Introduction

In recent years, the United States has seen an unprecedented surge in concern regarding the effectiveness of literacy instruction within its classrooms (*What Exactly Is the Science of Reading?*, 2024). This concern, deeply rooted in the desire to enhance literacy outcomes for all students, has prompted a significant shift towards reevaluating and restructuring the methods by which children learn to read (Westall & Cummings, 2022). This movement has gained further momentum in the wake of the COVID-19 pandemic, highlighting the urgent need for reforms to address the compounded educational challenges faced by students nationwide. Policymakers, educational leaders, and teachers are once again at the forefront of efforts to reimagine literacy instruction, guided by a growing recognition of the critical role that early literacy plays in shaping a student's academic trajectory.

Central to this discourse is the acknowledgment of persistent achievement gaps that disproportionately affect students from low-income families, multilingual learners, and students with disabilities (Hart & Risley, 2003; U. S. Department of Education, 2022). These disparities not only underscore the importance of equitable access to quality education but also call for a nuanced understanding of the barriers to literacy faced by diverse student populations. In response, there is an increasing focus on the development and implementation of early literacy interventions, buoyed by a substantial body of research indicating that foundational literacy skills acquired in the early years are indispensable for later academic success (Council, 1998; Cunningham & Stanovich, 1997; Hernandez, 2011).

Moreover, the debate extends to the quality of teacher preparation and the availability of ongoing professional development opportunities. With evidence-based practices in literacy instruction gaining prominence, there is a pressing need to ensure that educators are equipped with the knowledge and skills required to effectively teach reading (Moats, 2020). This includes a thorough understanding of the Science of Reading (SOR), a body of research that encompasses the cognitive processes involved in learning to read, including phonemic awareness, phonics, fluency, vocabulary, and comprehension.

Furthermore, access to high-quality instructional materials is another critical factor in the literacy equation. The disparities in resource allocation among schools and districts highlight a significant challenge in ensuring that all students have the opportunity to engage with books, technology, and other essential learning tools (Morgan & Amerikaner, 2018). Legislators and educational leaders are therefore tasked with addressing these inequities, striving to create a landscape where every child has the resources needed to become a proficient reader.

Recent history of literacy work in San Diego Unified

Beginning in the early 2000s, the San Diego Unified School District (SDUSD) adopted a Balanced Literacy (BL) approach, which was part of a broader educational reform aimed at improving literacy outcomes across diverse student populations, including a significant number of English learners (ELs). During the administration of Superintendent Alan Bersin and Chancellor of Instruction Anthony Alvarado, the district implemented BL as a central component

of its instructional reforms. This approach emphasized the creation of meaning through active reader interaction with text, supported by various instructional activities and interaction strategies.

The district's reforms sought to address the varying literacy needs of all students by incorporating elements such as explicit literacy skills instruction, differentiated instruction, and accountable talk. However, despite these efforts, challenges remained in adequately meeting the needs of EL students. O'Day's investigation of the district's literacy progress suggested that while BL held promise, there were gaps in effectively addressing the specific language and literacy needs of the district's diverse EL population (2009).

Shifting forward a few decades to 2020, SDUSD launched a new literacy strategy to rapidly accelerate literacy outcomes for all students. The strategy sought to remake early literacy instruction from top to bottom, including the adoption of a new curriculum, implementation of curriculum-aligned assessments, a universal screening system, investments in teacher professional development and support, and the hiring of dozens of early literacy coaches to provide additional instruction in foundational literacy skills.

SDUSD transitioned towards implementing evidence-backed literacy practices in 2021. This shift came as a response to long-standing reliance on BL approaches that were increasingly criticized for their ineffectiveness in teaching foundational reading skills. With this approach, the district adopted a new curriculum called Benchmark Advance, aligning with evidence-based practices that emphasized systematic phonics instruction and other structured literacy components. Despite this progress, the implementation has faced some challenges, particularly due to the disruptions caused by the COVID-19 pandemic and the complexity of applying new strategies across a large and diverse school district.

SDUSD has also taken steps to improve literacy outcomes by forming a Literacy Working Group to advise on curriculum and best practices. This group developed a SOR primer for teachers and conducted district-wide environmental scan to assess the effectiveness of current instructional methods. The goal was to ensure that successful practices can be replicated across schools while identifying and addressing any ineffective strategies.

Additionally, the district invested in professional development and resources through a Literacy Acceleration Plan, reflecting a broader movement within the district to balance school autonomy with adherence to proven instructional practices. Overall, while there have been strides towards embracing SOR, SDUSD continues to navigate the complexities of large-scale implementation and aims to refine its approach to ensure all students benefit from effective literacy instruction.

Before adopting this new literacy strategy, SDUSD capitalized on its connections with major support organizations such as the Diamond Education Excellence Partnership (DEEP), the California Reading and Literacy Project (CRLP), and other organizations that support literacy initiatives. CRLP is a statewide professional development network that focuses on improving reading and literature instruction in K-12 schools. It is one of several California Subject Matter Projects aimed at enhancing subject-specific teaching skills. CRLP offers training

and resources to teachers to help them develop effective literacy instruction strategies, with a special emphasis on English language development and culturally responsive teaching to meet the diverse needs of California's student population. CRLP along with the support from DEEP have been instrumental in advancing SDUSD's strategic goals to enhance literacy results.

Key features of the DEEP program

The Diamond Educational Excellence Partnership, or DEEP, is a collaborative undertaking of multiple partner organizations and community leaders who are committed to improving educational outcomes for students attending public schools in the Diamond Community, a cluster of neighborhoods in southeastern San Diego that is home to many of the city's lower performing schools. The Diamond Community is an area known for its diverse population, rich cultural heritage, and vibrant community life. It includes a collection of communities such as Lincoln Park, Valencia Park, Emerald Hills, Chollas View, and Mount Hope. The Diamond Community is served by several public schools including Lincoln High School and 10 elementary schools including the four DEEP-supported schools: Chollas Mead, Johnson, Webster, and Encanto.

DEEP has several strategic priorities, including:

1. Strengthening the capacity of caregivers of young children to provide enriching early learning and healthy development experiences
2. Supporting schools to ensure that children have access to high-quality in-school literacy instruction and social-emotional support
3. Offering children extended learning opportunities that expand upon strong in-school teaching and learning

To achieve priority 2, DEEP partners with the CRLP and SDUSD to bring phonics-focused early reading instruction into four partner schools: Chollas-Mead, Encanto, Johnson, and Webster elementary schools. SDUSD's early literacy initiative collaborates and focuses on tailored small-group instruction in essential literacy skills for grades K-3.

Teachers use the Systematic Instruction in Phonological Awareness, Phonics, and Sight Words (SIPPS) curriculum to systematically teach phonological awareness, phonics, and sight words, enhancing students' decoding abilities, comprehension, and fluency with level-specific explicit instruction. Additionally, DEEP leaders in collaboration with DEEP principals collaborate and engage in monthly classroom walkthroughs in order to observe, review and refine classroom literacy practices. DEEP and CRLP also offer professional development centered on evidence-based reading strategies. Teachers at DEEP-affiliated schools benefit from continuous training in effective literacy techniques, data-driven instruction, and curriculum support, all based on the SOR framework. This first-year report and evaluation specifically focus on the literacy curriculum, K-3 instructional strategies, and the professional development aligned with DEEP priority two. Subsequent reports will address other DEEP priorities and aspects of literacy instruction.

The Science of Reading

SOR is a body of knowledge about teaching reading that draws on scientific research to understand how children learn to read and how reading instruction can be most effective. According to Petscher and colleagues (2020), SOR is defined as a phrase that represents the accumulated knowledge about reading and reading development, and best practices for reading instruction, all obtained through the scientific method. This includes a basis in neurological research, detailing how the brain functions to enable learners to connect sounds and oral language to letters and words, ultimately leading to the comprehension of a full passage.

SOR is often compared to a BL approach, heralded by SDUSD over 20 years ago. BL is an instructional approach that combines various methods to teach reading and writing. It includes components such as guided reading, shared reading, independent reading, read-alouds, and word study. The core philosophy of BL is to blend phonics (the systematic teaching of letter-sound relationships) with whole language approaches that emphasize comprehension and meaning-making through reading and writing. The goal is to develop students' literacy skills holistically, catering to a diversity of learners.

