



Legislating Literacy: Estimating the Impact of State-Level Science of Reading Reforms on Reading Test Scores

Florence Bason

University of Florida

Christopher Redding

University of Florida

This study examines the extent to which state adoption of Science of Reading (SoR)-aligned policies was associated with changes in fourth-grade reading achievement. Using longitudinal state-level panel data from 2003-2024 and quasi-experimental difference-in-differences (DiD) methods, the study estimates relationships between SoR policy adoption and NAEP Grade 4 reading outcomes. Findings suggest that SoR-aligned reforms are associated with modest but gradually emerging improvements in reading achievement, with stronger effects observed at longer post-adoption periods. Effects are not uniformly distributed across student groups, with more consistent positive associations observed for White students and students not eligible for free or reduced-price lunch. Results further suggest that coordinated policy environments may matter more than individually adopted literacy policies.

VERSION: October 2026

Suggested citation: Bason, Florence, and Christopher Redding. (2026). Legislating Literacy: Estimating the Impact of State-Level Science of Reading Reforms on Reading Test Scores. (EdWorkingPaper: 26-1616). Retrieved from Annenberg Institute at Brown University: <https://doi.org/10.26300/dbgl-p217>

Legislating Literacy: Estimating the Impact of State-Level Science of Reading Reforms on Reading Test Scores

Florence Bason¹ and Christopher Redding²

¹University of Florida Literacy Institute, ² Human Development and Organizational Studies in Education, University of Florida

Author Note

I have no known conflict of interest to disclose.

Correspondence concerning this article should be addressed to Florence Bason, University of Florida Literacy Institute, University of Florida, Norman Hall, Room 0810, 618 SW 12 Street, PO Box 117050, Gainesville, FL, 32611. Email: socogirl@ufl.edu

Abstract

This study examines the extent to which state adoption of Science of Reading (SoR)-aligned policies was associated with changes in fourth-grade reading achievement. Using longitudinal state-level panel data from 2003-2024 and quasi-experimental difference-in-differences (DiD) methods, the study estimates relationships between SoR policy adoption and NAEP Grade 4 reading outcomes. Findings suggest that SoR-aligned reforms are associated with modest but gradually emerging improvements in reading achievement, with stronger effects observed at longer post-adoption periods. Effects are not uniformly distributed across student groups, with more consistent positive associations observed for White students and students not eligible for free or reduced-price lunch. Results further suggest that coordinated policy environments may matter more than individually adopted literacy policies.

Keywords: Science of Reading; literacy policy; reading achievement; state education policy; difference-in-differences, NAEP

Legislating Literacy: Estimating the Impact of State-Level Science of Reading Reforms on Reading Test Scores

When the National Assessment of Educational Progress (NAEP) released data in 2024, only 31% of fourth graders performed at or above the Proficient level in reading (NAEP, 2025). Average scores declined by two points compared to 2022 and by five points compared to 2019—both statistically significant decreases. That contemporary reading proficiency levels are not significantly different from 1992 when the NAEP reading assessment was first administered underscores long-term stagnation in national reading achievement (NAEP, 2025).

Differences in reading achievement also remain evident by race/ethnicity and free or reduced-price lunch (FRPL) eligibility. Average fourth-grade reading scores ranged from 199 for Black students to 237 for Asian students, while FRPL-eligible students scored 27 points lower on average than students who were not FRPL eligible (NAEP, 2025). Civil rights organizations have identified disparities in early literacy as a critical dimension of unequal educational opportunity, with possible implications for later outcomes including special education referrals, exclusionary discipline, and high school graduation (Carr, 2022; Hernandez, 2011).

Concerns over reading achievement levels and persistent disparities have renewed attention to questions about how reading should be taught. Although the importance of literacy is widely recognized, consensus regarding effective reading instruction has proven far more elusive. For decades, researchers, policymakers, and educators have engaged in what are often referred to as the “reading wars,” marked by disagreements between whole language and phonics-based approaches to literacy instruction (Pearson, 2004). These debates have contributed to competing instructional recommendations and persistent disagreements about how reading research should inform educational practice (Castles et al., 2018).

As these debates have evolved, the term “Science of Reading” (SoR) has gained widespread recognition, referring to evidence across psychology, linguistics, neuroscience, and education that identifies how reading develops and which instructional practices are most effective (Adams, 1990; Ehri, 2005; Kim & Snow, 2025; Shanahan, 2020). Once confined to scientific and professional circles, SoR has increasingly entered public discourse, with the *Sold a Story* podcast (Hanford, 2022) and Mississippi’s dramatic improvement in 4th grade reading scores (Burk, 2020; Spencer, 2024) emblematic of this shifting policy discourse.

SoR has also become a powerful organizing framework for state literacy legislation, emerging from Mississippi’s systemic approach to changing the instructional infrastructure around literacy (Cohen & Bhatt, 2012), including evidence-based curricula in teacher preparation programs (TPPs), changes to elementary teacher licensure requirements, professional development (PD) and instructional coaching, universal screening and assessment systems, targeted intervention for struggling readers, and educator accountability structures (Burk, 2020). While some states had enacted literacy policies prior to the COVID-19 pandemic, the majority of legislative activity around SoR occurred after 2020. As of 2025, 41 states and the District of Columbia (DC) have enacted SoR-aligned literacy laws aimed at improving reading outcomes.

Despite the rapid and widespread adoption of these policies, evidence on their effectiveness remains limited. Studies of Mississippi’s Literacy-Based Promotion Act (LBPA) have shown promising associations with improved reading scores (Burk, 2020; Spencer, 2024). Leveraging district-level variation in policy implementation, Novicoff and Dee (2025) evaluated California’s statewide SoR initiative, finding modest but significant improvements in early reading scores. This work suggests that comprehensive, evidence-based reforms can improve

reading outcomes; yet, because they focus on individual states, the question remains about whether similar effects would emerge across the broader landscape of SoR policy adoption.

This question is particularly important because evidence from other large-scale literacy reform efforts has been less conclusive. Examining state literacy policies such as literacy screening, intervention requirements, PD, and third-grade retention provisions, Westall and Cummings (2023) found that comprehensive policies, particularly those including third-grade retention requirements, were associated with improvements in high-stakes reading test scores, though effects on low-stakes assessments like NAEP were more limited. Similarly, the Reading First initiative, which provided funding to states to encourage the use of evidence-based reading instruction, was shown to increase time spent on phonemic awareness, phonics, vocabulary, fluency, and comprehension, but showed no evidence of improved student reading comprehension in early elementary school and only evidence on first graders' decoding skills in a single year of the study. This more muted evidence is more consistent with a model of policy effects in which policies are interpreted and enacted locally, with implementation shaped by organizational context, instructional capacity, educators' existing knowledge and beliefs, and the supports available to translate policy into classroom practice (Coburn, 2001, 2004).

Despite mixed success of these previous efforts, there are conceptual reasons to expect that SoR policies may operate differently. Rather than targeting a single aspect of reading instruction, state-level SoR policies often combine multiple policy levers, including teaching preparation, teacher certification, PD, instructional coaching, progress monitoring, curriculum, and intervention systems (Burk, 2020). When these policy levers are aligned, they have the potential to create an educational infrastructure that supports changes in instructional practice

(Woulfin & Gabriel, 2020). This perspective is consistent with instructional coherence frameworks emphasizing that reforms are most likely to influence classroom practice when curriculum, assessment systems, professional learning, and instructional expectations reinforce a shared vision of teaching and learning (Coburn et al., 2016; Cohen & Bhatt, 2012; Newmann et al., 2001; Woulfin, 2020).

The purpose of this study is to estimate the effects of state-level SoR-aligned policies on student achievement on the NAEP using a difference-in-differences (DiD) approach, both overall and disaggregated by race/ethnicity and free and reduced-price lunch (FRPL) eligibility. For the purposes of this analysis, SoR-aligned policies are defined as state legislation or regulations that require or promote instruction, curricula, assessments, or teacher preparation grounded in evidence-based reading practices consistent with the five essential components of reading: phonemic awareness, phonics, fluency, vocabulary, and comprehension. To capture the complexity of these reforms, SoR-aligned policies are further categorized into six key areas: (1) teacher preparation, (2) teacher certification or license renewal, (3) PD and coaching, (4) assessment, (5) instructional materials, and (6) instruction and intervention. This study addresses the following research questions:

- 1) To what extent do state-level SoR-aligned policies impact fourth-grade NAEP reading achievement?
- 2) To what extent does this overall relationship vary by race/ethnicity and FRPL status?
- 3) Do specific categories of SoR-aligned policies have differential effects on fourth grade reading outcomes?

As states translate this research base into legislation, SoR has increasingly become a policy movement – one that carries its own interpretations, priorities, and implementation

challenges (McLean et al., 2026). For policy researchers, this distinction is critical: what is enacted under the banner of SoR may vary substantially in content and fidelity. Rigorous empirical study is therefore needed to evaluate whether these laws, as implemented, are improving literacy outcomes and reducing longstanding inequities in reading achievement. The literature review below situates contemporary SoR policy within both the research on reading instruction and the broader literature on educational policy implementation.

The Science of Reading: From Research to Policy Implementation

Decades of research consistently demonstrate the effectiveness of explicit, systematic reading instruction. The National Reading Panel (2000) reviewed 38 studies yielding 66 treatment-control comparisons and found that systematic phonics instruction produced significantly greater gains than instruction that included unsystematic or no phonics, with effects varying across reading outcomes and student characteristics (Ehri et al., 2001).

The translation of SoR research into policy has not been without controversy, particularly as states have increasingly relied on legislative mandates to influence instructional practice. Questions remain about how successfully the SoR-aligned reforms can be translated into practice in ways that change teacher practice and improve student outcomes. Scholars have noted that top-down implementation efforts often produce superficial compliance or outright rejection rather than meaningful instructional practice (Supovitz & Weinbaum, 2010). For instance, Coburn's (2004) early work identified a range of teacher responses, including assimilating new messages about reading with their own beliefs, creating parallel structures reflecting different instructional philosophies, and responding symbolically. Rejection of policy messages about reading instruction was also common. Others have come to a similar conclusion, pointing to the ways in which scripted, phonics-based reading programs can challenge teachers' conceptions of

themselves as skilled professionals who adapt instruction to meet individual student needs (Narayanan et al., 2024; Powell et al., 2017; Valli & Buese, 2007). Thus, policy efforts run up against deeply held beliefs about teacher autonomy (Ingersoll, 2009), which dovetail with teachers' own professional identity and views about how reading should be taught.

Taking a more system-level approach, some scholars have noted “chronic incoherence in U.S. public education” as limiting instructional reforms when teachers face conflicting messages across preparation programs, standards, curricula, textbooks, assessments, observation protocols, school and district leadership priorities, PD, and professional learning networks, among others (Cohen & Bhatt, 2012, p. 123). Though there has been longstanding optimism around instructional coherence as a way to foster policy systems that promote high-quality instruction (e.g., Cohen & Bhatt, 2012; Moon et al., 2022; Newmann et al., 2001; Spillane et al., 2002), achieving this aim has been elusive (Coburn, 2001, 2004).

In addition to broader concerns about how literacy reforms are translated into classroom practice, critics have questioned whether the SoR's often standardized, scripted curricula conflict with culturally sustaining pedagogies that value the linguistic and literacy practices students bring from home (Elzy-Palmer et al., 2025). Consider how a teacher responds when a Spanish-speaking student reads *ship* as *sheep*. A culturally responsive teacher might recognize that the student is applying Spanish phonemic rules, affirm that knowledge, and explain how the vowel sound differs in English. In contrast, a less culturally affirming teacher might simply correct the error and move on without validating and bridging why they had read *ship* in the way they had. This example helps illustrate that even if SoR-aligned curricula are adopted, the extent to which they are implemented in a culturally responsive manner may have a bearing on whether they affect the achievement of students whose home literacy practices and cultural knowledge have

been traditionally marginalized in school settings. Amidst these various challenges with translating literacy reforms into practice, in the next section, we describe Mississippi as a proof-of-concept for how a racially diverse state that historically underperformed on the NAEP initiated comprehensive SoR-aligned reform.

Mississippi's Literacy-Based Promotion Act

Historically among the nation's lowest-performing states, Mississippi entered the 2010s amid longstanding concerns about low reading achievement and educational inequities. In 2011, 78% of the state's fourth-grade students scored below proficient on NAEP, contributing to a growing sense of urgency among policymakers (Burk, 2020). In response, a coalition of state leaders and stakeholders pursued a comprehensive approach to early literacy reform, culminating in the passage of the LBPA in 2013. The law established multiple coordinated interventions designed to strengthen foundational reading instruction, improve teachers' knowledge of SoR, and ensure that students were reading on grade level by the end of third grade (Burk, 2020). Although the law included grade retention provisions for students not meeting this bar, its primary focus was preventive, emphasizing early identification and support, with benchmarking and progress monitoring beginning in kindergarten and instructional adjustments and interventions when students are not on track (Burk, 2020).

Another component of the LBPA provides support for teacher professional learning and teacher training in foundational literacy. The state provided in-service training, such as Language Essentials for Teachers of Reading and Spelling (LETRS) (Scorrano, 2021). They also hired and deployed literacy coaches, prioritizing schools performing in the bottom quarter of the state reading assessment (Burk, 2020). As Mississippi implemented the LBPA, state leaders also identified gaps in preservice teachers' preparation. In response, Mississippi updated licensing

requirements to assess preservice elementary teachers' knowledge of the science of learning to read, further aligning teacher preparation, licensure, and classroom expectations (Burk, 2020).

