spring
- Greater reductions in students scoring well below grade level than observed statewide

Figure 13. MobyMax - IRI Proficiency Outcomes



Note. Percentage of all matched students by category.

Growth Beyond Proficiency: Score Analysis

Despite low usage, students in grades 2 and 3 outperformed state average growth and remained within the expected growth range.

Table 16. MobyMax - Summary Outcomes for Students with Matched Scores and at Least 30 Minutes Usage

Grade Level	Average IRI Score Increase	Increase Compared to State Average	Is Average within Expected Growth Range	Average Time in Product (mins in SY)	% Students Losing Ground (Fall to Spring)
Kindergarten	Too Few to Report	Too Few to Report	Too Few to Report	Too Few to Report	N/A
Grade 1	Too Few to Report	Too Few to Report	Too Few to Report	Too Few to Report	N/A
Grade 2	71	Above	Yes	115	N/A
Grade 3	62.7	Above	Yes	124	N/A
TOTAL	66.8	N/A	N/A	128	2.84%

Note. Includes percentages of students with matched scores who had at least 30 minutes in product.

Only 2.84% of students lost ground between the Fall and Spring, which is within the expected 2%–4% range.

Dosage and Differential Impacts

Due to the small number of students in higher usage bands, definitive conclusions about dosage effects are limited. However, small samples suggest that increased time in product was associated with higher proficiency rates.

Table 17. MobyMax - Average IRI Proficiency Rates by Usage Category

Usage Bucket	n	Fall Level Proficient (n)	Spring Level Proficient (n)	Fall % Proficient	Spring % Proficient	% Change
<30 min	341	204	239	59.8%	70.1%	10.3%
31 min–3 hours	309	197	228	63.8%	73.8%	10.0%
3–5 hours	78	55	63	70.5%	80.8%	10.3%
>5 hours	21	15	17	71.4%	81.0%	9.5%

Note. Includes percentages of students with matched scores.

While the gains are consistent across usage bands, the small sample size in higher tiers makes it difficult to attribute effects definitively to dosage.

Conclusion

MobyMax demonstrated early promise in grades 2 and 3, where students using the product for 5+ hours made meaningful gains. However, extremely low match rates, poor implementation fidelity, and minimal time in product among most users limit the strength of conclusions.

We recommend MobyMax be reinstated to the Approved Vendor List. To justify continued list placement beyond 2024–25, significant improvements in data quality and implementation will be required—particularly a match rate of at least 85% and greater alignment to usage benchmarks. MobyMax's flexible, low-cost structure may position it well for targeted remediation, but broader fidelity is essential for evaluating system-wide impact.

Reading Horizons (Discovery)

Recommendation

Remove from List - Insufficient Data

Vendor-Reported Outcomes

Reading Horizons reported minimal usage of its Discovery platform across Idaho K–3 schools during the 2023–24 school year. Of the 1,686 student accounts registered statewide, 93% had no recorded usage, and only 125 students logged any time in the program. Critically, no students reached the recommended 40-hour usage threshold, which the vendor identifies as the minimum required for instructional impact based on internal research. Among users, the average time spent was just 55 minutes, with a maximum of 8.5 hours recorded.

The Discovery program was used in only two schools statewide:

- Albion Elementary: 58 general education students had accounts but recorded no usage; 2 special education students logged brief sessions
- Declo Elementary: Of 240 registered students, 123 logged usage; most activity came from first graders and third graders

Even among these users, average usage remained far below the program's impact threshold:

- Grade 1: 71 minutes
- Grade 2: 36 minutes
- Grade 3: 45 minutes
- Kindergarten: 0 usage

These patterns suggest that the program was not meaningfully implemented in 2023–24 and that reported outcomes cannot be used to assess effectiveness.

Data Matching and Evaluation Scope

Reading Horizons submitted 125 student records, of which 123 were matched to state data, yielding a 98.4% match rate. However, only 94 students logged at least 30 minutes of use, resulting in an evaluable sample of 75.2%. All data came from a single district, further limiting generalizability.

Implementation and Usage Patterns

Among matched students with sufficient usage data, no students met the vendor's recommended 40-hour threshold, nor the state's 30-hour or 60-hour benchmarks. The average time in product for the matched sample was 105 minutes (approximately 1.75 hours). Only grades 1 and 3 had enough students to warrant grade-level summary:

- Grade 1: 144 minutes average usage
- Grade 3: 143 minutes average usage

These levels fall significantly short of instructional thresholds and do not support impact analysis.