Both SOR and BL have similar goals, include phonics instruction, and attend to comprehension. However, these methods address phonics instruction differently. A BL approach uses a mix of phonics and holistic reading practices (such as combining reading, writing, speaking, and listening activities around thematic units), while SOR advocates for explicit, systematic instruction in phonics and other foundational reading skills. SOR emphasizes systematic and explicit instruction in the key components of reading, often referred to as the "Big Five" or "Five Pillars of Reading," and is grounded in research from various fields such as cognitive psychology, linguistics, neuroscience, and education.

The Five Pillars of Reading, as highlighted by the National Reading Panel (2000) and other research, include:

1. **Phonemic Awareness:**

   - The ability to recognize and manipulate individual sounds (phonemes) in spoken words. This involves tasks such as identifying, blending, and segmenting sounds.

2. **Phonics:**

   - The relationship between letters and their corresponding sounds. Phonics instruction teaches students how to decode words by recognizing the sound-symbol relationships.

3. **Fluency:**

   - The ability to read with speed, accuracy, and expression. Fluency is developed through repeated practice, often involving reading aloud.

4. **Vocabulary:**

- A rich vocabulary is crucial for comprehension. Instruction in vocabulary involves teaching the meanings of words, including strategies for understanding and acquiring new words.

5. **Comprehension:**

- The ability to understand and make meaning from text. Comprehension instruction includes strategies for understanding the structure of sentences and passages, making inferences, and summarizing information.

Key research-based principles of the SOR include (Child Health & Human Development, 2010; Joyce & Showers, 2002; Panel et al., 2000):

- **Systematic and Explicit Instruction:** SoR advocates for systematic and explicit instruction, where each skill is taught explicitly and in a planned sequence.

- **Multisensory Approaches:** Engaging multiple senses (auditory, visual, and kinesthetic) in learning to read is often recommended to reinforce connections between letters and sounds.

- **Early Intervention:** Identifying and addressing reading difficulties early on is crucial for preventing long-term reading challenges.

- **Evidence-Based Practices:** SoR emphasizes the importance of using instructional methods and interventions that have been proven effective through research.

- **Professional Development:** Educators are encouraged to receive ongoing professional development to stay informed about the latest research and effective instructional practices in reading.

Another common distinction between these two approaches involves instructional practices. A BL approach uses read-alouds, guided reading, independent reading, and writing activities, often focusing on student choice and engagement, while SOR focuses on a sequenced, skills-based approach, ensuring students master each component before moving on to the next.

Evidence of Effectiveness

The effectiveness of SOR is supported by a body of literature that underscores the positive impact of its key components on student reading development. Systematic phonics instruction, an integral part of SOR, has been shown to significantly enhance students' ability to read when compared to unsystematic or no phonics instruction. The National Reading Panel's (2000) phonics-centered meta-analysis of 38 experimental and quasi-experimental studies demonstrates that phonics instruction significantly boosts students' reading growth relative to those who received unsystematic or no phonics instruction.

Furthermore, the National Reading Panel's meta-analysis along with other empirical research corroborate the beneficial effects of phonics instruction among students facing reading challenges, who are low income, and multilingual learners (Bowers, 2020). This evidence showcases

the vital role of SOR components in informing instruction to positively influence student reading outcomes for diverse learners. Notably, Novicoff and Dee (2023) examined the impact of California's Early Literacy Support Block Grant (ELSBG) on improving early literacy among K-3 students in the state's lowest-performing elementary schools. The study found significant improvements in ELA achievement and smaller spillover improvements in math achievement in the first two years of implementation, suggesting the effectiveness of the SOR-based interventions when applied at scale. Other studies have highlighted the success of interventions focused on phonemic awareness, phonics, fluency, and comprehension in high-poverty schools (Suggate, 2016). Students who are low-income often face additional challenges that impact their literacy development, such as poor health, unstable living environments, and limited access to enrichment activities. These factors contribute to the risk of poor literacy outcomes for these students, evidenced by national data showing they are less likely to read proficiently by the end of Grade 4 compared to their more advantaged peers (Ferguson et al., 2007).

Recently, SOR has gained prominence in education policy discussions and teacher training programs. Advocates argue that a strong foundation in the key components of reading is essential for all students, including those with dyslexia and multilingual learners. However, other researchers disagree on the value and benefit of SOR for multilingual learners. Their criticisms center around SOR's failure to address cultural contexts, inadequate attention to bilingualism, and pedagogical rigidity, and how these issues might hinder the learning process for multilingual learners. These criticisms posit the need for more inclusive and flexible educational approaches that leverage the linguistic and cultural assets of all students (Tsokalidou & Skourtou, 2020; Xia & Haas, 2024). Empirical evidence aligns with some of these concerns, showing that while the SOR enhances decoding skills, it might not fully address the comprehensive literacy needs of multilingual learners. Research suggests that more effective strategies for these students often combine phonics with instruction focused on meaning, and actively incorporate elements of bilingualism. (Bialystok, 2018; Goldenberg, 2020; Great City Schools, 2023; O'Day, 2009; Polanco & Luft de Baker, 2018).

The academic and empirical debate regarding SOR is complex, with both proponents and critics using empirical evidence to support their perspectives. Against this backdrop, SDUSD partnered with the San Diego County Office of Education (SDCOE) in 2023 to evaluate its literacy programs across the Diamond community. The Diamond community serves as a proving ground to assess the effectiveness of various literacy curricula, including SOR, alongside professional development and coaching resources. This evaluation is particularly promising given four of the ten schools in the community were supplementing the curriculum with a SOR-based phonics approach, while the others employed a curriculum-based approach following Benchmark Advance. The aim of the evaluation is to gain comprehensive insights into the implementation and impact of the literacy curricula, including SOR, within this district cluster by understanding the added value of SIPPS and CRLP for students, especially multilingual learners.

Methods

This evaluation employed a mixed-methods approach to assess the impact of literacy curricula and supports on student language development within the district. This approach was appropriate to comprehensively understand the multifaceted literacy program by combining quantitative analysis of student performance data with qualitative insights from educators and stakeholders.

We collected achievement data using standardized reading assessments (aReading assessment), supplemented by school records of student demographics and attendance. We collected state assessment data to support a Beating the Odds (BTO) analysis. The quantitative data allowed for statistical analysis to evaluate literacy outcomes across the district and to understand the quality of literacy instruction relative to other schools in San Diego. Additionally, we conducted 9 semi-structured interviews with teachers and surveyed 150 teachers and administrators to gather in-depth information on their experiences and perceptions of the literacy programs. Finally, we observed 55 grade 2-4 literacy classrooms using the CLASS observational rubric to assess the fidelity of curriculum implementation and instructional quality.

We used multilevel modeling to analyze the reading achievement and CAASPP data, accounting for individual and school-level variables. This method was chosen for its ability to handle the hierarchical structure of educational data and to robustly estimate the effects of the literacy program while controlling for confounding factors. Thematic analysis was conducted on the interview and survey responses to identify common themes and insights into the program's implementation and impact. Classroom observation data were analyzed using descriptive statistics to summarize the instructional practices observed. The mixed-methods design provided a holistic view of the literacy program's implementation and outcomes. Quantitative analysis of achievement data allowed for objective measurement of program efficacy, while qualitative data offered contextual insights that helped explain the quantitative results. This combination was crucial for addressing the complex evaluation questions and providing actionable recommendations for the district.

Findings

We designed a mixed methods evaluation study to investigate the implementation and impact of literacy curricula and supports on student language development. These formative findings reflect perspectives and evidence gathered from the first year of a multiyear evaluation. Below, we have arrayed our findings by evaluation questions.

How are SDUSD schools in the Lincoln cluster implementing the early literacy program and what school-and-district-wide conditions support or hinder effective implementation?

Rolling out a curriculum or instructional practice across a large, urban school district like SDUSD involves numerous challenges, such as acquiring materials, training teachers, ensuring

administrators provide feedback, and offering ongoing support through instructional coaches. Like many districts across the country, SDUSD has navigated numerous obstacles while implementing its comprehensive literacy program. Over the past three years, this effort has required detailed planning and coordination across various levels, from district and site leadership down to the classroom. This comprehensive approach has encompassed not only the procurement of materials and professional development for teachers but also the provision of literacy coaching. Beyond academic supports, the district has enhanced its offerings with extended learning opportunities and collaborations with partners to enrich students' learning experiences and support their social-emotional development. Consequently, the effectiveness of the literacy efforts at both the school and district levels has been deeply influenced by these conditions, which have both facilitated and hindered the successful implementation of the literacy initiative.