In the wake of these reforms, average grade 4 NAEP reading scores increased from 209 in 2013 to 219 in 2019, improving its national ranking from 49th to 29th. By 2022, Mississippi outperformed the national average and had a lower percentage of students at or below NAEP basic. These improvements drew significant attention from policymakers, educators, and researchers and prompted both celebration and scrutiny regarding replicability and equity. In response, a range of studies examined the LBPA's implementation and outcomes, from descriptive accounts to quasi-experimental analyses (Burk, 2020; Spencer, 2024). At present, the strongest evidence comes from Spencer's (2024) synthetic difference-in-differences methodology to estimate the LBPA's causal impact. Compared to a synthetic control group, the author found a statistically significant gain of 9.8 NAEP points in Grade 4 reading (approximately 0.25-0.3 standard deviations). These findings supported the notion that comprehensive, coordinated reform can produce meaningful improvements, though it should be noted that Spencer's (2026) follow-up work found no effect on longer term outcomes (e.g., ACT scores; AP test participation; Spencer, 2026).

The Spread of State Literacy Reforms

As Mississippi's gains drew national attention, states pursued their own SoR-aligned packages. By 2025, another 40 states adopted SoR-aligned literacy laws. While these laws share a similar orientation, they differ substantially in scope, implementation timelines, accountability mechanisms, and the policy levers used to influence local literacy practices, including teacher preparation and certification requirements; others emphasize PD, curriculum adoption, or assessment systems. A smaller subset integrates all these elements into comprehensive, system-

wide reforms. Because education governance is decentralized in the U.S., the same policy label (i.e., SoR) may entail different expectations, supports, and enforcement across states.

For example, Mississippi and Arkansas enacted comprehensive literacy laws nearly a decade ago, giving districts time to phase in training, materials, and assessment systems. Later adopters, such as California and Oregon, passed legislation more recently, leaving implementation in the early stages. Some states, like Colorado, focus primarily on teacher preparation and professional learning, while others, such as North Carolina and Tennessee, link literacy reforms to statewide curriculum and assessment mandates. This variation provides an opportunity to study not only whether SoR-aligned laws affect achievement, but also which policy levers are most influential. To capture this landscape of SoR policies, Appendix Table A1 summarizes SoR-aligned legislation passed across states, while Table 1 reports the timing and key components of these policies.

Conceptual Framework

Our conceptual framework draws on policy implementation theory and instructional coherence to explain whether and how SoR reforms may affect student achievement. Policy implementation theory explains how policies move from policy adoption to enactment in schools and classrooms (Honig, 2006; McLaughlin, 1987). Even well-designed policies can falter if the supports, incentives, and capacities needed for implementation are absent. This perspective is critical for SoR reforms: passing a literacy bill does not guarantee improved instruction; impact depends on how policies shape the professional knowledge, instructional materials, assessments, and intervention systems that teachers encounter (Woulfin & Gabriel, 2020). Implementation theory anchors this study's conceptual framework: policies influence outcomes only to the extent they are effectively implemented and supported across organizational levels. This helps explain

why similar reforms can produce different effects across states and underscores the importance of examining variation in policy components, scope, and strength.

While implementation theory explains why policy effects vary, instructional coherence helps specify how multiple policy levers operate together. Instructional coherence refers to the degree to which the elements of an instructional system are mutually reinforcing around a shared vision of teaching and learning (Newmann et al., 2001). Coburn et al. (2016) emphasize that coherence must be built vertically across policy, district, school, and classroom levels, noting that reforms often fail due to misalignment among the tools and messages educators receive.

Applied to SoR policy, instructional coherence suggests that policy components function not in isolation but as an interdependent system. Woulfin and Gabriel (2020) conceptualize reading improvement as dependent on an interconnected instructional infrastructure that includes curriculum, PD, and school and district leadership as mutually supportive pillars of reform. From this perspective, evidence-aligned practices or materials alone are insufficient if the organizational conditions surrounding them are disconnected. Effective reading reform depends not only on the quality of individual supports but also whether they are coherent, transparent to educators, and feasible to sustain at scale. This framework is particularly relevant to state SoR legislation, which often included a range of policies related to TPP reform, PD and coaching, teacher certification and licensure, curriculum and instructional materials, assessment systems, and interventions and support systems (Schwartz, 2024).

A major focus of SoR legislation has been on teacher knowledge and pedagogy related to early literacy. As shown in Table 1, 26 states have adopted SoR reforms targeted towards TPPs to strengthen beginning teachers' knowledge of evidence-based reading instruction. SoR legislation has thus increased oversight of how TPPs prepare candidates to teach reading.

However, this effect may be muted given that any impact would be concentrated among beginning teachers and that effects may be muted given that TPPs vary considerably in the extent to which their coursework addresses the research base on reading development and instruction (Moats, 2020; National Council on Teacher Quality, 2023).

For educators already in the classroom, state policies aimed at PD and coaching serve as the primary mechanism for improving instructional practice. A strong evidence base supports these practices (Desimone, 2009). Basma and Savage's (2018) systematic review and meta-analysis of the impact of teacher PD on student reading achievement found a 0.225 standardized effect. Kraft et al.'s (2018) meta-analysis of 60 causal studies on instructional coaching found a 0.18 standard deviation impact on student achievement, with larger effects for content-specific (particularly reading-focused) coaching models. In some states, PD and coaching were supplemented with changes in teacher licensure renewal policies. Together, these policies sought to increase teacher knowledge and capacity related to literacy instruction.

Other policies developed system-level supports for teachers, including curriculum and instructional materials, assessment systems, and intervention systems. Curriculum is one of the most direct ways states and districts influence daily classroom instruction (Chingos & Whitehurst, 2012). Yet, evidence suggests that adopting higher quality materials alone is often insufficient to improve student outcomes (Blazar et al., 2020). An early What Works Clearinghouse review and more recent evaluations found limited evidence for many widely used reading curricula and substantial variation in the quality and usability of instructional materials (EdReports, 2024; WWC, 2007). Even when districts adopt materials rated as high-quality, teachers often use only select components due to time constraints, misalignment with assessments, or limited training (Doan et al., 2022). Consistent with Woulfin and Gabriel (2020),

these findings suggest that curriculum is only one pillar of a broader instructional infrastructure whose effectiveness depends on connections to teacher professional learning.

System-level supports adopted as part of SoR legislation include assessment and intervention systems designed to identify reading difficulties early and connect students with appropriate support. Although screening, progress monitoring, and multi-tiered systems of support (MTSS) predate the contemporary SoR policy movement, they provide a mechanism for translating evidence about reading development into instructional decisions. Accordingly, many states embed assessments within an MTSS framework that relies on early screening and continuous progress monitoring to guide instruction and intervention, particularly for students at risk of reading difficulties such as dyslexia. These systems range from formative assessments that inform classroom instruction (i.e., Tier I) to targeted, evidence-based support (i.e., Tier II and III; Wanzek & Vaughn, 2007). Their inclusion in SoR policies reflects evidence that early, intensive interventions can improve foundational reading skills for struggling readers (Gersten et al., 2020; WWC, 2022).

Whether any benefits of SoR legislation have disparate effects by race/ethnicity and FRPL is less clear. Universal screening enables early identification of reading difficulties, potentially preventing the compounding gaps that often affect low-income students and students of color. When materials, training, and coaching are aligned, schools reduce variability in core instruction, which might benefit low-income students and students of color who are most often assigned to less effective teachers (Clotfelter et al., 2005; Goldhaber et al., 2015). Moreover, some states have prioritized lower performing schools when rolling out supports (e.g., coaching), which may disproportionately benefit low-income students and students of color.

That said, the rigidity of SoR policies may limit teachers' ability to adapt instructional resources for diverse learners. SoR policies also must contend with other systemic factors. For instance, additional training and support may prove fruitless amidst broader teacher labor market patterns that not only contribute to the inequitable distribution of high-quality teachers and high turnover (Nguyen & Redding, 2026). High turnover might undermine SoR implementation by fragmenting the collaborative structures and institutional knowledge needed to maintain instructional coherence (Ronfeldt et al., 2013). As a result, schools serving large concentrations of low-income students and students of color may face greater implementation challenges, potentially limiting the extent to which SoR reforms reduce educational inequality.

Data and Method

This study uses several data sources combined into a balanced state-year panel from 2003 to 2024: (1) NAEP: State-level scale scores for grade 4 reading and math, both overall and disaggregated by race/ethnicity and FRPL, (2) SoR-aligned policy adoption, and (3) state-level public school enrollment from the Common Core of Data (CCD), used to construct enrollment weights. NAEP data are particularly valuable for the purposes of this study because they provide a consistent, comparable measure of reading performance across states, independent of state standards and accountability systems (NCES, 2024). However, because NAEP is administered biennially in most years (i.e., 2003, 2005, 2007, 2009, 2011, 2013, 2015, 2017, 2019, 2022, and 2024), we constructed an annual state-year panel by linearly interpolating state-level means in non-assessment years.

Dependent Variables

The primary outcome in this study is NAEP Grade 4 reading achievement, measured using state average scale scores. NAEP reading scores are estimated using item response theory

(IRT), which combines student responses across multiple test forms into a common developmental scale that permits valid comparisons across states and over time. Because the outcome is reported on the NAEP scale, treatment effects are interpreted in NAEP points. Across the study period, the standard deviation of state mean reading scores was approximately 6.5 NAEP points, providing a benchmark for interpreting effect sizes. NAEP reading assesses students' ability to comprehend literacy and informational texts and includes skills related to vocabulary, interpretation, and reading comprehension. Although NAEP is not designed to directly assess foundational reading skills such as phonics, these skills provide the foundation for skilled reading and are expected to contribute to later reading comprehension. Consequently, improvements in foundational reading resulting from SoR-aligned reforms should ultimately be reflected in broader measures of reading achievement such as NAEP.

To examine heterogeneous policy effects (RQ2), separate treatment effects will be estimated for racial/ethnic subpopulations and by eligibility for the National School Lunch Program (NSLP). Racial subgroups include White, Black, Hispanic, Asian, American Indian/Alaska Native, and students identifying with two or more races.

Defining SoR-Aligned Policies

To determine whether each state had adopted SoR-aligned legislation, we first reviewed the Schwartz (2024) *Education Week* article summarizing state literacy laws and policies, followed by The Reading League Compass (2025). Using these resources as a springboard, we reviewed each state's legislation and policy documents, noting the year, bill title, and relevant provisions (see Appendix Table A1). Each law was coded according to the definition of SoR-aligned policy (i.e., policies requiring or promoting instruction, curricula, assessments, or teacher

preparation grounded in the five essential components of reading: phonemic awareness, phonics, fluency, vocabulary, and comprehension), described in greater detail in Appendix Table A2.

Most states were clearly aligned with this definition and were therefore straightforward to code as SoR-aligned (e.g., Colorado, Florida, and West Virginia). States with no legislation or only dyslexia-screening requirements were coded as not SoR-aligned (e.g., Maine, Montana, and New York). Some states required additional consideration. For example, Idaho's *Literacy Achievement and Accountability Act (2021)* mandates evidence-based reading instruction, assessment, and educator preparation but does not explicitly reference SoR, structured literacy, or the five essential components of reading. Similarly, Nebraska's *Reading Improvement Act (2018)*, Nevada's *Read by Grade 3 Act (2015, revised 2019)*, New Jersey's *Student Literacy Act (2024)*, and Oregon's *Early Literacy Success Initiative (2023)* emphasize foundational or evidence-based literacy instruction and professional learning but stop short of requiring explicit, systematic instruction in phonemic awareness, phonics, fluency, vocabulary, and comprehension. These states were coded as SoR-aligned because their legislation promotes evidence-based early literacy instruction consistent with SoR principles. In sensitivity analyses described below, we exclude these borderline cases.

A potential limitation of our treatment coding is that implementation timelines varied across states. Many reforms were phased in over multiple years, particularly in areas such as teacher preparation, curriculum adoption, and PD requirements. As an example, Missouri incrementally adopted reforms from 2018 to 2023, with later legislation adding approved instructional materials and SoR-focused PD. In such cases, we coded the first year of relevant legislation as the treatment year even when the full scope of SoR policies was adopted later.

For our main time-varying treatment variable, states were coded using the first year in which SoR-aligned legislation was adopted, with 0 coded as not having adopted a policy in that year. Table 1 summarizes enactment years, implementation periods, and final treatment classifications for each state included in the analysis. We supplement this binary treatment variable with a dynamic treatment variable. This variable is constructed by calculating the difference between each observation year and the state's SoR policy adoption year ($k = \text{Observation year} - \text{Adoption Year}$). The resulting variable represents years relative to adoption, with negative values indicating pre-adoption, $k = 0$ indicating the adoption year, and positive values indicating years since initial SoR policy adoption. For event-time analyses, the year immediately preceding adoption ($k = -1$) served as the reference period. In TWFE models, event times beyond ± 6 years were grouped into endpoint bins to improve precision in sparsely populated leads and lags. For models using Callaway and Sant'Anna's estimator, dynamic treatment effects were reported over an event-time window of -6 to +6 years to focus inference on periods with sufficient cohort support.