IRI Outcomes and State Comparison

Given the extremely limited and inconsistent usage of Reading Horizons (Discovery), meaningful comparisons to statewide IRI growth are not feasible. Matched IRI outcomes are unlikely to show any statistically or practically significant differences attributable to program exposure in 2023–24.

Dosage and Differential Impacts

Because no students reached the recommended dosage, and only one student exceeded even 8 hours of use, it is not possible to assess any dosage-related effects. The vendor has indicated that usage levels have already increased in the 2024–25 school year, which may provide a foundation for future evaluation.

Conclusion

Although Reading Horizons (Discovery) may offer a structured and research-based literacy framework, the 2023–24 implementation in Idaho was too limited to support any conclusions about its effectiveness.

Due to insufficient data, we recommend Reading Horizons be removed from the list.

Amira (Previously Istation)

Recommendation

Remain on List — Monitor Participation Rates and Dosage Implementation in 2024–25

Summary Highlights

- Widespread implementation across 129 districts and 409 schools
- Excellent match rate (99.2%), but less than half of enrolled students used the program
- No students reported as meeting vendor usage thresholds, though 7.9% met the state's 30-hour benchmark and 2.6% exceeded 60 hours
- IRI proficiency rates improved across all grades, closely mirroring state trends
- All average IRI growth scores fell within expected ranges, though slightly below state averages in most grades
- Strong internal results from vendor's Amira ISIP benchmark, but data need to be validated with consistent IRI-aligned usage and reporting

Vendor-Reported Outcomes

Following the merger of Amira Learning and Istation, the newly integrated Amira platform reported substantial literacy gains among Idaho students using Amira Instruct during the 2023–24 school year. According to the vendor, students using Amira for the recommended 20–40 minutes per week consistently outperformed peers on internal ISIP benchmark assessments, with higher score and percentile gains across grades K–3.

Highlights from the vendor study included:

- 99% of students showed positive ISIP score growth
- 84% showed percentile gains, with second graders benefiting most from exceeding the recommended time
- Students meeting or exceeding usage targets in Tiers 1 and 2 gained over 24 percentile points, reinforcing Amira's effectiveness for intervention
- The optimal time window for most grades was 20–40 minutes/week; for grade 2, performance improved further beyond 40 minutes/week

These results support Amira Instruct as a targeted tool for both Tier 2 and Tier 3 intervention and general classroom use, though they rely on the vendor's internal benchmarks. IRI-aligned usage and outcomes will be critical for validating these findings in Idaho's accountability context.

Data Matching and Evaluation Scope

Amira submitted 88,358 student records, of which 87,638 (99.2%) were matched to IRI data. However, only 43,506 students (49.2%) had at least 30 minutes of recorded use, suggesting the

file may have included all enrolled students at participating schools regardless of actual program engagement. Despite strong coverage across 129 districts and 409 schools, this mismatch between enrollment and engagement complicates interpretation.

Implementation and Usage Patterns

Among students with matched IDs and at least 30 minutes of usage:

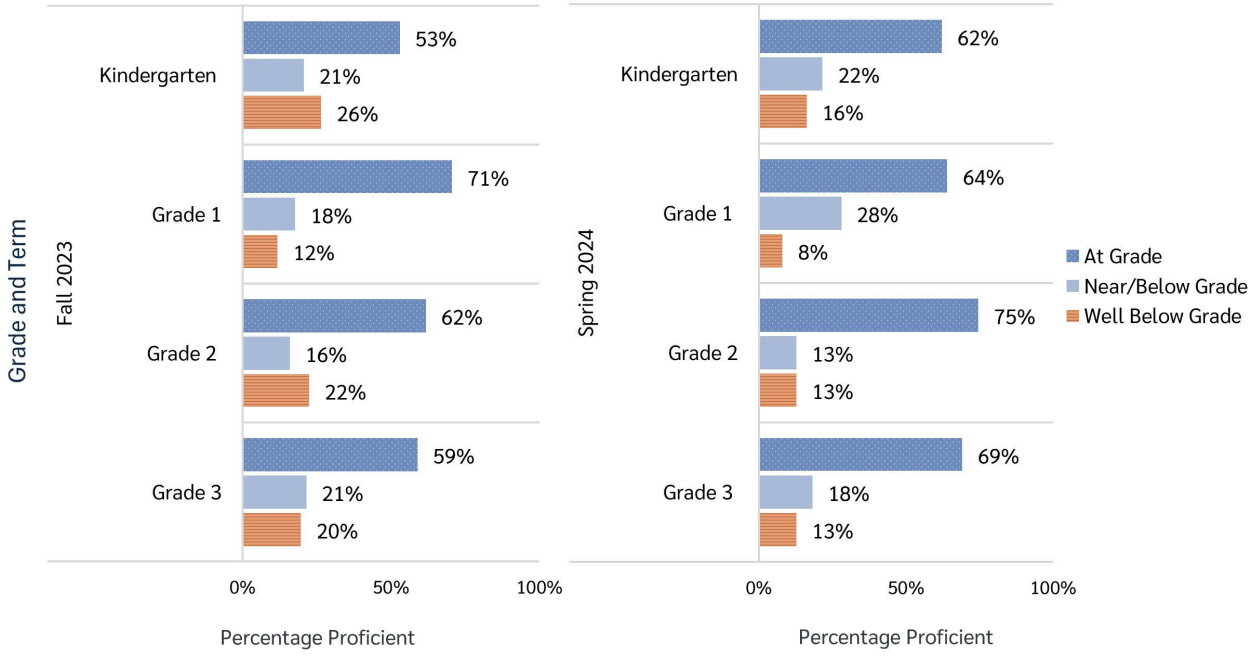
- No students were flagged by the vendor as meeting their recommended usage thresholds, possibly due to data reporting issues
- 7.89% met the state's 30-hour benchmark, and 2.55% exceeded 60 hours
- Average time in product was 903 minutes (approx. 15 hours), ranging from 721 minutes in grade 3 to 1,029 minutes in grade 1

These levels suggest moderate implementation but inconsistent fidelity, particularly relative to Amira's recommended engagement range.

IRI Outcomes and State Comparison

Across all grades, IRI proficiency rates improved over the year, and the percentage of students performing well below grade level decreased, mirroring statewide trends. While Spring IRI proficiency rates were equivalent to state averages, no separate analysis could be conducted for students meeting Amira's recommended usage due to missing indicator data.

Figure 14. Amira - IRI Proficiency Outcomes



Note. Percentage of all matched students by category.

Growth Beyond Proficiency: Score Analysis

Average IRI scale score growth was:

- Equivalent to the state average in grade 2
- Slightly below state averages in grades K, 1, and 3

However, all results fell within the state's expected growth range, and only 2.3% of students lost ground, which is consistent with statewide norms.

Table 18. Amira – Summary Outcomes for Students with Matched Scores and at Least 30 Minutes Usage

Grade Level	Average IRI Score Increase	Increase Compared to State Average	Is Average within Expected Growth Range	Average Time in Product (mins in SY)	% Students Losing Ground (Fall to Spring)
Kindergarten	92.2	Below	Yes	996	N/A
Grade 1	69.3	Below	Yes	1,029	N/A
Grade 2	66.5	Equivalent	Yes	867	N/A
Grade 3	54	Below	Yes	721	N/A
TOTAL	70.1	N/A	N/A	903	2.30%

Note. Includes percentages of students with matched scores who had at least 30 minutes in product.

Dosage and Differential Impacts

A wide distribution of time in product allowed for meaningful comparisons across usage bands. Students in higher time bands consistently showed greater increases in proficiency, with the 15–40-hour range appearing most effective. A small number of students logged more than 100 hours; however, diminishing returns were observed at the very highest usage levels.

Table 19. Amira - Average IRI Proficiency Rates by Usage Category

Usage Bucket	n	Fall Level Proficient (n)	Spring Level Proficient (n)	Fall % Proficient	Spring % Proficient	% Change
<10 hours	69,829	37,991	45,083	54.4%	64.6%	10.2%
10–19.99 hours	11,213	6,100	7,666	54.4%	68.4%	14.0%
20–29.99 hours	3,876	2,267	2,861	58.5%	73.8%	15.3%
30–39.99 hours	1,526	900	1,142	59.0%	74.8%	15.9%
40–69.99 hours	942	528	674	56.1%	71.5%	15.5%
70–99.99 hours	213	103	137	48.4%	64.3%	16.0%
100–149.99 hours	298	135	178	45.3%	59.7%	14.4%
150–199.99 hours	157	67	97	42.7%	61.8%	19.1%
200–299.99 hours	129	64	73	49.6%	56.6%	7.0%
300+ hours	175	86	112	49.1%	64.0%	14.9%

Note. Includes percentages of students with matched scores.