DEEP Schools Offer More Systematic Phonics Instruction and Have Better Access to Support

In the initial year of this evaluation, we examined the implementation of curriculum and literacy practices across the district by interviewing and surveying teachers, conducting classroom observations during literacy blocks, and analyzing literacy data. Several conditions stood out in our analysis.

First, DEEP schools offer a systematic and structured approach to literacy instruction compared to other schools in the Lincoln cluster, especially with respect to phonics and phonemic awareness. DEEP schools' literacy approach integrates Systematic Instruction in Phonological Awareness, Phonics, and Sight Words (SIPPS) and the tenets from California Reading and Language Project (CRLP). Teachers acknowledge the value of their professional development experiences in SIPPS and CRLP.

For example, one teacher indicated:

> "We have our required PD and for the last five years we have completed professional development on foundational reading skills like phonics… it really goes into depth about patterns, and how language works the way it doesn't. And we really use that training, for example, we worked on comprehension…we found a really rich text for students, and we engaged in a whole group process of analyzing the text… examining keywords and key understandings. We go through the whole text using a planning protocol then we teach. It's how we use what we learned through PD."

This teacher's insights underscore the critical role of a comprehensive approach to literacy instruction, involving PD, curriculum, and instructional strategies, in equipping these teachers with the necessary skills to enhance student literacy and overcome curriculum-related obstacles.

Classroom observations corroborated these perceptions illustrating that teachers in DEEP schools provide greater *instructional support* for students developing language and literacy skills. Instructional Support (a domain from the Classroom Assessment Scoring System) evaluates the extent to which educators facilitate and enhance students' thinking, problem-solving

abilities, conversational skills, and vocabulary. Teachers who excel in this area engage students by linking concepts and skills to real-life contexts, posing questions that stimulate analysis and reasoning, offering appropriate assistance, and giving feedback that values students' efforts. Overall, DEEP schools have a more pronounced impact on fostering cognitive and language development compared to their non-DEEP counterparts. Importantly, classrooms, with this focused instructional support, are more likely to yield better student outcomes in terms of cognitive and language development compared to other classrooms. District leaders should consider scaling programs that provide comprehensive instructional support, as these have proven to be beneficial in improving student learning outcomes (Soto, 2023).

In contrast to specific references to SIPPS and CRLP, many teachers from both DEEP and non-DEEP classrooms described professional development as infrequent and often insufficiently detailed. For example, one teacher noted, "The professional development offered has been sporadic and not sufficiently tailored to our needs." Teachers expressed a need for more practical, classroom-focused PD sessions on trauma-informed instruction, English Language Development (ELD), and using the Benchmark curriculum. Strengthening professional development offerings could enhance teacher effectiveness and their use of evidence-based instructional practices.

Teachers in DEEP schools also described having more regular access to literacy coaches and peer mentoring. DEEP teachers emphasized the importance of coaching for implementing ongoing professional development focused on foundational reading skills. One teacher noted:

> "The literacy coach pushes in as needed as teachers have questions and needs. One week she might be meeting with kids, coaching teachers, leading professional development, making copies for parent meetings…she would go on a couple of walkthroughs of classrooms to look at how we are all rolling out, see what we're doing…checking on the data giving feedback about which kids are in the red so to speak. Asking what are we going to do for them? And then providing that support."

These teachers recognized how peer coaching enhanced their ability to support student literacy development. Whether coaches work with a small group of struggling students or follow coaching cycles involving observations of student engagement, the regular access to these coaches (and professional development) has played a crucial role in the successful implementation of structured literacy practices. These are important findings given that ongoing professional development and access to coaching support are essential ingredients for a comprehensive teacher support system and for the effective implementation of any schoolwide initiative. High quality professional development and coaching support provides teachers with access to the latest research-based strategies and techniques to support student literacy development (Kraft et al., 2018; Neufeld & Roper, 2003). A comprehensive support system not only improves instructional quality but also fosters a culture of continuous learning and improvement, ultimately leading to better literacy outcomes for all students.

Inconsistency in Curriculum Adoption and Literacy Practices Exist Across the Cluster

While some teachers acknowledged SIPPS and CRLP as curricula and tools that support their literacy instruction, most teachers identified how inconsistencies in curriculum adoption and literacy practices hinder their ability to create a cohesive literacy experience for their students. Interviews and survey results highlighted variability in how schools and teachers adopted the curriculum. While some teachers adhered strictly to the Benchmark curriculum, others incorporated other programs, such as i-Ready and the University of Florida Literacy Institute. Many teachers supplemented their instruction with resources from platforms like Teachers Pay Teachers. This inconsistency was observed both across different schools and within individual schools. The lack of a cohesive curriculum increases variability in student literacy development and increases teachers' reliance on supplementary materials. This is important because the variability in implementation of research-backed curriculum materials exacerbate educational inequities, particularly affecting students from under-resourced communities (Bugler, 2017).

Table 1: Frequency of Instructional Materials by Type

Curriculum	Resource Type	Frequency
Benchmark (Advance, Adelante)	Core Curriculum	89%
Heggerty	Supplemental/Intervention	60%
Okapi	Supplemental/Intervention	57%
Fountas and Pinnell	Supplemental/Intervention	43%
Teachers Pay Teachers	Teacher Resource	29%
Ready to Advance	Core Curriculum	28%
iReady	Core Curriculum	21%
CRLP	Professional Development	18%
Handwriting without Tears	Supplemental/Intervention	17%
SIPPS	Supplemental/Intervention	16%
Achieve 3000	Supplemental/Intervention	5%
Lucy Calkins Reading and Writing Workshop	Core Curriculum	2%
Orton Gillingham	Supplemental/Intervention	3%
Lucy Calkins Units of Study	Core Curriculum	3%
Joyful Literacy	Supplemental/Intervention	2%
Fundations	Supplemental/Intervention	2%
Scholastic News	Teacher Resource	1%
Write Source	Teacher Resource	0%

This table illustrates the frequency and type of curriculum teachers use in the cluster. We derived the resource type labels from examining and researching these resources, and as such, these are not labels used by the district. For example, i-Ready is identified as a core curriculum on the Curriculum Associates i-Ready website and by the California Department of Education. Based on the information in this chart, Benchmark is the most frequently used curriculum (89% of classrooms). While the frequent use of the Benchmark curriculum indicates a strong dependency on this foundational material, the extensive adoption of resources like Heggerty,

Okapi, and Teachers Pay Teachers highlights the critical role of supplemental and intervention tools. These observations suggest a potential need to develop a more explicit and uniform district-wide literacy framework. Additionally, there is a need for more inclusive and adaptable curriculum materials that not only align with the core curriculum but also meet the varied needs of the district's diverse student population.

In addition to curriculum resources, teachers must also leverage various evidence-based instructional practices to deliver the curriculum. The What Works Clearinghouse's (WWC) Educator's Practice Guide in foundational reading skills identifies several high leverage, evidence-based instructional practices that support literacy and language development (Foorman et al., 2016). These include:

- Using read-aloud conversations to develop inferential language skills by asking open-ended questions that encourage students to think beyond the immediate context
- Using Elkonin sound boxes and letter tiles to help students link their knowledge of letter sounds with phonemic awareness
- Teaching students to blend letter sounds and sound-spelling patterns from left to right within a word to produce a recognizable pronunciation
- Teaching regular and irregular high-frequency words so that students can recognize them efficiently
- Engaging students in activities that involve adding, removing, or changing letters to create new words
- Providing opportunities for oral reading practice with feedback to develop fluent and accurate reading with expression
- Teaching students to self-monitor their understanding of the text and to self-correct word-reading errors
- Modeling strategies, scaffolding, and providing feedback to support accurate and efficient word identification

While teachers deploy many of these instructional practices across the cluster, implementation varies significantly both within and between schools. These variations exist due to gaps in teacher awareness and knowledge, professional development offerings, and the demands placed on teachers to support diverse learners. In interviews, teachers expressed that they spent considerable time modifying materials and resources to meet the diverse needs of their students, which diverted their attention from exploring and integrating new or innovative evidence-based instructional strategies. One teacher highlighted the challenges with the Benchmark curriculum, noting, "The Benchmark curriculum… gives you a lot of information, but not always useful information. It's not always clear how to implement all of the different products that they have." Teachers are finding that while the curriculum is rich in content, it often lacks clear instructions on how to effectively use the various components in a real classroom setting. This leads to a situation where teachers have to spend extra time figuring out how to best apply these resources, which can be especially challenging in classrooms with diverse learning needs. As a result, there is a need for the curriculum developers to refine these materials to

provide clearer, more actionable guidance that aligns with the day-to-day realities of teaching diverse groups of students.