Empirical Analysis

Because SoR policies were adopted at different times across states, this study employs a staggered DiD design. A traditional TWFE estimator is used as a benchmark and is commonly used to evaluate policy reforms rolled out over multiple years; however, recent econometric research has shown that when treatment is adopted at different times across units – and especially when treatment effects vary over time – TWFE estimators can be sensitive to treatment effect heterogeneity under staggered adoption (Borusyak et al., 2021; Goodman-Bacon, 2021). We supplement the canonical TWFE estimator with more recent estimators (i.e., Callaway & Sant'Anna, 2021; Sun & Abraham, 2021; Wing et al., 2024).

The baseline DiD model is estimated using the following TWFE specification:

$$Y_{st} = \alpha + \beta_1 SoR_{st} + \gamma_s + \delta_t + e_{st} \quad (1)$$

where Y_{st} is the fourth grade NAEP reading score in state s in year t , SoR_{st} is an indicator equal to 1 in years at or after the state's first SoR-aligned year and 0 otherwise, γ_s represents state fixed effects, δ_t represents year fixed effects, and e_{st} is an error term. β_1 , the coefficient of interest in this model, can be interpreted as the average treatment effect of SoR-aligned policy adoption. Standard errors are clustered at the state level. This and other models are weighted by state-year public school enrollment so estimates can be interpreted as population-average effects.

To address potential bias from staggered treatment adoption, we estimate group-time average treatment effects using the Callaway and Sant'Anna (CS; 2021) and dynamic responses using Sun and Abraham (SA; 2021). These estimators restrict comparisons to untreated states (either not-yet-treated – the primary comparison set – or never treated), thereby avoiding comparisons to already-treated states that can bias TWFE estimates under staggered adoption. These estimators avoid the bad comparison problem by ensuring treated states are compared only to states that are untreated at time t (not-yet-treated) or never treated. As a final specification, we also employ a stacked DiD design, in which each treated unit defines a cohort based on its adoption year (Wing et al., 2024). For each cohort, the treated state is compared only to “clean controls,” defined here as states that are never treated or that adopt SoR only after the cohort's full post-treatment window, ensuring that control states remain untreated throughout the estimation period. This design avoids comparisons between already-treated and newly treated units that can bias traditional TWFE estimators in staggered adoption settings (Wing et al., 2024).

All the DiD estimators we used in this study rely on the assumption that, absent policy adoption, treated and comparison states would have followed similar trends in reading achievement. Although this assumption cannot be tested directly, we evaluate its plausibility using event-study specifications that estimate pre-treatment coefficients relative to the year immediately preceding adoption. Under the parallel trends assumption, these coefficients should be statistically indistinguishable from zero. We report event-study estimates and joint tests of pre-treatment coefficients for each specification. Additional threats to identification include the possibility of contemporaneous policy changes that coincide with SoR adoption and differential trends across states. To mitigate these concerns, all models include state and year fixed effects, which account for time-invariant differences between states and common shocks affecting all states in a given year. As an additional falsification test, we replicate the primary analyses using fourth-grade NAEP mathematics as the outcome. Because SoR policies are specifically designed to influence reading instruction, evidence of comparable effects on mathematics achievement would suggest that estimated reading effects may reflect broader policy or contextual changes rather than a specific influence of SoR-aligned reforms.

To assess the dynamic effects of SoR-aligned legislation and to formally test for parallel trends, we estimate the following model:

$$Y_{st} = \alpha + \sum_{\ell \neq -1} \beta_{\ell} D_{st}^{\ell} + \gamma_s + \delta_t + \varepsilon_{st} \quad (2)$$

where Y_{st} is the fourth-grade NAEP reading scores in state s and year t , D_{st}^{ℓ} is an indicator equal to 1 when state s is ℓ years relative to policy adoption in year t and 0 otherwise, γ_s represents state fixed effects, δ_t represents year fixed effects, and ε_{st} is an error term. The year immediately preceding policy ($\ell = -1$) serves as the omitted reference period. The coefficients β_{ℓ} trace the evolution of treatment effects before and after adoption, with coefficients for $\ell < -1$ providing

a test of the parallel trends assumption and coefficients for $\ell \geq 0$ capturing the dynamic effects of SoR-aligned legislation following adoption.

To answer the second research question regarding differences in the effect of SoR-aligned legislation by race/ethnicity and FRPL status, we re-estimate Equations 1 and 2 using the state-level scores for each subgroup.

Because SoR legislation varies substantially in both structure and comprehensiveness, additional models will probe heterogeneity in policy scope (Research Question 3). Four complementary approaches are used. First, we explore whether specific aspects of SoR legislation are more or less strongly associated with improved reading achievement. Specifically, we estimate separate staggered models for each component of SoR legislation that we coded—teacher preparation, certification/licensure, PD, assessment, instructional materials, intervention—to identify the degree to which these different policy levers are more or less strongly associated with observed changes in reading achievement. Because multiple components were often co-adopted, these models will not include other components simultaneously, avoiding multicollinearity and interpretation challenges.

We then conduct a series of exploratory analyses to understand the extent to which the scope of SoR policy adoption, different bundles of policies, or policy strength are differentially associated with student reading achievement. Our “Policy Scope” analysis substitutes our main binary treatment variable with a categorical variable indicating the maximum number of SoR-aligned policy components adopted over the study period (0; 1 – 2; 3 – 4; 5 – 6). Treatment effects are then estimated separately by policy scope. Because scope is defined based on the full observation period, these estimates reflect differences across states with varying overall policy intensity rather than year-specific policy exposure. This analysis provides our first step in

understanding whether the number of policies matter, though we urge caution when interpreting these results as components are not necessarily additive.

To avoid conflating dissimilar policies, our next exploratory analysis groups policy components will also be grouped into theoretically motivated policy bundles: (a) teacher knowledge supports (teacher preparation, certification/licensure, and PD), (b) system supports (instructional materials and assessment), and (c) comprehensive multi-component reforms (five or more components). These bundles are restricted to states whose reform profiles align with predefined configurations. Nine treated states (Delaware, Iowa, Louisiana, Nebraska, Nevada, New Jersey, North Dakota, South Carolina, and Texas) did not align with these configurations and were excluded from the bundle analysis.

Last, we used principal component analysis (PCA) to generate a data-driven measure of policy strength. The PCA was conducted on six indicators capturing whether each state adopted teacher preparation, certification/licensure, PD, assessment, instructional materials, and instruction/intervention policies. The first principal component was retained as a summary measure of overall reform intensity and explains 57.0% of the variance in policy adoption patterns. For interpretability, treated states were grouped into low-, medium-, and high-strength categories based on the distribution of the PCA score (bottom, middle, and top thirds).

Results

Event Study Results

Figure 1 presents the results from event-study models from our four estimators, with estimates of changes in NAEP Grade 4 reading achievement relative to the timing of state's adoption of SoR-aligned legislation. Pre-trends are stable or slightly negative in the case of the SA and stacked DiD models. Pre-treatment estimates from our preferred model applying the CS

estimator were generally small and centered around zero, providing support for the parallel trends assumption. Joint Wald tests failed to reject the null hypothesis that the pre-treatment coefficients from event times -6 through -1 were jointly equal to zero ($W = 7.58, p = .27$). Thus, there was no statistical evidence of systematic differences in reading achievement trends between adopting and comparison states prior to policy adoption.

Post-adoption estimates indicate that changes in reading achievement emerged gradually rather than immediately following policy implementation, though only significant at the 0.05 level for 4-5 years after adoption in the CS and SA estimators, and year 6 for the SA estimator. In the CS model, estimated effects reached 2.82 NAEP points four years after adoption ($SE = 0.85, p < .01$) and 2.68 points five years after adoption ($SE = 0.74, p < .01$). Relative to the state-level standard deviation in Grade 4 NAEP reading scores in the analytic sample (6.52 points), these estimates correspond to approximately 0.43 and 0.41 standard deviations, respectively. At six years after adoption, the estimated effects increased to 3.60 points ($SE = 1.86$), or approximately 0.55 state-level standard deviations, although the estimate was slightly less precise ($p = .052$). Post-adoption estimates from the stacked DiD and TWFE models are consistently positive but estimated with imprecision and do not increase 4-5 years after adoption.

Together the event-study results provide little evidence of immediate changes in statewide reading achievement following adoption of SoR-aligned legislation but suggest that positive effects may emerge after several years of policy exposure. Results from alternative staggered adoption estimators produced a broadly similar pattern of delayed effects.

Overall Results and Sensitivity Analyses

Table 2 presents estimates of the association between SoR-aligned legislation and fourth-grade reading achievement across alternative DiD estimators. The preferred CS specification

produced a small, but statistically insignificant average treatment effect of 0.33 NAEP points ($SE = 0.35$), equivalent to approximately 0.05 state-level standard deviations. Estimates were larger under the alternative staggered-adoption estimators. The SA estimate was 1.04 points ($SE = 0.20$, $p < .01$), while the stacked DiD estimator was 2.35 points ($SE = 0.90$, $p < .01$). The conventional TWFE estimator was 0.82 points ($SE = 0.72$) and was not significant.

Although the magnitude and precision of the estimates varied across estimators, all four estimates were positive. Together with the dynamic estimates in Figure 1, these results provide some evidence of improved reading achievement following adoption of SoR-aligned legislation, but the size of the estimated association is sensitive to the estimator used.

The remaining columns of Table 2 show that the overall relationship between SoR-aligned legislation and reading achievement was consistently positive but not statistically significant using alternative specifications of the CS estimator. The unweighted CS estimate was 0.50 points, compared with 0.33 points in the preferred enrollment-weighted specification. Using never-treated states rather than not-yet-treated states as the comparison group produced a similarly modest estimate of 0.60 points. Estimates were also similar when using the full panel (0.34), excluding states with borderline policy classifications (0.30), and restricting the analysis to observed NAEP years without linear interpolation (0.43), although none of these estimates was statistically distinguishable from zero.

Calendar-time estimates provide additional context for the potential influence of the COVID-19 pandemic. The estimated effect was 1.59 points ($SE = 0.32$) in the pre-COVID period and 0.52 points ($SE = 0.49$) in the post-COVID period (2023 – 2024).¹ Although the relatively short post-COVID period limits comparisons across periods, the positive estimates observed

¹ We do not include the 2020-2022 school years for this estimation window given that they depend completely on interpolated data.

before and during the pandemic suggest that the overall pattern cannot be attributed solely to pandemic-era changes in achievement.

As a falsification test, we re-estimated the primary models using fourth-grade NAEP mathematics scores, an outcome not directly targeted by SoR-aligned legislation. Appendix Table A3 shows that the estimated effects were small, negative, and statistically indistinguishable from zero using both not-yet-treated ($ATT = -0.77$) and never-treated comparison states ($ATT = -0.43$). Pre-treatment Wald tests likewise provided no evidence of differential trends, and the event-study estimates remained statistically indistinguishable from zero throughout the post-treatment period. The absence of a corresponding pattern in mathematics provides some evidence that the reading results do not simply reflect broader changes in state academic achievement.

SoR-Aligned Legislation and Subgroup Differences in Reading Test Scores

In the next set of analyses, we disaggregate treatment effects by FRPL eligibility (Figure 2A) and race/ethnicity (Figure 2B), with overall treatment effects reported in Table 3. Figure 2A shows relatively little evidence of changes in reading achievement among FRPL-eligible students following policy adoption. Post-adoption estimates were generally positive but imprecise, with 95% confidence intervals spanning zero throughout the event-study period. In contrast, estimates for students not eligible for FRPL became increasingly positive with longer exposure to SoR-aligned legislation, particularly in later post-adoption periods.

The overall treatment effects reported in Table 3 are consistent with these dynamic patterns. For FRPL-eligible students, the estimated average treatment effect was 1.00 NAEP points and was not statistically significant. For students not eligible for FRPL, the estimated effect was larger at 2.37 points ($p < 0.05$). Although estimated effects were positive for both

groups, the estimate was larger for students not eligible for FRPL and statistically distinguishable from zero. Importantly, however, the difference between subgroup-specific estimates was not statistically significant; therefore, these results do not provide evidence that the effects of SoR-aligned legislation differed between FRPL-eligible and non-FRPL students.

When disaggregating by race/ethnicity, estimates for Black and Hispanic students remained relatively close to zero across most post-adoption years. For White students, estimates were modest during the initial post-adoption period but became larger and positive in later years. Estimates for American Indian/Alaska Native, Asian/Pacific Islander, and multiracial students were substantially less precise, as reflected in the wider confidence intervals. These patterns warrant cautious interpretation, particularly for groups represented by fewer state-year observations. Average treatment effects in columns 3-8 of Table 3 were positive across all six groups but varied considerably in magnitude and precision. The estimated effect was statistically significant only for White students ($ATT = 2.06, p < .05$). Estimates for Black ($ATT = 0.99$), Hispanic ($ATT = 1.02$), Asian/Pacific Islander ($ATT = 0.63$), American Indian/Alaska Native ($ATT = 3.91$), and students identifying with two or more races ($ATT = 1.49$) were positive but not statistically distinguishable from zero.

Together the subgroup analyses provide more consistent evidence of positive effects among students not eligible for FRPL and White students than among their respective comparison groups. However, confidence intervals overlapped substantially across subgroups, indicating no statistical evidence of differences across subgroups. Accordingly, these results are best interpreted as evidence of larger effects for White and FRPL students, but not necessarily significant from students of color and non-FRPL students, respectively.