This trend reinforces findings from the vendor study: most students benefit from 20–40 minutes per week, while usage beyond that range may offer limited additional benefit depending on grade and context.

Conclusion

Amira demonstrated broad implementation across Idaho and high-quality internal assessment results. However, participation and usage fidelity were inconsistent, and vendor usage thresholds were not clearly marked in submitted data, limiting analysis. Despite this, IRI outcomes were positive and aligned with state expectations, and students across all usage levels generally made progress.

We recommend Amira remain on the Approved Vendor List. Emphasis for 2024–25 should include:

- Improved tracking of vendor usage benchmarks in submitted data
- Continued analysis of dosage effects
- Focused efforts to increase implementation fidelity across districts

With these improvements, Amira has strong potential to demonstrate its effectiveness more fully in future review cycles.

Waterford (Reading Academy)

Recommendation

Remain on List — Strong Evidence of Impact for Students with Sufficient Usage

Summary Highlights

- High data quality with a 92.6% match rate and 89.3% of matched students having measurable usage
- Increases in IRI proficiency rates across all grades, with fewer students testing well below grade level in Spring than in Fall
- Above-average IRI score growth in grades 1–3, and all grades within expected growth ranges
- Strong positive impact for students meeting vendor usage thresholds, especially in grade 2
- Evidence supports vendor-reported dosage thresholds, reinforcing the importance of implementation fidelity

Vendor-Reported Outcomes

Waterford reported that Idaho students who used its Waterford Reading Academy consistently, particularly those logging between 1,500 and 3,000+ minutes annually, achieved substantial reading gains. According to internal vendor analysis:

- Kindergarten students using the program for at least 1,500 minutes gained the equivalent of one year of growth
- First and second graders using it for 2,000–3,000 minutes achieved an average of 1 year and 2 months of growth
- Students exceeding 3,000 minutes showed gains up to 1 year and 5 months
- Eight students completed the Waterford program, exiting at or above a grade 3 reading level

Vendor data also showed a strong linear relationship between usage time and reading growth, reinforcing the product's recommended dosage guidelines. These findings highlight Waterford's potential, especially for early readers who benefit from targeted foundational skill development.

Data Matching and Evaluation Scope

Waterford submitted usage data for 1,452 students. After removing 93 records outside of the K–3 grade span, 1,259 students remained for matching, of which 1,214 (89.3%) had at least 30 minutes of usage. Data covered 12 schools across 8 districts. The match rate was 92.6%, indicating high data quality and reliable reporting.

Implementation and Usage Patterns

Waterford recommends:

- 1,500 minutes/year for kindergarten
- 2,000+ minutes/year for grades 1 and 2

In the matched dataset:

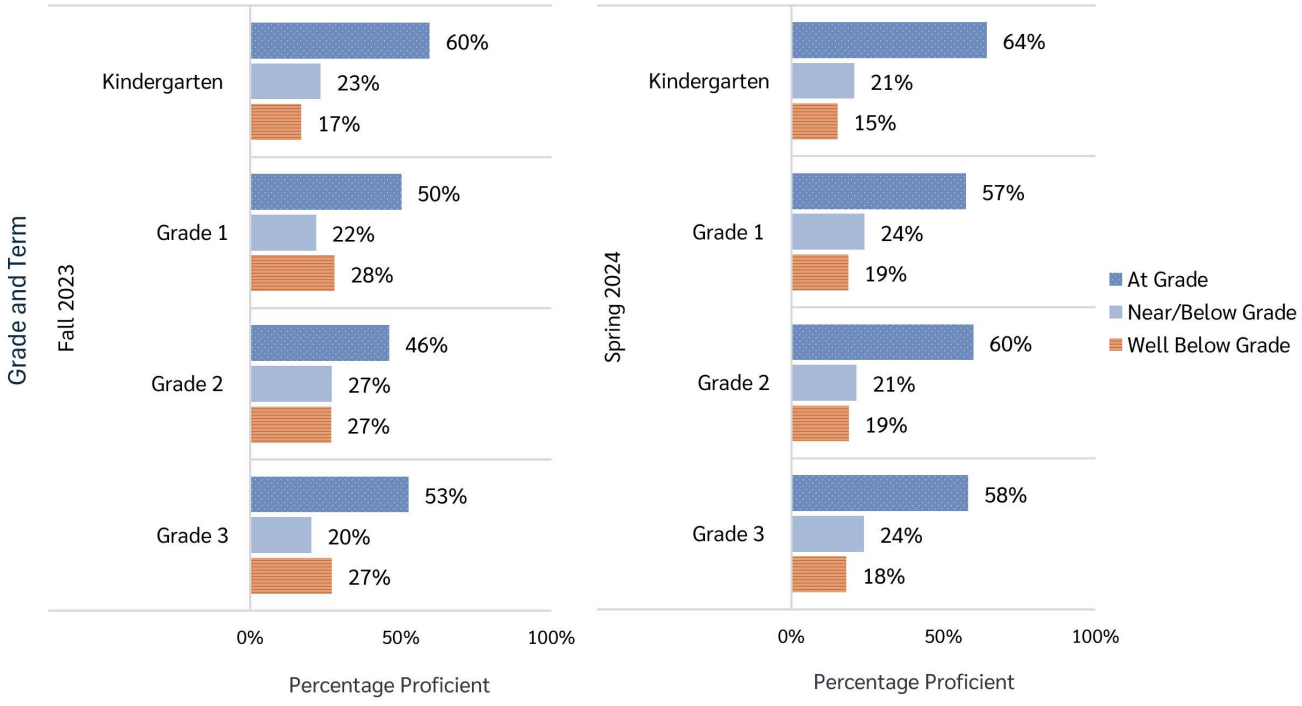
- 20.2% of students exceeded these thresholds
- 12.3% met the state recommendation of 20 hours (1,200 minutes)
- 0.64% met the 60-hour state target

Average time in product ranged from 748 minutes in kindergarten to 995 minutes in grade 2, with an overall average of 893 minutes (~15 hours). While below the recommended thresholds for many students, higher-usage groups demonstrated significantly better outcomes.

IRI Outcomes and State Comparison

IRI proficiency gains mirrored state patterns, with improvements in the percentage of students performing at or above grade level, and reductions in those testing well below grade level. Grade 2 students in particular outperformed the state average with a 14-point gain in proficiency.

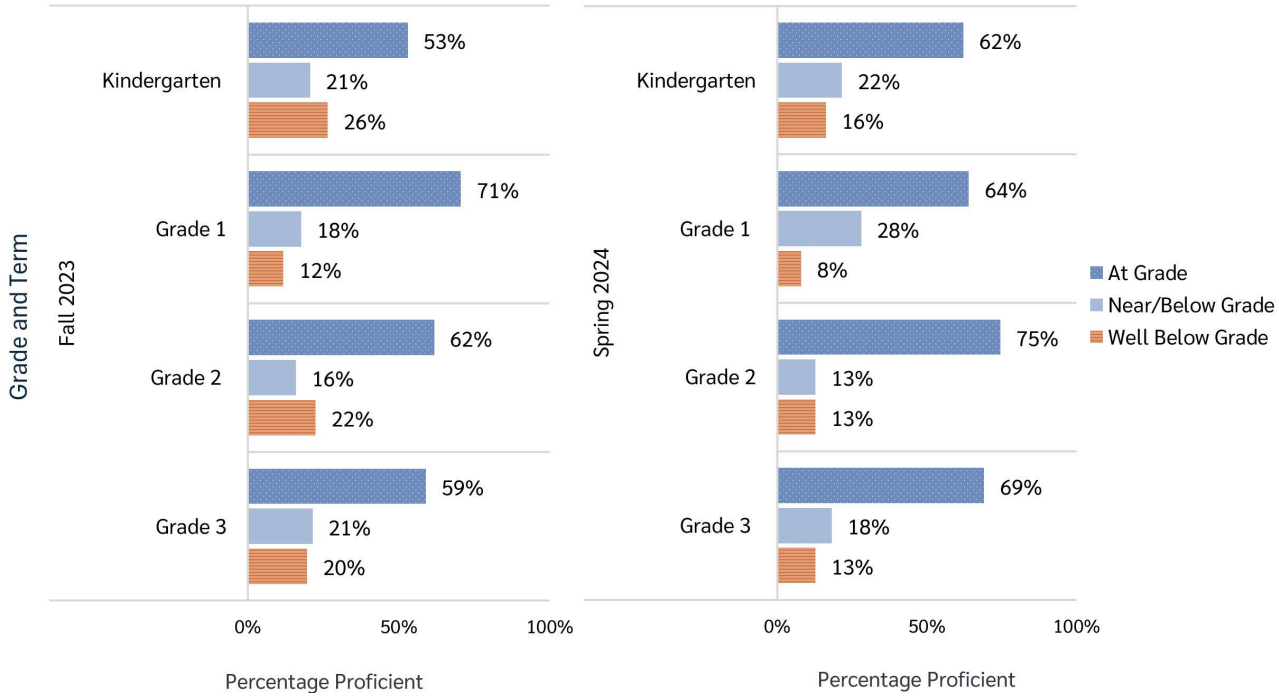
Figure 15. Waterford - IRI Proficiency Outcomes