Additionally, many teachers indicated feeling unprepared and overwhelmed to independently learn new literacy practices and strategies and align these practices to the diverse needs of their students. For instance, only 14% of teachers indicated in our survey how their professional development opportunities sufficiently equipped them to support students with dyslexia. Similarly low levels of confidence were reported for supporting students with disabilities and those experiencing homelessness.

The variability in instructional strategies suggests that there is no standardized approach across the cluster to teaching literacy, which is likely to lead to differences in the quality and effectiveness of the instruction students receive. However, a one-sized-fits-all literacy approach may be equally disruptive to instructional quality and effectiveness. The district might find the right balance of standardization and autonomy within the four Lincoln cluster schools who are outperforming expectations (see evaluation question 2). Additionally, the adaptation of materials and resources by teachers indicates a proactive effort to meet diverse student needs, but it also underscores a lack of uniformity that could impact overall educational outcomes. Some teachers are employing a mix of strategies to deliver content, which reflects flexibility and responsiveness to student needs. However, this also points to a potential gap in the Benchmark curriculum, as it requires teachers to modify and supplement materials on their own. These modifications can lead to inconsistencies and may not always result in the most effective instructional practices. There is a clear need to standardize some practices across the cluster to ensure instruction is consistent, evidence-based, and equitable. This might involve developing a more uniform and district-wide literacy framework and curriculum that provides structured support for diverse learners. Additionally, district leaders need to enhance professional development offerings to better prepare teachers for supporting students with diverse needs, including specialized training for supporting vulnerable students with dyslexia, other disabilities, and multilingual learners.

Implementing curriculum and evidence-based instructional practices across a large, urban school district like SDUSD is inherently complex, involving meticulous coordination and support at multiple levels. The district's efforts in planning, providing materials, professional development, and ongoing support are commendable, yet the effectiveness of these efforts is significantly influenced by the conditions at both school and district levels. This interim analysis shows that DEEP schools, with their emphasis on phonics, phonemic awareness, and professional development, are making notable strides in literacy implementation and enhancing student cognitive development. Non-DEEP schools might benefit from integrating similar literacy practices and adopting DEEP's approach to professional development. Despite progress in DEEP schools, inconsistencies in curriculum implementation across the cluster and insufficient ongoing support for evidence-based literacy practices highlight the need for a more standardized approach across all schools. Providing instructional materials, enhancing professional development, and ensuring consistent support will be essential for improving educational outcomes and addressing the diverse needs of all students.

How is the DEEP literacy model affecting student literacy outcomes?

Examining the impact of literacy curricula, practices, and strategies on student achievement is crucial for several reasons. Literacy is a foundational skill that underpins all areas of education and is essential for students' academic success and future opportunities. Understanding which curricula and instructional practices are most effective can help educators and policymakers make informed decisions that enhance teaching and learning outcomes. Additionally, as diverse student populations present varying needs, it is vital to identify strategies that support all learners, including those with learning disabilities, multilingual learners, and students from under-resourced backgrounds. By evaluating the effectiveness of different literacy approaches, schools can ensure they are providing equitable and high-quality education that fosters student achievement and prepares them for lifelong learning.

Several Lincoln Cluster Schools Outperform Expectations Relative to Other Elementary Schools

A Beating the Odds (BTO) analysis is a method used by educational researchers to identify schools that outperform expectations based on their unique student populations. Essentially, this analysis involves predicting a school's performance using demographic and prior academic data, and then comparing these predictions to the school's actual performance. Schools that perform significantly better than predicted are said to be "beating the odds," indicating that they are achieving better results than similar schools with comparable student demographics (R Core Team, 2023).

By spotlighting schools that outperform their predicted outcomes, BTO analysis provides valuable insights into effective educational strategies and practices that can be replicated in other schools. It serves as a tool for recognizing and learning from success stories, enabling policymakers, educators, and communities to adopt and implement strategies that have proven successful in similar contexts. This approach encourages a more equitable assessment of school performance, focusing on resilience and effectiveness in fostering student achievement against all odds.

We conducted a BTO analysis on all elementary schools in San Diego County using a regression model with student demographic information and prior CASSPP achievement data to forecast current CASSPP scores. This allows us to determine which schools are performing better than expected. We specifically highlighted how schools in the Lincoln cluster compare to other San Diego elementary schools. The results are visualized in a rain plot, with schools performing below, at, and above expectations. Lincoln cluster schools are represented by light and dark blue dots, while other San Diego elementary schools are shown as grey dots.

In this plot, performance is plotted on the x-axis as percentiles, which help us rank schools based on the difference between their actual and predicted performance. The y-axis labels indicate qualitative differences in performance: schools in the 70th percentile or higher are considered to have significantly outperformed expectations (increased performance), schools between the 0 and 30th percentiles are underperforming (decreased performance), and those in the 30th to 70th percentiles are performing as expected (Not BTO).

San Diego Elementary Schools Beating the Odds in Reading Achievement

Both DEEP (dark blue) and nonDEEP (light blue) schools perform mostly as expected while some are beating the odds.
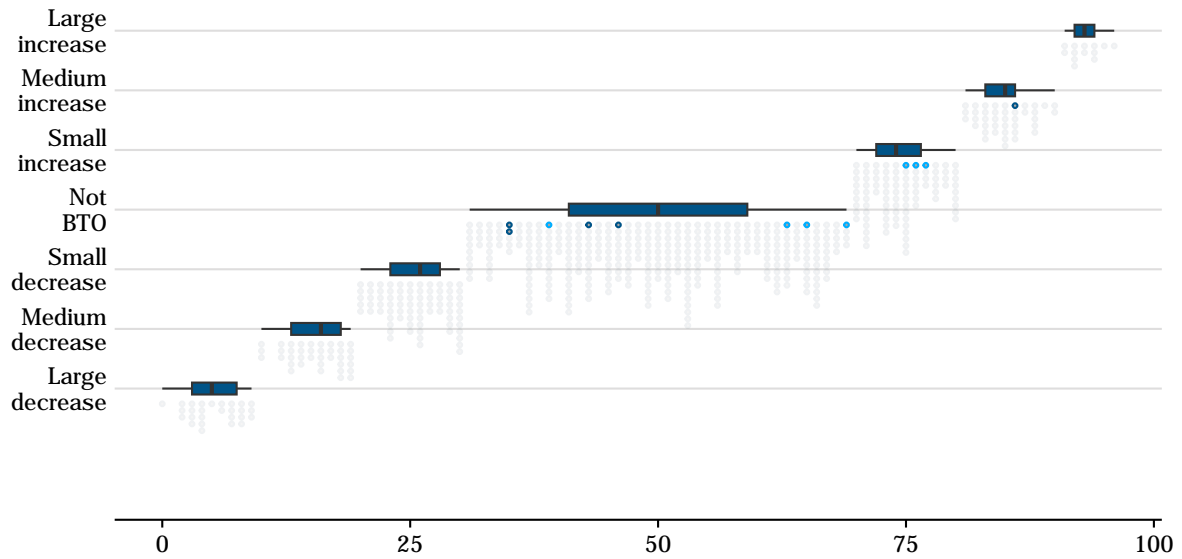
Source: Data and Impact Center for Excellence

SDCOE

Note: DEEP and nonDEEP schools ordered from largest decrease to largest increase: Johnson (DEEP), Horton, Encanto (DEEP), Webster (DEEP), Valencia Park, Chavez, Baker, Nye, Porter, Balboa, and Chollas/Mead (DEEP).