Variation in the Effect of SoR Legislation by Policy Features

The preceding analyses estimate the average association between adoption of SoR-aligned legislation and reading achievement, but states varied considerably in the content and scope of their reforms. We therefore examined whether effects differed by policy components, the number of components adopted, policy bundles, and policy strength. For these analyses, never-treated states serve as the comparison group because the goal is to distinguish among policy configurations. Using not-yet-treated states could confound differences in treatment timing, where never-treated states provide a common benchmark across policy configurations.

Policy Components

Panel A of Table 4 examines each of the six policy components separately. Estimated average treatment effects were positive for five of the six components, although only certification/licensure and PD were statistically significant. Adoption of certification/licensure policies was associated with a 1.43-point increase in fourth grade reading scores ($p = .04$), while PD policies were associated with a 1.66-point increase ($p = .01$). Estimates for teacher preparation (ATT = 0.75), instructional materials (ATT = 0.28) and instruction/intervention (ATT = 0.55) were positive but not statistically significant. The estimate for assessment was negative but likewise not statistically significant (ATT = -0.57).

The dynamic estimates provide additional evidence that effects associated with some components may emerge over time (Appendix Figure 1). Pre-treatment estimates were generally centered near zero across components. Following adoption, certification/licensure and PD exhibited increasingly positive estimates at longer exposure periods. The materials estimates also became substantially larger in later post-adoption years despite the small overall ATT. Estimates for assessment and instruction/intervention showed less consistent evidence.

Policy Scope

Policy scope did not exhibit a simple dose-response relationship (Appendix Figure 2; Table 4, Panel B). States adopting one to two components had an estimated ATT of 2.26 points, although this estimate was imprecise and not statistically significant. States adopting three to four components experienced a statistically significant estimated increase of 2.13 points ($p < .01$). The estimate for states adopting five to six components was smaller and not statistically significant (ATT = 0.71). The event-study estimates show different dynamic patterns across scope categories, with later positive estimates particularly apparent among states adopting three to four components. These findings provide no evidence of a monotonic relationship in which adopting a greater number of policy components is associated with larger reading gains.

Policy Bundles

Panel C of Table 4 examines whether effects varied across different policy bundles. States adopting the teacher-knowledge bundle (comprising teacher preparation, certification/licensure, and PD) had an estimated ATT of 0.65 points ($p < .01$); however, the teacher-knowledge estimate is based on only two treated states and should therefore be interpreted cautiously. In contrast, the estimate for the systems support bundle (comprising instructional materials, assessment, and intervention) was -0.16 points and not significant. Comprehensive reforms containing five or more components were positive but not statistically significant. These estimates suggest that the composition of policy reforms, rather than simply the number of components adopted, may be relevant to their association with reading achievement.

Policy Strength

Finally, Panel D examines a data-driven measure of policy strength based on the first principal component of the six policy indicators. Estimated effects increased across the low-, medium-, and high-strength categories. The ATT was 0.24 points among low strength states and

1.04 points among medium-strength states; neither estimate was statistically significant. In contrast, the high-strength policy environments were associated with a 2.61 point increase in fourth-grade reading achievement ($p < .01$), equivalent to approximately 0.40 state-level standard deviations. Although these categories are descriptive rather than direct measures of implementation quality, the pattern suggests that the configuration of policies captured by the strength index may be more consequential than the number of adopted components alone.

Discussion

This study provides national evidence on the relationship between state adoption of SoR-aligned legislation and fourth-grade reading achievement. Across specifications, estimated effects were generally positive but varied in magnitude and precision, with the preferred CS model producing a small, statistically insignificant overall estimate. Event-study results, however, suggest meaningful variation by time since adoption. Estimated effects grew several years after implementation, reaching roughly 0.4 state-level standard deviations four to five years after the initial policy year. The pattern is broadly consistent with Spencer's (2024) analysis of Mississippi's LBPA, which found larger effects among students exposed to the policy throughout grades K-3. Although differences in the units of standardization preclude direct comparison of effect magnitudes, both studies point to the importance of time and sustained exposure in understanding the effects of early literacy policy.

This pattern is aligned with the policy implementation framework motivating this study, in which formal adoption is only the beginning of a process through which policy must alter the knowledge, tools, organizational supports, and instructional routines educators encounter in practice (Honig, 2006; McLaughlin, 1987). At the same time, time since initial adoption should not be interpreted as duration of implementation alone. States often enacted different literacy

reforms, reforms that often had different implementation timelines, meaning that later estimates may also reflect the time required for reforms to influence instructional practice.

The subgroup findings complicate the expectation that statewide literacy reform will necessarily produce comparable benefits across student populations. Estimated effects were generally positive across racial/ethnic and FRPL groups, but evidence of gains was larger and more consistent for White students and students not eligible for FRPL. These patterns should not be interpreted as differences in the effects of statewide SoR policies across groups given the lack of significant between-group differences. The findings therefore suggest an important limit to what statewide literacy reform alone can be expected to accomplish. Statewide SoR policies can influence curricula, assessments, and intervention systems, among other areas, but they do not necessarily compensate for systemic differences in the opportunities through which children develop language, knowledge, and familiarity with the world represented in texts. Our findings also suggest that for White and more affluent students, SoR reforms may leverage background knowledge in ways that accelerate their reading achievement absent these reforms. These findings also raise questions about the extent to which SoR-aligned policies attend not only to the mechanics of learning to read, but also to the knowledge, language, cultural experiences, and identities students bring to the act of reading.

Examining variation in the effect of SoR by policy features lent additional insight aligned to the study's conceptual framework. First, the policy component findings point toward educator knowledge and capacity as a potentially important mechanism through which SoR policy operates. Certification/licensure and PD were the two components most consistently associated with positive student reading achievement. Though teacher preparation also targets teacher pedagogical content knowledge, the longer time horizon for changes to take hold among a large

segment of the student population may explain the lack of demonstrable evidence for this policy component.

At the same time, efforts to change teachers' knowledge and instructional practice may carry unintended consequences that warrant consideration. Two concerns may be particularly relevant to contemporary reading reform: threats to teachers' professional identity and reductions in professional autonomy. Changing what teachers know is not necessarily the same as changing what teachers believe or do (Woulfin & Gabriel, 2020). Legislation can require participation in PD, establish licensure expectations, and mandate instructional materials, but it cannot mandate conceptual change. New knowledge must be reconciled with what teachers already believe to be true, including beliefs developed and reinforced through decades of personal and classroom experience. In Emily Hanford's *Sold a Story*, the difficulty of this reconciliation is evident in the voices of veteran educators reflecting on their prior reading instruction. One teacher with 29 years of experience described how their instructional beliefs aligned with what they already believed about how children learn, while another expressed sadness for former students after reconsidering practices they had once believed were effective (Hanford, 2022).

Recent research further illustrates the tension between consistency and professional judgement. Naidu et al. (2026) found that teachers implementing commercial reading programs varied considerably in their perceived autonomy: some valued the structure the programs provided, while others described constraints on their ability to adapt instruction to individual students. The challenge for policy, then, is not simply whether teachers should have autonomy, but where professional judgment should operate within an evidence-aligned instructional system.

Notably, the policy scope results do not support a simple dose-response relationship in which adding components consistently produces larger achievement gains: states adopting three

to four components showed positive effects, whereas estimates for states adopting five to six components were smaller and less precise. In contrast, the policy-strength analysis showed an increasingly positive pattern across low-, medium-, and high-strength policy environments, with the largest estimate among high-strength states. This latter pattern is consistent with our conceptualization of SoR reforms as interdependent instructional systems; preparation and licensure, PD, instructional materials, assessment, and intervention function as mutually reinforcing policy levers when they communicate consistent instructional expectations.

This interpretation is consistent with conceptualizations of instructional improvement as dependent on interconnected educational infrastructure (Cohen and Bhatt, 2012; Woulfin & Gabriel, 2020). Woulfin and Gabriel (2020), writing specifically about SoR reform, argue that curriculum, PD, and leadership function as mutually supportive pillars and that successful reading reform depends on the connections among them rather than any single element. The instructional coherence literature similarly emphasizes that educators are more likely to change practice when the policies, materials, professional learning, and expectations surrounding their work communicate a coherent vision rather than competing message (Coburn et al., 2016; Cohen & Bhatt, 2012; Newmann et al., 2001). From this perspective, the effects of SoR-aligned legislation may depend less on any one component than on whether reforms collectively build on the capacity and coherence needed to translate policy into practice. Importantly, our measures capture policy content and configuration rather than implementation quality or instructional coherence directly. Thus, the findings are consistent with this systems-based interpretation but do not directly test coherence as the mechanism producing achievement gains.

Implications for Research, Practice, and Policy

The findings have several implications for future research. First, our findings underscore the importance of viewing SoR-aligned reforms as heterogeneous in content and scope, as policy configurations and strength were more consistently associated with increased reading achievement than the number of reform components. Future research can not only examine how policy components are sequenced, combined, and sustained over time, but also how state policies affect the policy and instructional decisions of educators and school and district leaders. Research on policy implementation and sensemaking has long demonstrated that practitioners do not simply receive policy directives; they interpret, adapt, and sometimes transform them within local organizational and professional contexts (Coburn, 2001, 2004; Spillane et al., 2002). Future studies should therefore connect state policy measures to more proximal indicators of implementation, including participation in professional learning, curriculum use, coaching and implementation supports, teacher knowledge and practice, and the extent to which reforms reach schools and districts consistently. Such data would help distinguish whether variation in outcomes reflects differences in policy design, exposure, implementation, or the organizational conditions through which policies are enacted.

For policymakers, these findings underscore that state-level SoR reforms can improve reading achievement, though the effects are not immediate. The findings caution against equating policy comprehensiveness with the number of provisions enacted. Rather than continually adding requirements, states should consider whether policy levers are aligned and whether the infrastructure necessary to support their enactment is coherent, feasible, and sustainable (Woulfin & Gabriel, 2020). For practitioners and educational leaders, these findings reinforce that implementation should not rest on the individual teachers' ability to reconcile multiple policy demands on their own. Teachers interpret policy within professional communities and

organizational contexts, while school leaders can shape both access to policy ideas and the conditions under which teachers learn to enact them (Coburn, 2001, 2004). District and school leaders may therefore play an important role in creating coherence among instructional materials, assessments, professional learning, intervention systems, and other policy requirements. Implementation supports should provide sustained opportunities for educators to learn, enact, evaluate, and refine new practices rather than treating compliance with individual policy provisions as evidence that instructional change has occurred.

The subgroup findings underscore the importance of examining who benefits from statewide literacy reform. Although estimates were generally positive across several student groups and confidence intervals overlapped, statistically significant gains were more consistently observed among White students and students not eligible for FRPL. These results do not establish differential effects across groups, but neither should universal adoption be assumed to produce uniform implementation or outcomes. Future research should examine whether implementation conditions and access to reform supports vary systematically across schools and student populations, including whether deficit beliefs about students from low-income backgrounds shape expectations, instructional opportunities, or the enactment of reform. For policymakers and educational leaders, equitable implementation may therefore require attention not only to the content of state policy but also to the distribution of resources, expertise, organizational capacity, and opportunities for educators to examine beliefs and expectations that may shape implementation.

Finally, future research should also examine whether the achievement gains associated with these reforms are sustained over time. The relatively recent adoption of many SoR policies limits the post-policy periods available in this study, making it difficult to determine whether the

larger effects observed several years after adoption persist, continue to grow, or attenuate as reforms mature. This question is particularly salient given recent evidence from Mississippi. Although Spencer (2026) continued to find substantial effects of the state's early literacy policy on fourth-grade reading achievement, these gains did not appear to translate into improvements in later outcomes, raising questions about whether early literacy gains can be sustained without continued investment in subsequent grades.

Limitations

Several limitations should inform interpretation of these findings. First, the analysis identifies changes in state-level achievement associated with the timing of SoR-aligned legislation but cannot directly measure how those policies were enacted. Legislative requirements may be implemented unevenly across local contexts, and formal adoption does not indicate whether educators received the resources, professional learning, instructional materials, or organizational supports necessary to translate policy into changes in instruction. Consequently, the estimates capture the effects of state policy environments rather than the effects of implementation fidelity or specific instructional practices.

Second, assigning a single treatment date to multifaceted state literacy reform necessarily simplifies a policy process that often unfolds over several years. We defined treatment onset as the first year in which legislation meeting our operational definition of SoR-aligned reform became law. However, states frequently enacted literacy policies incrementally, with different components (such as PD, teacher preparation requirements, instructional materials, or intervention provisions) taking effect in different years. Other policy inventories may therefore identify different adoption dates because they capture broader early-literacy legislation, different policy features, or later stages of implementation. Our sensitivity analyses using alternative

classifications and timing decisions produced similar patterns, suggesting that the findings are not driven by a single treatment-timing decision. At the same time, the event-study results make transparent that alternative treatment dates would shift the interpretation of event time: what is labeled as the first, second, or third year following treatment depends on which point in an evolving policy process is designated as treatment onset. Thus, no single date fully captures the evolution of these policy environments.

Accordingly, the dynamic estimates are best interpreted as years relative to the first qualifying SoR-aligned policy under our operational definition rather than as precise measures of implementation duration. Increasingly positive estimates several years after initial adoption may reflect both the time required for enacted reforms to influence instructional systems and student outcomes and the continued accumulation or implementation of policy components after the initial treatment date. This study cannot fully separate these mechanisms, and later event-time coefficients should therefore be interpreted as effects associated with increasing exposure to an evolving policy environment rather than necessarily as the effect of a fully implemented policy package that began at time event zero.