Note. Percentage of all matched students by category.

Among students meeting vendor usage benchmarks, grade 2 Spring proficiency reached 75%, nearly 10 percentage points higher than the state average, supporting the effectiveness of the product when implemented with fidelity.

Figure 16. Waterford - IRI Proficiency Outcomes for Students Who Met Vendor Recommended Time Targets



Note. Percentage of all matched students by category.

Growth Beyond Proficiency: Score Analysis

Students in grades 1–3 posted above-average scale score increases on the IRI. All grades demonstrated gains within the state's expected growth range.

Table 20. Waterford - Summary Outcomes for Students with Matched Scores and at Least 30 Minutes Usage

Grade Level	Average IRI Score Increase	Increase Compared to State Average	Is Average within Expected Growth Range	Average Time in Product (mins in SY)	% Students Losing Ground (Fall to Spring)
Kindergarten	76.6	Below	Yes	748	N/A
Grade 1	79.6	Above	Yes	882	N/A
Grade 2	70.9	Above	Yes	995	N/A
Grade 3	59.3	Above	Yes	959	N/A
TOTAL	70.8	N/A	Yes	893	0.66%

Note. Includes percentages of students with matched scores who had at least 30 minutes in product.

Waterford had one of the lowest percentages of students losing ground between Fall and Spring: just 0.66%, well below the 2%–4% expected range.

Dosage and Differential Impacts

A strong positive correlation between time in product and outcomes was evident in both proficiency and scale score gains. Students who met Waterford's usage guidelines gained an average of 82.7 points, compared to 69.2 points for those who did not.

Table 21. Waterford - Average IRI Proficiency Rates by Usage Category

Usage Bucket	n	Fall Level Proficient (n)	Spring Level Proficient (n)	Fall % Proficient	Spring % Proficient	% Change
<25 hours	1,156	444	523	38.4%	45.2%	6.8%
25–33.3 hours	154	81	89	52.6%	57.8%	5.2%
33.3–41.6 hours	62	29	31	46.8%	50.0%	3.2%
41.6+ hours	81	43	55	53.1%	67.9%	14.8%

Note. Includes percentages of students with matched scores.

Students in the highest usage tier showed nearly 15 percentage points of improvement, supporting Waterford's dosage recommendations and highlighting its potential as a high-impact intervention.

Conclusion

Waterford Reading Academy demonstrated strong, consistent outcomes across all grades, with especially compelling results for students who met usage guidelines. Its high data quality, clear dosage-response patterns, and above-average IRI growth in multiple grades provide strong justification for keeping it on the list.

We recommend Waterford remain on the Approved Vendor List. Continued support for implementation fidelity, especially in reaching vendor-recommended usage thresholds, will be important for realizing full program impact in future years.

MARZANO
RESEARCH

Marzano Research is a women-owned small business dedicated to working with educators and system leaders to learn, evolve, and thrive.

We offer sophisticated education research and consulting capabilities with a collaborative, small-business approach.

As a certified benefit corporation, we believe in using business as a force for good and are committed to creating a positive impact for our employees, clients, stakeholders, community, and environment.

Denver, CO + Portland, OR | 720.463.3600 | www.MarzanoResearch.com