Figure 1: BTO Analysis of Lincoln Cluster and Other San Diego County Schools

From our observations, four schools in the Lincoln cluster show small to medium increases in performance, indicating that they are beating the odds. Three schools are performing as expected but are close to outperforming, and five others are performing precisely as predicted. It is important to note that a "Not BTO" designation means that a school is performing in line with what is expected based on its student population.

The value of a BTO analysis lies in its ability to account for the specific challenges and circumstances of each school, providing a fairer comparison than simply ranking schools by test scores alone. This adjustment considers factors like the economic backgrounds of students or the resources available at each school, ensuring a more equitable assessment of school performance. The standout performance of Chollas-Mead, Balboa, Porter, and Nye indicates that these schools have implemented effective teaching practices and strategies that allow students to excel academically. These schools are likely using innovative approaches or have strong support systems that help students overcome barriers commonly associated with under-resourced environments. The success of these schools provides a model that other schools in similar circumstances can emulate.

Table 2: BTO Percentiles and Performance Levels of Lincoln Cluster Schools

School Name	Estimated Performance	Percentile (th)	Performance Level
Chollas/Mead Elementary	7.947140	86	Medium increase
Balboa Elementary	5.274959	76	Small increase
Porter Elementary	5.132473	77	Small increase
Nye Elementary	4.974271	75	Small increase
Baker Elementary	3.837845	69	Not BTO
Chavez Elementary	2.796907	65	Not BTO
Valencia Park Elementary	2.503307	63	Not BTO
Webster Elementary	-0.601550	46	Not BTO
Encanto Elementary	-1.143286	43	Not BTO
Horton Elementary	-1.816496	39	Not BTO
Johnson Elementary	-2.738606	35	Not BTO

For instance, in the table above, Chollas-Mead, designated as a DEEP school, is performing in the 86th percentile, *meaning it outperforms 86% of elementary schools in the county in terms of reading achievement.* Chollas-Mead, along with Balboa, Porter, and Nye, are all identified as beating the odds. This analysis showcases their exceptional performance given their contexts. Standard evaluations often use average test scores to measure school performance, which can unfairly penalize schools in less advantaged areas. BTO analysis goes deeper, accounting for the specific challenges faced by schools and highlighting those that are performing unexpectedly well. This analysis helps shift focus from what schools are lacking to what they are accomplishing, promoting a more equitable approach to evaluating educational success. It rec-

ognizes and rewards schools for doing more with less, shifting policy focus towards supporting and expanding what works in these high-achieving, under-resourced schools.

As diverse student populations present varying needs, it is vital to identify strategies that support all learners, including those with learning disabilities, multilingual learners, and students from under-resourced backgrounds. By evaluating the effectiveness of different literacy approaches, schools can ensure they are providing equitable and high-quality education that fosters student achievement and prepares them for lifelong learning. Chollas-Mead, Balboa, Porter, and Nye are performing better than expected, given the student populations they serve. This suggests that these schools are likely implementing effective literacy practices and strategies that are successfully meeting the needs of their students. Understanding these practices can help disseminate successful approaches across the district. Schools that perform well despite challenges presented by economic, social, and racial inequities demonstrate that equitable education is achievable with the right strategies and support systems. District and school leaders should consider replicating the literacy practices of these high-performing BTO schools across other schools in the district.

Some Evidence of Enhanced Literacy Outcomes in DEEP Schools

Previously, we identified how teachers perceived the investment in professional development around SIPPS and CRLP led to significant improvements in teaching practices and student literacy outcomes, how coaching and peer mentoring were crucial to adopting and improving instructional practices to meet diverse student needs, and how DEEP schools showed better performance in instructional support (as defined by the CLASS observation instrument) compared to non-DEEP schools (La Paro & Pianta, 2003). Armed with these findings, we turned our attention to local literacy achievement results and how Lincoln cluster schools performed relative to each other.

We collected data that SDUSD uses to track literacy progress across all its elementary schools. The FastBridge aReading assessments provide a reliable gauge of a student's overall literacy abilities. The aReading assessment is a computer-adaptive test designed to evaluate a wide range of literacy skills, including phonemic awareness, phonics, vocabulary, comprehension, and fluency. It adjusts the difficulty of questions in real time based on the student's responses, providing a precise measure of their reading ability. The results from aReading are used to identify students' strengths and weaknesses, monitor progress over time, and inform instruction tailored to meet each student's needs. This makes it a valuable tool for educators in tracking literacy development and making data-driven decisions to enhance student learning outcomes.

Below are descriptive statistics for both the DEEP and non-DEEP schools along with testing results.

Metric	DEEP	Non DEEP
Total Students	4440	9422
Female (%)	48	48

Hispanic (%)	71	73
EL (%)	29	37
SPED (%)	19	19
Mean Pre aFast	475	471
Mean Post aFast	486	481
Mean Days Absent	17	18
Difference	11	10

As indicated in the table, our data included approximately 4,400 records from DEEP students participating in the aReading assessment and 9,422 non-DEEP students. SDSUD assesses students twice a year, so these totals reflect multiple records per student. The demographics of four DEEP and six non-DEEP schools are similar including the percentages of females, students who are Hispanic, English learners, and students with disabilities. Other data includes attendance rates, where on average students miss 17 to 18 days of instruction per year, which is a significant amount of time given that 18 days is the threshold for chronic absenteeism. The average pretest aReading scores for DEEP and non-DEEP students were 475 and 471 respectively. The scale scores from the assessment range from 350 to 750 and vary according to grade level. Average post-test scores were 486 and 471 respectively. While non-DEEP students' scores were slightly below those of DEEP students, both groups experienced a 10-point increase between the fall and spring administrations. These results suggest steady improvements in reading scores in 2023-2024 in both programs and that students are growing on average as expected.

To ascertain whether the DEEP schools' emphasis on phonics and phonemic awareness (with SIPPS and CRLP) improved literacy outcomes more than non-DEEP schools, we conducted a multilevel model (MLM) analysis. In addition to examining differences between the two literacy approaches, we also focused on each program's impact on ELs since ELs represent a special case for literacy programs. English learners face unique challenges in acquiring literacy skills because they must simultaneously learn a new language and develop fundamental reading and writing abilities. Unlike native English speakers, ELs need targeted support to build their vocabulary, understand grammar and syntax, and develop phonological awareness in English. This multifaceted challenge can hinder their overall academic progress if not addressed adequately. With this in mind, we were interested in seeing if the DEEP model was more effective for multilingual learners.

After excluding students in grades K-2, our final dataset included data on 4,319 unique students. We further removed 1,788 records for students who were missing either a post-test score or a corresponding pre-test score, with most exclusions due to missing post-test scores. We checked to ensure the remaining data was balanced across various factors in our model, and with the remaining 2,531 records, we formed groups that could be fairly compared.

To ensure comparable groups, we performed propensity score matching (PSM) using pretest scores along with demographic variables, including race/ethnicity, gender, status for free and

reduced-price lunch and special education. This method paired students in the DEEP schools with those in non-DEEP schools, ensuring baseline equivalence across all covariates and pretest measures. PSM is particularly valuable in experimental designs that do not allow for random assignment, as it aims to mitigate bias in estimating treatment effects (Leite, 2017). Our model predicted post-test literacy scores from a combination of student (e.g., race, gender, socioeconomic status, and others) and teacher factors (e.g., degree type and years of experience), along with the literacy program in which students participated.