Additionally, the analysis relies on state-level NAEP outcomes, which provide a common measure of achievement across states but limit both precision and the frequency with which outcomes are observed. Because NAEP is generally administered biennially, scores for intervening years were linearly interpolated to construct the annual panel. Interpolation assumes relatively smooth changes between observed assessment years and may be particularly problematic during periods of abrupt disruption (such as the COVID-19 pandemic). Estimates excluding interpolated observations were similar in direction to the primary results, but uncertainty about achievement trajectories between NAEP administrations remains.

Policy adoption is not random. States that enact SoR-aligned legislation, and states that adopt particular combinations of policy components, may differ from other states in political context, administration capacity, prior investments in literacy, or broader commitment to educational reform. The staggered-adoption estimator and pre-treatment diagnostics address important sources of bias associated with treatment timing, but they cannot eliminate confounding from concurrent state-level changes. Similarly, the component, bundle, scope, and policy-strength analyses should not be interpreted as isolating the causal contribution of individual policy features because components are frequently adopted together and their adoption may itself reflect broader state capacity.

Finally, subgroup analyses are constrained by the availability and precision of state-level NAEP estimates. Reporting limitations and smaller subgroup populations result in fewer state-year observations from some racial and ethnic groups, contributing to wide confidence intervals and limiting the ability to identify differences between groups. The subgroup findings should therefore be interpreted as evidence of variation in the pattern and precision of estimated effects rather than definite evidence that SoR-aligned policies produce different effects across student populations.

Conclusion

The rapid expansion of SoR legislation represents one of the most substantial shifts in state literacy policy in recent decades. Our findings suggest that evaluating this movement requires looking beyond whether states have adopted SoR-aligned legislation to how these reforms develop and operate as policy systems. Although average treatment effects were modest, improvements became more apparent with longer policy exposure, and exploratory analyses suggest meaningful variation according to the composition and strength of state policy

environments. At the same time, evidence of gains was not equally apparent across student populations. Together, these findings caution against treating the adoption of evidence-aligned policy as equivalent to its successful translation into practice. As states continue to invest in literacy reform, understanding how policy components are aligned, implemented, sustained, and experienced across educational contexts will be critical to determining whether the current wave of SoR legislation produces lasting and equitable improvements in reading achievement.

References

- Adams, M.J. (1990). *Beginning to read: Thinking and learning about print*. MIT Press.
- Basma, B., & Savage, R. (2018). Teacher professional development and student literacy growth: A systematic review and meta-analysis. *Educational Psychology Review*, 30(2), 457–481. <https://doi.org/10.1007/s10648-017-9416-4>
- Blazar, D., Heller, B., Kane, T. J., Polikoff, M., Staiger, D. O., Carrell, S., Goldhaber, D., Harris, D. N., Hitch, R., Holden, K. L., & Kurlaender, M. (2020). Curriculum reform in the Common Core era: Evaluating elementary math textbooks across six US states. *Journal of Policy Analysis and Management*, 39(4), 966-1019. <https://doi.org/10.1002/pam.22257>
- Burk, K. (2020). The perfect storm: Mississippi’s momentum for improving reading achievement. *The Reading League Journal*, 27–47.
- Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, 19(1), 5–51. <https://doi.org/10.1177/1529100618772271>
- Carr, S. (2022, December 27). NAACP targets a new civil rights issue—Reading (The Hechinger report). <https://hechingerreport.org/naACP-targets-a-new-civil-rights-issue-reading/>
- Chingos, M. M., & Whitehurst, G. J. (2012). *Choosing Blindly: Instructional Materials, Teacher Effectiveness, and the Common Core*. Brookings Institution.
- Clotfelter, C. T., Ladd, H. F., & Vigdor, J. (2005). Who teaches whom? Race and the distribution of novice teachers. *Economics of Education Review*, 24(4), 377-392. [10.1016/j.econedurev.2004.06.008](https://doi.org/10.1016/j.econedurev.2004.06.008)

- Coburn, C. E. (2001). Collective sensemaking about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis*, 23(2), 145–170. <https://doi.org/10.3102/01623737023002145>
- Coburn, C. E. (2004). Beyond decoupling: Rethinking the relationship between the institutional environment and the classroom. *Sociology of Education*, 77(July), 211-244.
- Coburn, C. E., Hill, H. C., & Spillane, J. P. (2016). Alignment and accountability in policy design and implementation: The Common Core State Standards and implementation research. *Educational Researcher*, 45(4), 243–251.
<https://doi.org/10.3102/0013189X16651080>
- Cohen, D. K., & Bhatt, M. P. (2012). The importance of infrastructure development to high-quality literacy instruction. *The Future of Children*, 22(2), 117-138. <http://doi.org/10.1353/foc.2012.0012>
- Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181–199. <https://doi.org/10.3102/0013189X08331140>
- Doan, S., Kaufman, J. H., Woo, A., Prado Tuma, A., Diliberti, M. K., & Lee, S. (2022). *How states are creating conditions for use of high-quality instructional materials in K–12 classrooms: Findings from the 2021 American Instructional Resources Survey* (RR-A134-13). RAND Corporation.
- EdReports. (2024). *State of the instructional materials market 2023: Teachers weigh in on how they're using materials and the support they need*.
<https://edreports.org/resources/article/state-of-the-instructional-materials-market-2023>

- Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research*, 71(3), 393-447.
<https://doi.org/10.3102/00346543071003393>
- Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading*, 9(2), 167–188. https://doi.org/10.1207/s1532799xssr0902_4
- Elzy-Palmer, J., Babino, A., & Hubbard, T. (2025). Dimensions of equity in the Science of Reading research: A systematic review of actual, artificial, and absent up-takes of equity. *Reading Research Quarterly*, 60(4), e70070.
- Gamse, B. C., Jacob, R. T., Horst, M., Boulay, B., Unlu, F., Bozzi, L., Caswell, L., Rodger, C., Smith, W. C., Brigham, N., & Rosenblum, S. (2008). Reading first impact study. Institute of Education Sciences. <https://ies.ed.gov/pubsearch/pubsinfo.asp?pubid=NCEE20094038>
- Gersten, R., Haymond, K., Newman-Gonchar, R., Dimino, J., & Jayanthi, M. (2020). Meta-analysis of the impact of reading interventions for students in the primary grades. *Journal of Research on Educational Effectiveness*, 13(2), 401–427.
<https://doi.org/10.1080/19345747.2019.1689591>
- Goldhaber, D., Lavery, L., & Theobald, R. (2015). Uneven playing field? Assessing the teacher quality gap between advantaged and disadvantaged students. *Educational Researcher*, 44(5), 293-307. <https://doi.org/10.3102/0013189X15592622>
- Hanford, E. (2022). *Sold a story* [Audio podcast series]. American Public Media.
- Hernandez, D. (2011). Double Jeopardy: How Third-Grade Reading Skills and Poverty Influence High School Graduation. *Annie E. Casey Foundation*

- Honig, M. I. (Ed.). (2006). *New directions in education policy implementation: Confronting complexity*. State University of New York Press.
- Ingersoll, R. M. (2009). *Who controls teachers' work? Power and accountability in America's schools*. Cambridge, MA: Harvard University Press.
- Kaufman, D. (2022, October 6). In Mississippi, a broad effort to improve literacy is yielding results. *The New York Times*.
- Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, 88(4), 547–588. <https://doi.org/10.3102/0034654318759268>
- McLean, C., Gallagher, T., & Hattan, C. (2026). An anatomy of the SoR: A systematic review of the conceptualizations and instrumental uses of the science of reading. *Reading Research Quarterly*, 61(2), e70095. <https://doi.org/10.1002/rrq.70095>
- McLaughlin, M. W. (1987). Learning from experience: Lessons from policy implementation. *Educational Evaluation and Policy Analysis*, 9(2), 171-178. <https://doi.org/10.3102/01623737009002171>
- Moats, L. C. (2020). Teaching Reading" Is" Rocket Science: What Expert Teachers of Reading Should Know and Be Able to Do. *American Educator*, 44(2), 4.
- Moon, J.M., Camburn, E.M., Sebastian, J. (2022). Streamlining your school: Understanding the relationship between instructional coherence and school performance. *School Effectiveness and School Improvement*, 33(2), 260-279.
- Naidu, R., Campbell, S., & Walker, S. (2026). Balancing the Chalk and the Check: Teacher Autonomy Amid the Rise of Scripted Reading Programs. *The Reading Teacher*, 80(2), e70074. <https://doi.org/10.1002/trtr.70074>

- Narayanan, M., Shields, A. L., & Delhagen, T. J. (2024). Autonomy in the spaces: Teacher autonomy, scripted lessons, and the changing role of teachers. *Journal of Curriculum Studies*, 56(1), 17-34. <https://doi.org/10.1080/00220272.2023.2297229>
- National Center for Education Statistics. (2024). *NAEP Reading: National Trends and Student Skills, Grades 4 & 8*. The Nation's Report Card. U.S. Department of Education. https://www.nationsreportcard.gov/reports/reading/2024/g4_8/national-trends/
- National Council on Teacher Quality. (2023). *Teacher prep review: Strengthening elementary reading instruction*. National Council on Teacher Quality. <https://www.nctq.org/research-insights/teacher-prep-review-strengthening-elementary-reading-instruction/>
- Newmann, F. M., Smith, B., Allensworth, E., & Bryk, A. S. (2001). Instructional program coherence: What it is and why it should guide school improvement policy. *Educational Evaluation and Policy Analysis*, 23(4), 297-321. <https://doi.org/10.3102/01623737023004297>
- Nguyen, T. D., & Redding, C. (2026). *Ending Teacher Shortages*. Johns Hopkins University Press.
- Novicoff, S., & Dee, T. S. (2025). The Achievement Effects of Scaling Early Literacy Reforms. *Educational Evaluation and Policy Analysis*. <https://doi.org/10.3102/01623737251349178>
- Pearson, P. D. (2004). The reading wars. *Educational Policy*, 18(1), 216-252. <https://doi.org/10.1177/0895904803260041>
- Powell, R., Cantrell, S. C., & Correll, P. (2017). Power and Agency in a High Poverty Elementary School: How Teachers Experienced a Scripted Reading Program. *Journal of Language and Literacy Education*, 13(1), 93-124. <https://eric.ed.gov/?id=EJ1141492>

Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement.

American Educational Research Journal, 50(1), 4–36.

<https://doi.org/10.3102/0002831212463813>

Scorrano, D. (2021). Ask an expert: the magic of Mississippi’s literacy change. *The*

Beacon. [https://www.thewindwardschool.org/the-windward-institute/the-](https://www.thewindwardschool.org/the-windward-institute/the-beacon/article/~board/beacon-archives/post/ask-an-expert-the-magic-of-mississippis-literacy-change)

[beacon/article/~board/beacon-archives/post/ask-an-expert-the-magic-of-mississippis-literacy-change](https://www.thewindwardschool.org/the-windward-institute/the-beacon/article/~board/beacon-archives/post/ask-an-expert-the-magic-of-mississippis-literacy-change)

Shanahan, T. (2020). What constitutes a science of reading instruction? *Reading Research*

Quarterly, 55, S235-S247. <https://doi.org/10.1002/rrq.349>

Spencer, N. (2024). Comprehensive early literacy policy and the “Mississippi Miracle”.

Economics of Education Review, 103, 102598.

<https://doi.org/10.1016/j.econedurev.2024.102598>

Spencer, N. (2026). Long-Run Effects of Early Literacy Policy: Evidence from Mississippi.

<http://dx.doi.org/10.2139/ssrn.6760038>

Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition:

Reframing and refocusing implementation research. *Review of Educational Research*, 72(3), 387–431. <https://doi.org/10.3102/00346543072003387>

Supovitz, J. A. & Weinbaum, E. H. (2010). *The implementation gap: Understanding reform in high schools*. New York, NY: Teachers College Press.

Valli, L., & Buese, D. (2007). The changing roles of teachers in an era of high-stakes accountability. *American Educational Research Journal*, 44(3), 519-558.

<https://doi.org/10.3102/0002831207306859>

- Wanzek, J., & Vaughn, S. (2007). Research-based implications from extensive early reading interventions. *School Psychology Review*, 36(4), 541–561.
<https://doi.org/10.1080/02796015.2007.12087917>
- Westall, J., & Cummings, A. (2023). *The effects of early literacy policies on student achievement* (EdWorkingPaper No. 23-788). Annenberg Institute at Brown University.
<https://edworkingpapers.com/sites/default/files/ai23-788.pdf>
- What Works Clearinghouse (2007). WWC topic report. Beginning reading. U.S. Department of Education, Institute for Education Sciences.
- Woulfin, S. L. (2020). Coach professional development in the urban emergent context. *Urban Education*, 55(10), 1355–1384. <https://doi.org/10.1177/0042085917714513>
- Woulfin, S., & Gabriel, R. (2020). Interconnected infrastructure for improving reading instruction. *Reading Research Quarterly*, 55(S1), 109-117.
<http://doi.org/10.1002/rrq.339>

Figure 1. Event study analysis of the effects of SoR-aligned legislation on fourth-grade reading scores.