Factors Influencing Reading Achievement in Lincoln Cluster Schools
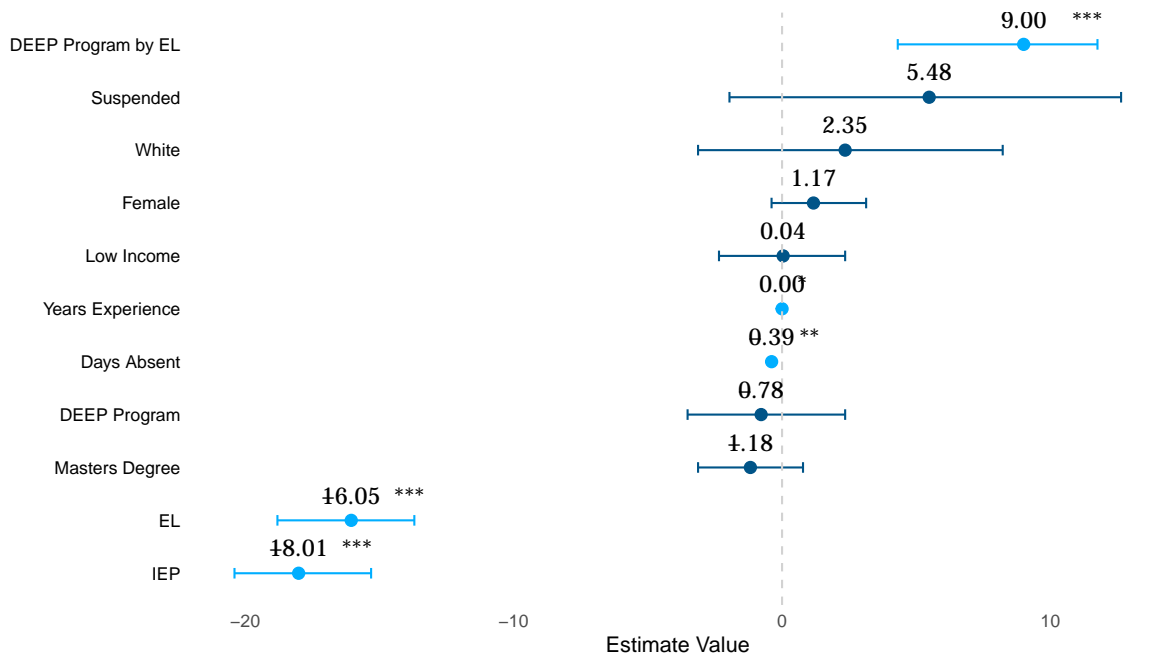
Source: Data and Impact Center for Excellence

Note: This plot shows 12 factors and their influence on reading achievement. The dots represent the difference between the factor and its counterpart (e.g., DEEP compared to nonÐEEP schools). The bars represent the range of differences from the minimum difference to the maximum difference.

Figure 2: Coefficient Plot of Factors Influencing Reading Achievement

Based on the results of our analysis, some effects are worth noting. First, students in DEEP schools perform slightly lower compared to students in non-DEEP schools on the aReading assessment by 0.78 points after controlling for demographic, teacher, and school factors that typically influence reading achievement. However, this difference is not statistically significant (p = 0.67), which means that on average students performed about the same regardless of the literacy models.

Additionally, both English learners and students in special education programs (IEP) scored significantly lower by 16.05 and 18.01 points respectively on the aReading test compared to their non-English learner and non-special education counterparts (p < 0.001). Both results

indicate a negative impact of being an English learner or having a disability on reading performance regardless of the school attended. English learners and students with disabilities struggle to develop language and literacy skills on-time in the cluster.

Furthermore, student absences significantly impacted aReading results, with each additional day absent reducing the aReading score by 0.39 points. For the average student with 17 absences, *the student's score was 6.63 points lower than it would have been had the student not missed any school.* Absences have a pronounced negative effect on reading achievement, regardless of the literacy model in place.

Finally, the interaction between attending a DEEP school and being an EL significantly improves reading performance by an additional 9 points (p = 0.0002). EL students who participate in the DEEP literacy model tend to score about 9 points higher on reading tests compared to those who are in the non-DEEP model. The difference is significant enough to *suggest that the DEEP model helps improve reading scores for the average EL student better than the literacy models in non-DEEP schools.*

Interaction Plot of Program and EL Status on Reading Assessment
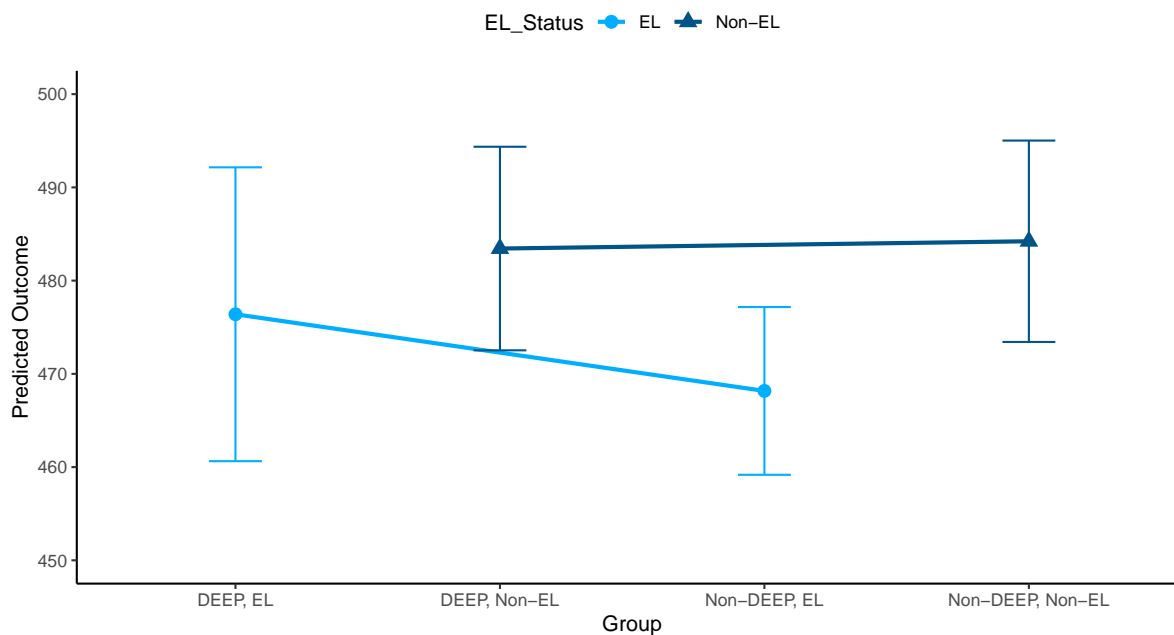
Source: Data and Impact Center for Excellence

Note: This plot shows the interaction between EL status and literacy model on reading achievement. The dots represent the difference between literacy outcomes for DEEP students compared to nonÐEEP students. The triangles reflect the difference between nonELs in DEEP schools compared to nonELs in nonÐEEP schools in the cluster.

Figure 3: Interaction Plot of Program and EL Status on Reading Assessment

The interaction plot illustrates the impact of participation in the DEEP literacy model (DEEP vs. Non-DEEP) and EL status (EL vs. Non-EL) on the aFast reading assessment scores. The

vertical axis presents the model-predicted reading scores, while the horizontal axis categorizes participants into four groups based on their DEEP enrollment and EL status. The plot reveals that ELs in DEEP schools tend to have higher predicted reading scores compared to ELs in Non-DEEP schools. This suggests that participation in a DEEP school may be beneficial for the literacy development of ELs. Conversely, for Non-ELs, the program does not exhibit a definitive advantage, indicating that its benefits may be more pronounced or specific to EL students.

The analysis sheds light on the efficacy of the DEEP model compared to its non-DEEP counterparts, highlighting particular challenges faced by specific student groups and the influence of factors like absences and program involvement on their literacy outcomes. Notably, ELs in DEEP-affiliated schools exhibited higher literacy growth than those in non-DEEP schools. This finding tentatively suggests that an SOR approach, often criticized for its primarily monolingual focus, might actually be beneficial for multilingual learners. Although these findings challenge common perceptions about the suitability of SOR methods for ELs, it is important to clarify that our evaluation aimed to compare the effectiveness of literacy instruction and curriculum in the Lincoln cluster by analyzing student performance on the aReading assessment, rather than directly assessing the specific efficacy of SOR for ELs. Consequently, our evaluation may have overlooked factors necessary to definitively support claims about SOR's effectiveness for ELs. Future research should further investigate and validate this finding. There is an urgent need for more studies on multilingual learners, especially those integrating biliteracy principles into SOR frameworks. Such research could contribute to developing a more inclusive and effective literacy framework that better serves the diverse linguistic needs of multilingual learners (Cummins, 2000; Garcia & Lin, 2009).

Furthermore, the significant impact of absences on reading scores also highlights the critical need for strategies to improve student attendance. Schools should implement policies and support systems that encourage regular attendance and address barriers that keep students from attending school.

Finally, district leaders and school administrators should consider these findings when allocating resources and re-designing their literacy programs. Investing in programs that show potential benefits for vulnerable student populations, and ensuring consistent implementation across schools can help reduce any educational disparities.