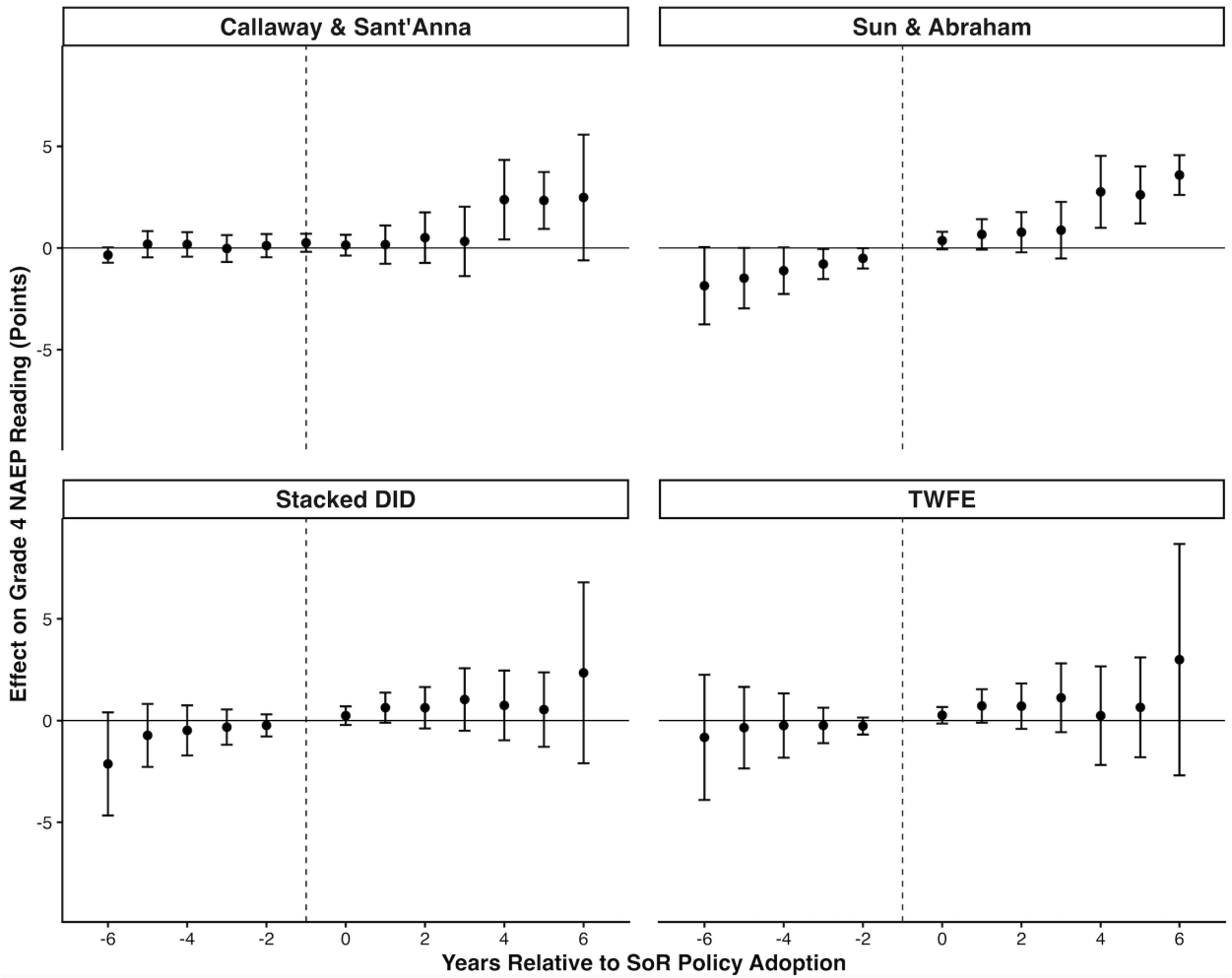
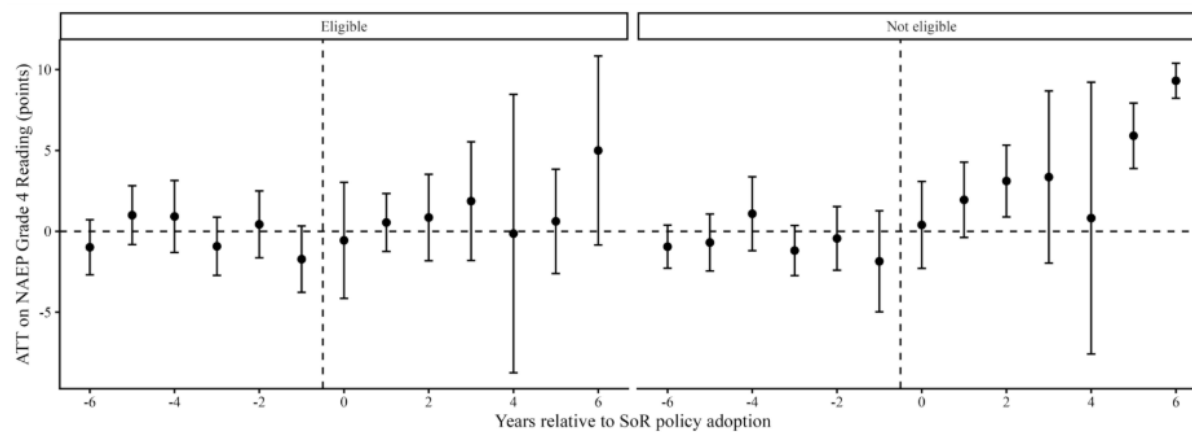
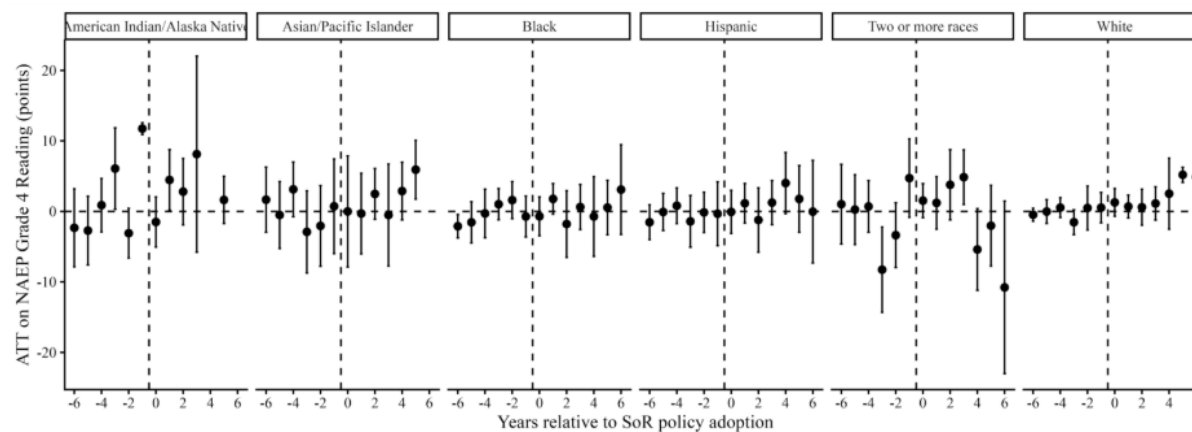


Figure 2. Event study analysis of the effects of SoR-aligned legislation on fourth-grade reading scores, by Free and Reduced Price Lunch Eligibility and Race/Ethnicity

Panel A. Free and Reduced Price Lunch Eligibility



Panel B. Race/Ethnicity



Notes. Estimates are from enrollment-weighted Callaway and Sant'Anna staggered-adoption DiD models using not-yet-treated states as the comparison group. Event time is measured in years relative to SoR policy adoption. Points represent coefficient estimates and bars indicate 95% confidence intervals. The dashed vertical line marks the reference period.

Table 1. Coding of State Science of Reading Legislation

State	Year Passed	SoR Legislation	Teacher Prep.	Teacher Certification	PD/Coaching	Assessment	Materials	Instruction/Intervention
AL	2019	X	X	X	X	X	X	X
AK	2022	X	X		X	X	X	X
AZ	2010 2022	X			X	X	X	X
AR	2017 2021	X	X	X	X		X	X
CA	2021, 2023	X	X	X	X	X		
CO	2019	X			X	X	X	X
CT	2021	X			X	X	X	X
DE	2022	X			X		X	
DC	2020	X			X	X	X	X
FL	2021, 2023	X		X	X	X	X	X
GA	2023	X	X	X	X	X	X	X
HI	--							
ID	2021	X			X	X	X	X
IL	2023	X	X	X	X	X	X	
IN	2023	X	X	X	X		X	X
IA	2024	X	X	X		X		X
KS	2022, 2023	X		X	X	X	X	X
KY	2022	X	X	X	X	X	X	X
LA	2021	X			X			
ME	--							
MD	2019	X	X	X	X	X	X	X
MA	--							
MI	2016	X	X	X	X	X		X
MN	2021	X			X	X	X	X
MS	2013, 2016	X	X	X	X	X	X	X
MO	2018, 2022, 2023, 2025	X	X	X	X	X	X	X
MY	--							

State	Year Passed	SoR Legislation	Teacher Prep.	Teacher Certification	PD/ Coaching	Assessment	Materials	Instruction/ Intervention
NE	2018	X				X		X
NV	2015, 2019	X	X	X	X	X		X
NH	--							
NJ	2024	X			X	X		
NM	2019, 2023	X	X		X	X	X	X
NY	--							
NC	2021, 2023	X	X	X	X	X	X	X
ND	2021	X			X		X	X
OH	2023	X	X		X	X	X	X
OK	2019, 2021	X	X		X	X	X	X
OR	2023	X			X	X	X	X
PA	2021, 2024	X	X		X	X	X	X
RI	2019	X	X	X	X			
SC	2014	X	X		X			X
SD	--							
TN	2021	X	X	X	X	X	X	X
TX	2019, 2023	X		X	X		X	X
UT	2022	X	X		X	X	X	X
VT	--							
VA	2022, 2023	X	X	X	X	X	X	X
WA	--							
WV	2019, 2023	X	X	X	X	X	X	X
WI	2023	X	X	X	X	X	X	X
WY	--							

Notes. SoR = Science of Reading; PD = professional development. Years indicate enactment of SoR-aligned legislation identified in the policy review. X indicates that the legislation included the specified policy component. Dashes indicate that no qualifying SoR legislation was identified during the study period.

Table 2. The effect of SoR-aligned legislation on fourth-grade reading scores

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	C&S	S&A	Stacked	TWFE	C&S Unweighted	C&S Never treated	C&S Full panel	Drop borderline states	No linear interpolation	Pre- COVID	Post- COVID
SoR legislation	0.33 (0.35)	1.04*** (0.20)	2.35** (0.90)	0.82 (0.72)	0.50 (0.54)	0.60 (0.37)	0.34 (1.21)	0.30 (0.34)	0.43 (1.52)	1.59 (0.32)	0.52 (0.49)
Observations	1122	1122	5214	1122	1122	1122	1122	1012	561	867	102

Notes. Standard errors are shown in parentheses. Columns 1-4 report the primary estimator comparisons. Columns 5-9 report robustness checks. Columns 10-12 report calendar-period aggregations of Callaway and Sant'Anna calendar-time treatment effects using the not-yet-treated comparison group. Observations in Columns 10-12 indicate the number of state-year observations contributing to each period-specific estimate. Estimates are aggregated across 7 pre-COVID years, 3 COVID-period years, and 2 post-COVID years, respectively. All weighted estimates use state-year public school enrollment weights. + $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 3. The effect of SoR-aligned legislation on fourth-grade reading scores, by FRPL eligibility and race/ethnicity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	FRPL- eligible	Not FRPL eligible	American Indian/ Alaska Native	Asian/ Pacific Islander	Black	Hispanic	Two or more races	White
SoR legislation	1.00 (0.96)	2.37* (0.98)	3.91 (4.10)	0.63 (2.42)	0.99 (0.97)	1.02 (1.17)	1.48 (1.49)	2.06* (0.86)
Observations	510	509	152	351	472	512	319	561

Notes. Standard errors are shown in parentheses. Estimates are overall average treatment effects from enrollment-weighted Callaway and Sant'Anna staggered-adoption difference-in-differences models using not-yet-treated states as the comparison group.

Observations indicate subgroup-specific state-year observations. † $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

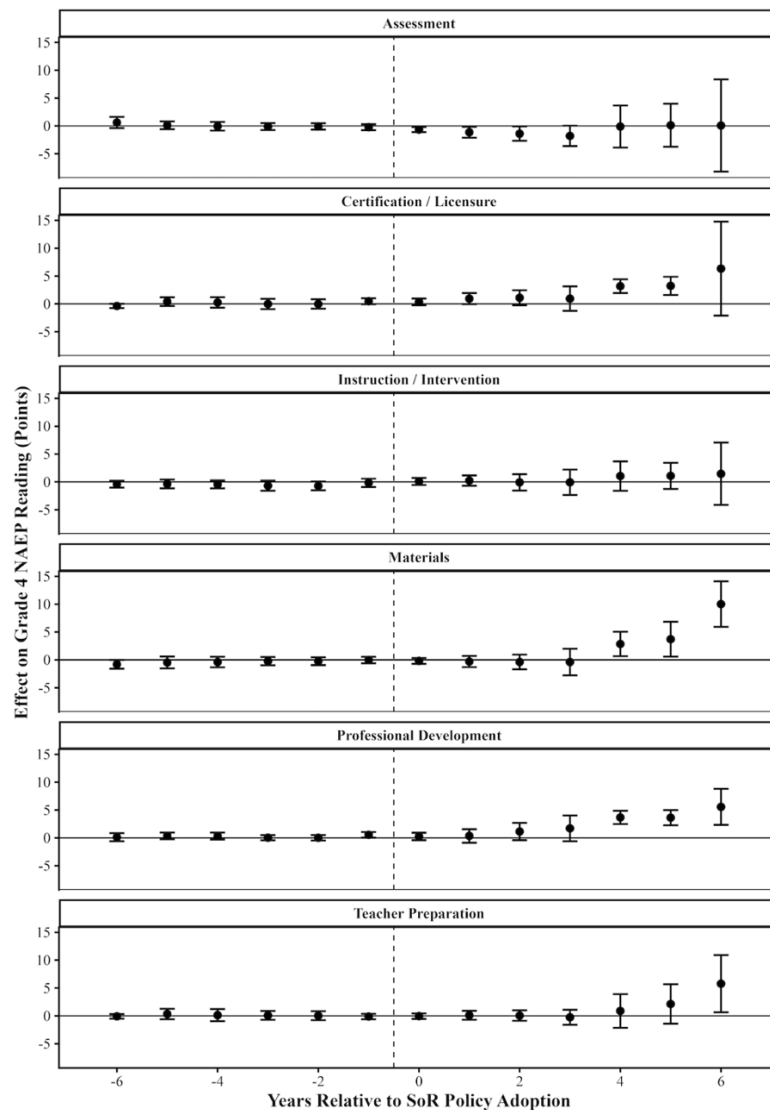
Table 4. Variation in the Effect of SoR-aligned legislation, by Policy Features

Panel A. Policy Components					
Component	N	ATT	SE	CI	p
Teacher Preparation	26	0.75	0.72	[-0.67, 2.17]	0.30
Certification / Licensure	21	1.43	0.68	[0.10, 2.76]	0.04*
Professional Development	38	1.66	0.64	[0.40, 2.93]	0.01
Assessment	33	-0.57	0.75	[-2.04, 0.91]	0.45
Instructional Materials	32	0.28	0.72	[-1.13, 1.69]	0.70
Instruction / Intervention	35	0.55	0.82	[-1.05, 2.15]	0.50
Panel B. Policy Scope					
Scope	N	ATT	SE	CI	p
1–2	4	2.26	2.13	[-1.91, 6.43]	0.29
3–4	14	2.13	0.80	[0.56, 3.71]	0.008**
5–6	23	0.71	0.86	[-0.97, 2.38]	0.41
Panel C. Policy Bundles					
Bundle	N	ATT	SE	CI	p
Teacher knowledge	2	0.65	0.24	[0.19, 1.12]	0.006**
Systems supports	7	-0.16	1.31	[-2.73, 2.41]	0.90
Comprehensive (5+)	23	0.71	0.85	[-0.95, 2.36]	0.40
Panel D. Policy Strength					
Strength	N	ATT	SE	CI	p
Low strength	17	0.24	1.34	[-2.39, 2.87]	0.86
Medium strength	12	1.04	1.01	[-0.93, 3.01]	0.30
High strength	12	2.61	0.89	[0.87, 4.35]	0.003**

Note. Enrollment-weighted C&S DiD models using never-treated states as the comparison group, with state-clustered standard errors. Component models are estimated from a common underlying state-year analytic sample; N indicates the number of states treated by each focal component. For policy scope, bundle, and strength models, N indicates the number of treated states in each category, with each estimate based on states in the indicated category and never-treated states. Policy scope represents the number of SoR policies. Policy bundles are defined as (a) teacher knowledge supports (teacher preparation, certification/licensure, and professional development), (b) system supports (instructional materials, assessment, and intervention), and (c) comprehensive multi-component reforms (five or more components). Policy strength is based on the first principal component from a principal component analysis of six final policy component indicators: teacher preparation, certification/licensure, professional development, assessment, instructional materials, and instruction/intervention. Treated states are grouped into low-, medium-, and high-strength categories based on tertiles of the PCA score. 95% confidence intervals are calculated as estimate $\pm 1.96 \times SE$. † $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

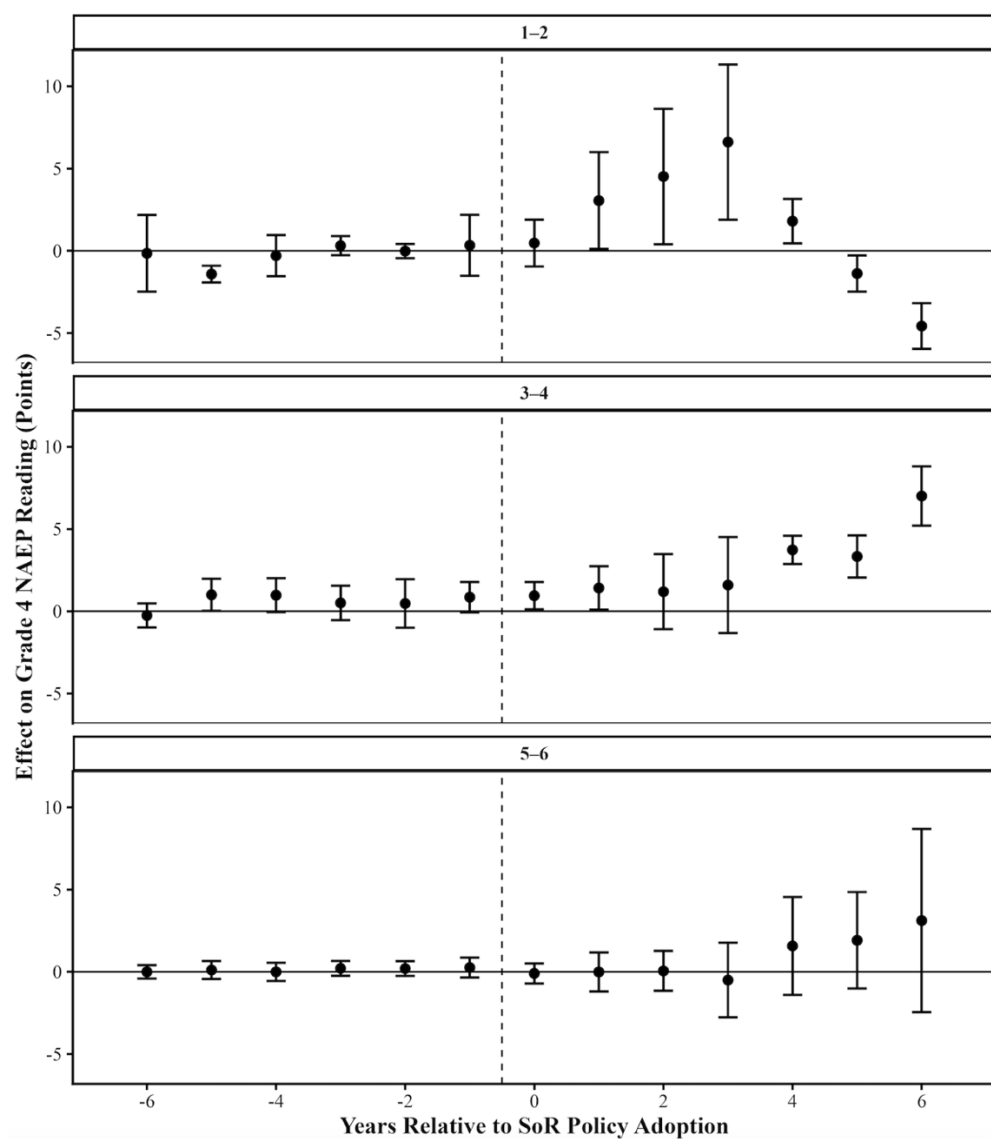
Online Appendix

Appendix Figure 1. Component-Specific Event-Study Estimates (Callaway and Sant'Anna, Never-Treated Comparison Group)



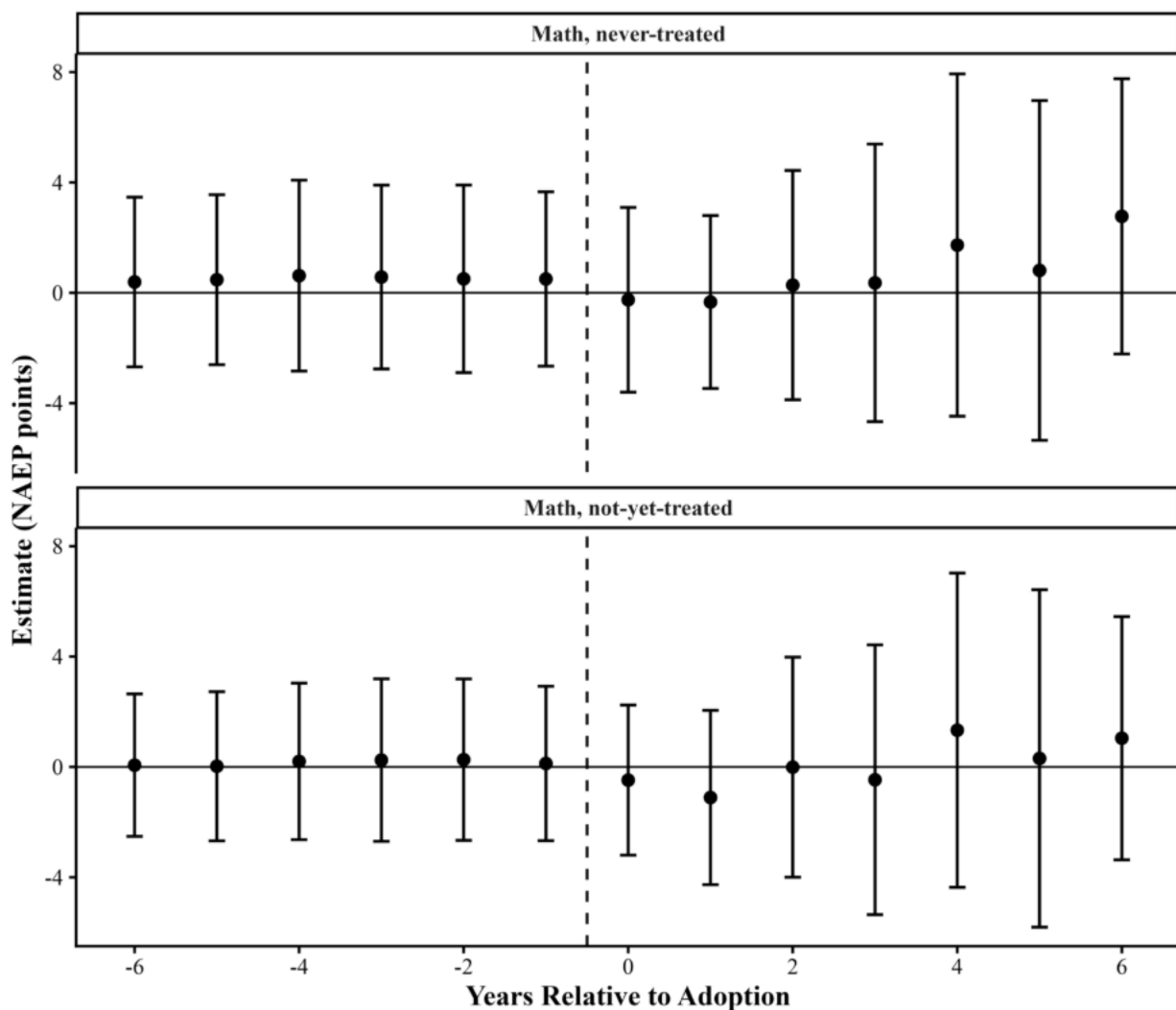
Note. Points represent dynamic treatment effects from the Callaway and Sant'Anna DiD estimator, plotted in event time relative to the first year of policy adoption (event time = 0). Error bars indicate 95% CI based on state-clustered SE. The dashed vertical line marks the first year of adoption. Estimates are enrollment-weighted and comparison is the never-treated group. The figure is presented over a common event-time window (−6 to +6) for comparability. Pre-treatment coefficients (−6 to −1) are used to assess the parallel trends assumption.

Appendix Figure 2. Event-Study Estimates for Policy Scope: Never-Treated



Note. Points represent dynamic treatment effects from the Callaway and Sant’Anna DiD estimator, plotted in event time relative to the first year of policy adoption (event time = 0). Error bars indicate 95% CI based on state-clustered SE. The dashed vertical line marks the first year of adoption. Estimates are enrollment-weighted and compare each treated scope bin (1 – 2, 3 – 4, 5 – 6 components) to the never-treated group. The figure is presented over a common event-time window (–6 to +6) for comparability; however, estimates are only shown for event times with sufficient data. Pre-treatment coefficients (–6 to –1) are used to assess the parallel trends assumption. Comparison is the never-treated.

Appendix Figure 3. Event-study estimates for the effect of SoR Policies on NAEP grade 4 mathematics (falsification test).



Note. Points represent dynamic treatment effects estimated using the Callaway and Sant'Anna DiD estimator on NAEP Grade 4 Mathematics. Event time is measured in years relative to policy adoption (event time = 0), with the year prior to adoption (-1) serving as the reference period. Error bars indicate 95% confidence intervals based on state-clustered standard errors. The dashed vertical line arks the first year of policy adoption. Estimates are enrollment-weighted and shown separately for models using never-treated and not-yet-treated comparison groups.

Appendix Table A1. State science of reading aligned laws.