Bright Spots and Lessons Learned

Effective Professional Development

Regular, structured professional development sessions focused on foundational reading skills and instructional strategies, such as those provided by SIPPS and CRLP, were highly effective in enhancing teacher practices and student outcomes.

In surveys and interviews, teachers and coaches highlighted the high-quality professional learning experiences provided by SIPPS and CRLP, which equipped them with effective strategies to enhance foundational reading skills. These sessions not only offered theoretical knowledge but also practical, hands-on approaches that teachers could directly apply in their classrooms. Teachers reported feeling more confident and equipped to address diverse student needs, particularly in phonics and phonemic awareness.

The success of SIPPS and CRLP professional development underscores the importance of investing in high-quality, continuous professional learning opportunities for teachers. It highlights how well-structured and supportive training programs can significantly enhance educational practices and student achievement.

Collaborative Support Systems

The presence of literacy coaches and teacher leaders providing ongoing support and mentorship was a significant factor in effectively implementing literacy practices.

Whether through formal literacy coaching or peer support within grade-level teams, teachers uniformly attributed their implementation success to these collaborative supports. Literacy coaches played a crucial role in offering personalized guidance, modeling instructional strategies, and providing timely feedback, which helped teachers refine their practices and address specific classroom challenges. Peer support through grade-level teams fostered a sense of community and collective responsibility, enabling teachers to share resources, discuss strategies, and problem-solve together.

Teachers also highlighted the challenges they faced when these support systems were lacking. The absence of coaching support was linked to struggles in implementing literacy practices effectively, underscoring the critical role that collaborative support systems play in teacher development and student success.

This collaborative support system is essential for the growth and development of teachers, as it provides a structured yet flexible framework for continuous learning and improvement.

Data-Driven Instruction

Utilizing data from assessments to inform instructional practices and interventions helped in addressing diverse student needs and improving literacy outcomes.

Teachers emphasized the importance of using data to monitor and respond to student progress. By employing flexible grouping and targeted interventions based on assessment data, they were able to tailor instruction to meet the specific needs of each student. This data-driven approach enabled teachers to identify areas where students were struggling and to implement timely and effective interventions to support their learning.

These findings underscore the need for streamlined assessments and opportunities to use data more effectively. Teachers highlighted how having access to clear, actionable data allowed them to make informed decisions about their instructional practices, thereby enhancing their ability to support diverse learners. Regularly analyzing assessment data helped in setting

realistic goals, tracking student progress, and adjusting teaching strategies to better align with students' evolving needs.

This focus on data-driven instruction is crucial for creating a responsive and adaptive learning environment. It ensures that all students, regardless of their starting point, receive the support they need to achieve literacy proficiency.

Challenges and Barriers

Inconsistent Curriculum Implementation

Variability in the adoption and use of the Benchmark curriculum has led to significant inconsistencies in student preparation across the district, resulting in disparities in educational outcomes. Some schools fully embrace the Benchmark curriculum, while others use it selectively or supplement it with other materials like Teachers Pay Teachers and iReady. This inconsistency arises from varying levels of training, differing interpretations of the curriculum's requirements, and a lack of resources to accommodate diverse learners.

Due to these inconsistencies, many teachers rely on supplementary materials to fill perceived gaps, leading to a fragmented educational experience for students. As a result, students receive different levels of instruction, affecting their readiness for subsequent grades, standardized tests, and overall academic performance. Students moving between schools or classrooms face challenges adjusting to differing expectations and instructional approaches.

Inconsistent training and support for teachers further contribute to the uneven implementation of the curriculum. Some educators report not feeling fully prepared to deliver the Benchmark curriculum as intended, leading them to depend more on supplementary materials and personal adaptations.

This disparity in curriculum implementation can exacerbate educational inequities, particularly for students in under-resourced schools who may lack access to high-quality supplementary materials, widening the achievement gap between different student groups within the district.

Limited Access to Essential Materials

Teachers often face significant challenges in accessing essential materials needed to meet their students' diverse literacy levels, hindering effective instruction. This impacts the quality and consistency of literacy education, particularly in classrooms with diverse needs.

Many teachers report a shortage of essential literacy materials, such as decodable texts, phonics kits, and manipulatives. Budget limitations restrict the purchase of necessary resources, forcing teachers to use personal funds, which is unsustainable and inequitable. Outdated or insufficiently stocked school libraries and digital resource collections limit students' access to varied reading materials, crucial for fostering a love of reading and improving literacy skills.

Even when materials are available, insufficient training on their effective use further complicates the issue. Professional development focusing on integrating these resources is lacking or inconsistent.

The absence of essential materials leads to gaps in instruction, preventing full implementation of the literacy curriculum and necessary interventions for struggling readers, ultimately affecting student engagement and learning outcomes.

Need for More Comprehensive Professional Development

Professional development in many instances was infrequent and not sufficiently tailored to the specific needs of teachers, underscoring the need for more practical, classroom-focused PD sessions.

Many teachers reported that PD sessions were too infrequent to provide ongoing support and growth. Sporadic training sessions do not allow for the continuous development of skills and the reinforcement of new teaching strategies. Also, when professional development follows a one-size-fits-all approach, it does not address the diverse needs of teachers. Teachers working with different student populations, such as ELs or students with disabilities, require specialized training that is relevant to their unique classroom contexts Teachers also benefit more from hands-on, classroom-focused training that demonstrates how to implement strategies in real classroom settings. The lack of practical application makes it difficult for teachers to translate what they learn into effective instructional practices.

Furthermore, high-quality professional development must include follow-up support for teachers. Through mentoring, coaching, and with continuous feedback teachers refine and sustain these new practices. Without ongoing support, the impact of professional development is significantly diminished.

Recommendations

Formalize Curriculum Materials and Implementation

Develop a uniform literacy framework to reduce inconsistencies in curriculum resources. Provide comprehensive training with hands-on workshops and continuous support to ensure consistent adoption across all schools. Create inclusive and adaptable curriculum materials to cater to a wide range of literacy levels. Utilize literacy coaches to support small group instruction and targeted interventions. Re-evaluate the use of existing curricula like Benchmark and consider integrating additional resources like SIPPS and CRLP. Address training gaps, especially for teachers hired post-curriculum adoption or during the COVID pandemic, to ensure all educators are proficient and effective in implementing the curricula.

Enhance Teacher Support Systems

Increase the frequency and quality of professional development sessions and implement a structured peer coaching program at the district or site level. Tailor professional development to

address the unique challenges of different schools, modeling successful DEEP school workshops for non-DEEP schools. Focus on specialized training to support teachers in addressing the needs of vulnerable student groups, particularly those with disabilities and dyslexia, including strategies for differentiated instruction and effective technologies. Expand the availability of instructional coaches and mentors, especially in non-DEEP schools, to support the application of learned strategies in the classroom. Organize more collaborative opportunities for teachers to share best practices and successes through regular professional learning communities or a district-wide mentorship program.

Leverage Data for Instructional Improvement

Integrate regular, streamlined assessments to monitor student progress consistently. Provide teachers with comprehensive training in data literacy, focusing on how to interpret results and apply them to instructional strategies. Equip classrooms with digital tools for easy access and analysis of student data, supporting data visualization and integration with learning management systems. Promote a culture of data-driven instruction by organizing regular data review sessions within professional learning communities (PLCs), where teachers can collaborate and develop targeted intervention plans. Offer ongoing support from data coaches or specialists to help teachers effectively incorporate data into their instructional planning, ensuring data-driven instruction is a sustainable practice that enhances student outcomes.

Evaluation Next Steps

As we move into the second year of the evaluation of literacy practices within SDUSD, our focus will pivot towards a more granular analysis of comprehension strategies, the impact of embedded coaching compared to district-level coaching, and an exploration of additional supports provided by DEEP that could further explore variations in student achievement. Specifically, we aim to:

- Understand Comprehension Practices: Assess the implementation and effectiveness of routines and structures that support reading comprehension across grades. This will involve a close examination of how teachers integrate comprehension strategies into daily literacy instruction and the effect of these practices on student learning outcomes.

- Assess Coaching Models: Compare the effectiveness of embedded literacy coaches (those who work within specific schools) versus district-wide coaches. This analysis will help determine the coaching model that most effectively builds teacher capacity and improves student literacy outcomes.

- Expand Our Scope of DEEP Supports: Investigate other components of DEEP's support, such as parent involvement initiatives, community literacy programs, and integration of digital literacy, to see how these elements contribute to the overall literacy achievements of students.

Conclusions

In conclusion, the efforts to revamp literacy instruction within the San Diego Unified School District highlight both the complex challenges and promising strides made in addressing the diverse needs of its student population. This evaluation of curricular and instructional implementation, as well as their impact on literacy outcomes, provides insights crucial for guiding the district's ongoing literacy initiatives. Most importantly, the emerging success of the DEEP model in alleviating adverse effects on ELs highlights the effectiveness of a tailored, evidence-based literacy approach for multilingual students. School and district leaders should persist in investigating the reasons behind the varying performance of EL students in DEEP-affiliated schools. Furthermore, it is essential for school and district leaders to thoroughly address the ongoing challenges related to inconsistent curriculum execution, variations in teacher training, and the necessity for comprehensive professional development, which continue to pose substantial hurdles throughout the cluster.

These findings call for a sustained commitment to refine literacy strategies that are inclusive and adaptive to the needs of all students, particularly those from linguistically diverse backgrounds. Moreover, the evident role of systemic support structures, such as professional development and literacy coaching, in facilitating effective teaching practices emphasizes the need for continuous investment in these areas. As SDUSD moves forward, it is imperative to leverage the lessons learned from this initiative to enhance the scalability and sustainability of literacy programs that promise equitable educational outcomes.

Finally, we would be remiss not to identify the important limitations of this evaluation. Notably, our evaluation thus far has focused primarily on two aspects of literacy development: phonics and phonemic awareness. In the coming year, we will expand our scope to include other critical aspects of literacy, such as fluency, vocabulary, and comprehension. Additionally, while this evaluation has concentrated on a subset of DEEP's theory of change, future analyses will incorporate a broader range of influences, including extracurricular activities and other community and school partnerships, which also play a vital role in student literacy and language development.

References

Bialystok, E. (2018). Bilingual education for young children: Review of the effects and consequences. *International Journal of Bilingual Education and Bilingualism*, *21*(6), 666–679. https://doi.org/10.1080/13670050.2016.1203859

Bowers, J. S. (2020). Reconsidering the evidence that systematic phonics is more effective than alternative method of reading instruction. *Educational Psychology Review*, *32*, 681–705. https://doi.org/10.1007/s10648-019-09515-y

Bugler, M., D. (2017). *How teachers judge the quality of instructional materials.* WestEd. https://www.wested.org/wp-content/uploads/2017/03/resource-selecting-instructional-materials-brief-1-quality.pdf

Child Health, E. K. S. N. I. of, & Human Development, D., NIH. (2010). *Developing early literacy: Executive summary of the national early literacy panel.* National Institute for Literacy; U.S. Government Printing Office.

Council, N. R. (1998). *Preventing reading difficulties in young children* (C. E. Snow, M. S. Burns, & P. Griffin, Eds.). The National Academies Press. https://doi.org/10.17226/6023

Cummins, J. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire.* Multilingual Matters.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*(6), 934–945.

Ferguson, H., Bovaird, S., & Mueller, M. (2007). The impact of poverty on educational outcomes for children. *Paediatrics and Child Health*, *12*(8), 701–706. https://doi.org/10.1093/pch/12.8.701

Foorman, B. et al. (2016). *Foundational skills to support reading for understanding in kindergarten through 3rd grade.* National Center for Education Evaluation; Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education. http://whatworks.ed.gov

Garcia, O., & Lin, A. (2009). *Translanguaging in bilingual education.* https://doi.org/10.1007/978-3-319-02324-3_9-1

Goldenberg, C. (2020). Reading wars, reading science, and english learners. *Reading Research Quarterly*, *55*(S1), S131–S144. https://doi.org/10.1002/rrq.340

Great City Schools, C. of. (2023). *A framework for foundational literacy skills instruction for english learners: Instructional practice and materials considerations.* Council of the Great City Schools. https://creativecommons.org/licenses/by-nd/4.0/

Hart, B., & Risley, T. R. (2003). The early catastrophe: The 30 million word gap by age 3. *American Educator*, 4–9. http://www.aft.org//sites/default/files/periodicals/TheEarlyCatastrophe.pdf

Hernandez, D. J. (2011). *Double jeopardy: How third-grade reading skills and poverty influence high school graduation.* The Annie E. Casey Foundation.

Joyce, B. R., & Showers, B. (2002). *Student achievement through staff development* (3rd ed.). Association for Supervision; Curriculum Development (ASCD).

Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruc-

tion and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, *88*(4), 547–588. https://scholar.harvard.edu/files/mkraft/files/kraft_blazar_hogan_2016_teacher_coaching_meta-analysis_wp_w_appendix.pdf

La Paro, K. M., & Pianta, R. C. (2003). Classroom assessment scoring system (CLASS). *APA PsycTests.* https://doi.org/doi.org/10.1037/t08945-000

Leite, W. (2017). *Practical propensity score methods using r.* SAGE Publications, Inc.

Moats, L. C. (2020). *Teaching reading is rocket science: What expert teachers of reading should know and be able to do.* American Federation of Teachers.

Morgan, I., & Amerikaner, A. (2018). *Funding gaps 2018.* The Education Trust. https://edtrust.org/resource/funding-gaps-2018/

Neufeld, B., & Roper, D. (2003). *Coaching: A strategy for developing instructional capacity.* Education Matters, Inc. https://files.eric.ed.gov/fulltext/ED577101.pdf

Novicoff, S., & Dee, T. (2023). *The achievement effects of scaling early literacy reforms.* Annenberg Institute at Brown University. https://doi.org/10.26300/jnmt-2093

O'Day, J. (2009). Good instruction is good for everyone—or is it? English language learners in a BL approach. *Journal of Education for Students Placed at Risk (JESPAR)*, *14*(1), 97–119. https://doi.org/10.1080/10824660802715502

Panel, N. R., Child Health, N. I. of, & Development, H. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* U.S. Dept. of Health; Human Services, Public Health Service, National Institutes of Health, National Institute of Child Health; Human Development; National Institute of Child Health; Human Development, National Institutes of Health.

Petcher, C., Y., & Wagner, R. K. (2020). How the science of reading informs 21st-century education. *Reading Research Quarterly*, *0*(0), 1–16. https://doi.org/10.1002/rrq.352

Polanco, P., & Luft de Baker, D. (2018). Transitional bilingual education and two-way immersion programs: Comparison of reading outcomes for english learners in the united states. *Athens Journal of Education*, *5*(4), 423–444. https://doi.org/10.30958/aje.5-4-5

R Core Team. (2023). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/

Soto, M. (2023). SIPPS impact on english language learners reading success. *Voices of Reform: Educational Research to Inform and Reform*, *6*(2), 68–82. https://www.voicesofreform.com/article/91137-sipps-impact-on-english-language-learners-reading-success

Suggate, S. P. (2016). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *Journal of Learning Disabilities*, *49*(1), 77–96. https://doi.org/10.1177/0022219414528540

Tsokalidou, R., & Skourtou, E. (2020). *Translanguaging as a culturally sustaining pedagogical approach: Bi/multilingual educators' perspectives* (J. A. Panagiotopoulou, L. Rosen, & J. Strzykala, Eds.; pp. 219–235). Springer. https://doi.org/10.1007/978-3-658-28128-1_13

U. S. Department of Education, N. C. for E. S., Institute of Education Sciences. (2022). *2022 reading assessment.*

Westall, J., & Cummings, A. (2022). *The effects of early literacy policies on student achievement.* Education Policy Innovation Collaborative. http://dx.doi.org/10.2139/ssrn.

4427675

What exactly is the science of reading? (2024). Harvard Graduate School of Education. https: //www.gse.harvard.edu/ideas/usable-knowledge/24/06/what-exactly-science-reading

Xia, R. J., & Haas, B. W. (2024). The effect of bilingualism and multicultural experience on social-cognitive processing: A meta-analytic review. *Journal of Cultural Cognitive Science*, *8*(1), 47–64. https://doi.org/10.1007/s41809-023-00138-y