State	Year Passed	Legislation Title	Description/ Key Provisions
AL	2019	Alabama Literacy Act (HB 388)	HB388 established the Alabama Literacy Act, creating a task force and requiring evidence-based reading programs for K – 3, monitoring of student progress, alerts for reading deficiencies, individualized intervention plans, summer camps, and retention policies tied to proficiency levels. Districts are required to provide core curriculum, PD, and coaching aligned with SoR. Teachers must deliver explicit and systematic instruction in reading skills. Requires preservice teachers to pass a foundational reading assessment (Alabama Legislature, 2019; Schwartz, 2024).
AK	2022	Alaska Read Acts (HB 114)	HB 114 requires schools to provide evidence-based support for students with reading deficiencies and mandates teacher preparation in the five components of reading identified by the National Reading Panel. The law establishes early identification and intervention procedures and requires professional learning aligned to structured literacy practices (Alaska Legislature, 2022; Schwartz, 2024)
AZ	2022	Early Literacy; Reading Plans (HB 2026; MOWR update)	Arizona’s Move On When Reading framework, originally enacted in 2010, established statewide literacy requirements for K–3, including evidence-based reading curricula and early screening for reading difficulties. Although the 2010 statute predates the modern SoR movement, its principles align closely with SoR by emphasizing the five essential components of reading identified by the National Reading Panel. The 2022 amendment (HB 2026) retained this foundation while revising implementation procedures—modifying the submission frequency of district literacy plans and expanding exemptions to the third-grade retention policy, particularly for English Learners (Arizona Legislature, 2022).

AR	2017	Right to Read Act (Arkansas Code § 6-17-429)	Right to Read Act requires all K-6 teachers to demonstrate proficiency in SoR and all other teachers and administrators to demonstrate awareness of SoR. It also established teacher training and licensure requirements tied to the SoR (Arkansas Legislature, 2017).
	2021	SB 349 (Act 606 of 2021)	SB 349 prohibits the use of three-cueing system model of reading (“MSV”). Schools found in violation may lose up to 10% of their state funding. It reinforces state alignment with SoR and strengthens enforcement mechanisms (Arkansas Legislature, 2021).
CA	2021	Teaching Credentialing: Reading Instruction (SB 488; Chapter 678, Statutes of 2021)	SB 488 requires teacher preparation programs to demonstrate they prepare candidates to deliver foundational reading skills instruction, including for ELL. Beginning in 2025 elementary and SPED candidates must pass a new literacy performance assessment aligned to structured literacy (California Legislature, 2021).
	2023	Literacy Coaches and Reading Specialists Program (2023-2024 State Budget Act via SB 114)	The 2023-2024 State Budget Act allocates state funds for literacy coaches, reading screening assessments, and development of statewide literacy roadmaps. It expands state-level investment in structured literacy implementation beyond teacher preparation (California Department of Finance, 2023).
CO	2019	READ Act Implementation Measures (SB 19-199)	SB 19-199 requires districts to provide evidence-based reading curriculum and instruction focused on the five components of reading. The law also mandates that all K – 3 teachers complete training grounded in SoR and that districts adopt scientifically based core and supplemental programs (Colorado General Assembly, 2019).
CT	2021	Right to Read (HB 6620; Public Act 21-2, Sec. 415-16)	HB 6620 requires the Connecticut State DOE to develop a research-based reading plan aimed at improving performance for struggling students and to establish a Center for Literacy Research and Reading Success. The Center must approve at least five evidence-based reading curriculum programs or models and five assessments. Beginning in the 2023-2024 school year, all districts must implement one of the approved

			curricula and use the approved assessments (Connecticut General Assembly, 2021).
DE	2022	Evidence-Based Reading Curricula Act (SB 4; Chapter 4, Volume 84)	SB 4 requires the DOE to publish and maintain a list of evidence-based K – 3 reading curricula aligned with SoR. By the 2027 – 2028 school year, all districts and charter schools serving grades K – 3 must (1) adopt a curriculum from the state approved list, (2) ensure all educators teaching reading complete aligned PD, and (3) designate a school-based supervisory role to support implementation. Districts must submit annual reports on implementation and DOE must publish a statewide summary report (Delaware General Assembly, 2023).
DC	2020	Addressing Dyslexia and Other Reading Difficulties Amendment Act of 2020 (D.C. Law 23-191)	D.C. Law 23-191 requires LEAs to adopt a SoR-based reading program, provide PD for educators to recognize and address reading difficulties, mandates universal screening for early identification of reading difficulties, assists on structured intervention, requires compliance reporting to OSSE (District of Columbia Council, 2020).
FL	2021	Student Literacy Act (HB 7011)	The Student Literacy act implemented statewide literacy screening and progress monitoring (grades pre K – 8); established literacy support regions and reading coaches; required teacher reading endorsements and professional learning grounded in evidence-based practices; revised licensure to include SoR content; expanded Just Read, Florida! And RAISE support programs (Florida Legislature, 2021).
	2023	Student Outcomes Act (HB 7039)	The Student Outcomes act reinforces SoR instruction; bans three-cueing in foundational reading and requires phonics-based primary; allocates funding for implementation; requires teacher preparation programs and district coaching/PD to align with SoR; mandates improved early warning and intervention structures (Florida Legislature, 2023).
GA	2023	Georgia Early Literacy Act (HB 538)	The Georgia Early Literacy Act requires the SBOE to approve a list of high-quality instructional materials for teaching K – 3 reading and a list of screening tools

			to identify struggling students. The DOE must also provide a free statewide screener. Districts are required to (1) develop individual reading plans for struggling students, (2) adopt instructional materials from the state-approved list or certify their chosen materials meet state standards for high-quality resources, and (3) ensure teachers receive state-developed or procured training on SoR, structured literacy, and foundational reading skills. Additionally, teacher certification assessments must align with evidence-based literacy instruction (Georgia General Assembly, 2023).
HI	N/A	N/A	Hawaii has not passed any legislation or non-legislated policy mandating SoR instruction.
ID	2021	Idaho Literacy Achievement and Accountability Act (Idaho Statutes § 33-1803 et seq.)	The Idaho Literacy Achievement mandates that all students must have access to evidence-based reading instruction focused on foundational reading skills, including phonemic awareness, phonics, fluency, vocabulary, and comprehension (Idaho Legislature, 2021, The Reading League, 2025).
IL	2023	Comprehensive Literacy Plan Act (CSB 2243 – Public Act 103-0402)	SB 2243 directs the Illinois SBOE to develop a comprehensive statewide literacy plan emphasizing systematic, explicit instruction in foundational reading skills, including phonemic awareness, phonics, fluency, vocabulary, and comprehension. The plan must include a rubric for evaluating curricula, a district literacy plan template, and guidance for literacy coaches. By January 2025, ISBE must make training in evidence-based reading instruction available to educators and align teacher preparation and certification assessments with these competencies (Illinois General Assembly, 2023).
IN	2023	Science of Reading Act (HB 1558)	Indiana’s HB1558 formally defines the “Science of Reading” in law, requiring instruction to include explicit, systematic teaching of the five essential components: phonemic awareness, phonics, fluency, vocabulary, and comprehension. Beginning in the 2024–2025 school year, schools must adopt curricula or supplemental materials aligned with SoR and are prohibited from using three-cueing models. The law also requires teacher

			preparation programs to integrate SoR principles, establishes literacy endorsements for newly licensed teachers (effective post-June 30, 2025), mandates annual reporting on reading and remedial programs, and directs the state to publish an advisory SoR-aligned curricula list (Indiana Legislature, 2023).
IA	2024	An Act Relating to Early Literacy Implementation, Including Personalized Reading Plans and Retention Notices (House File 2618)	HF 2618 transitioned Iowa's Early Literacy Implementation law to SoR by requiring teacher preparation programs to administer the Foundations of Reading assessment for licensure, mandating personalized reading plans for K–6 students not reading proficiently, and providing parental discretion for grade retention. The law emphasizes evidence-based reading instruction and accountability for both schools and preparation programs. Iowa Code §279.68 continues to require universal screening, progress monitoring, intensive instruction for students at risk, and parent notification (Iowa Legislature, 2024; Iowa Department of Education, 2024).
KS	2022	Every Child Can Read Act (HB 2567)	This law requires school to ensure reading literacy is attained through SoR and evidence-based instruction, specifically including phonics, phonemic awareness, vocabulary, fluency, and comprehension. Districts must use appropriate screening tools, interventions, and report on outcomes.
	2023	Kansas Blueprint for Literacy Act (SB 438)	SB 438 establishes a comprehensive literacy reform framework including deposit of SoR and structured literacy as the official literacy methodology, a literacy advisory committee and director of literacy education, required micro-credentials in SoR for all elementary teachers (by 2030), ban on three-cueing methods, centers of Excellence in Reading, and a comprehensive assessment system, and aggressive student literacy outcome goals.
KY	2022	Read to Succeed / Early Literacy Education Act (SB 9)	SB 9 requires every elementary school to provide comprehensive, schoolwide reading instruction aligned to the state standards; establishes MTSS for reading; mandates evidence-based intervention in phonemic awareness, phonics, fluency, vocabulary, comprehension; requires a universal screener

			and diagnostic assessment for K – 3 (includes training for teachers on how to administer and interpret); requires K – 3 schools to adopt a comprehensive reading program; creates a “Read to Succeed” fund for professional learning academies, coaching, models, and resources; requires new teachers to pass approved reading instruction tests (Kentucky General Assembly, 2022).
LA	2021	Early Literacy Professional Development Act (SB 216)	SB 216 requires the state DOE to create a list of approved PD courses grounded in SoR. Teachers and administrators must complete one approved PD course within two years of employment. Local school boards must report annual completion rates (Louisiana Legislation, 2021).
ME	N/A	N/A	Maine does not currently have a comprehensive statewide SoR law. While they have enacted policies related to early literacy and dyslexia, such as requiring dyslexia screenings, these do not mandate structured literacy or statewide adoption of SoR.
MA	2019,	Ready to Read Act (SB 734)	SB 745 required universal screening for kindergarten and early identification for students in grades 1 – 3 at risk for reading difficulties. It also mandated evidence-based supplemental reading instruction, progress monitoring, parent notification, and annual reporting by districts. Screening requirements must be based on foundational reading skills (Maryland General Assembly, 2019).
	2022	Code of Maryland Regulations (COMAR) Update	The COMAR Regulatory Update expanded requirements by explicitly extending screening to grades 1 – 3, adding evaluation of screening programs, and increasing accountability for LEAs (Maryland State DOE, 2022).
	2024	Statewide Adoption and Implementation of the SoR(Resolution 24-01)	Resolution 24-01 mandated full statewide alignment of literacy instruction with SoR. Directed the state superintendent to develop comprehensive policies on curriculum universal screening, interventions, teacher preparation and PD, early warning systems, and accountability. (Mary State Board of Education, 2024).
Massachusetts (MA)	N/A	N/A	Massachusetts does not currently have a comprehensive statewide SoR law. Existing

			policies focus on dyslexia screening and supports, such as early literacy screening regulations and the Act Relative to Students with Dyslexia, but does not mandated structured literacy or SoR.
MI	2016	Read by Grade Three Act (SB 12)	Read by Grade Three law required schools to provide systematic, explicit multisensory reading instruction and targeted interventions for K – 3 students. It also established a third-grade reading retention requirements for students who did not meet reading benchmarks and created a state literacy coaching network (Michigan Legislature, 2016).
	2023	Third-Grade Reading Proficiency Repeal (SB 12; Public Act 7 of 2023)	In March 2023, the legislature repealed the mandatory retention policy while maintaining early screening, intervention, and parent notification provisions (Michigan Legislature, 2023).
	2024	Dyslexia Screening and Teacher Preparation (Reading / SoR Bill Package; SB 567; SB 568)	In 2024, SB 567 and SB 568 expanded the policies. SB 567 requires schools to provide interventions grounded in SoR. SB 568 requires teacher preparation programs and alternative certification to provide instruction in structured literacy and SoR-aligned practices for candidates seeking certification in reading, language arts, and special education (Michigan Legislature, 2024a, Michigan Legislature, 2024b).
MN	2021	Teacher Training Appropriation (omnibus education finance law)	In 2021, the legislature appropriated \$3 million to provide teachers with PD in evidence-based reading instruction methods grounded in SoR (Minnesota Legislature, 2021).
	2023	Read Act / Reading to Ensure Academic Development Act (HF 2497)	The Read Act (HF 2497) in 2023 established a comprehensive statewide literacy framework. It requires districts to provide systematic, evidence-based instructional materials. The act also mandates universal screening of K – 3 students for reading difficulties, requires districts to create local literacy plans, and directs the MN DOE to support implementation and accountability. Implementation is phased, with statewide compliance expected by 2026 – 2028 (Minnesota DOE, 2023; Minnesota Legislature, 2023).
MS	2013	Literacy-Based Promotion Act (SB 2347)	Mississippi's Literacy-Based Promotion Act (LBPA) of 2013 established a statewide system

			for improving literacy outcomes. It requires schools to screen students in grades K – 3 for reading difficulties, develop individual reading plans for struggling students, and notify parents of any deficiencies. The law also created a system of literacy coaches to provide job embedded PD to teachers in evidence-based reading instruction. Start in 2015, promotion to 4 th grade was tied to reading proficiency, with students retained in 3 rd entitled to intensive interventions (Mississippi DOE, 2024).
	2016	Amendments to LBPA (SB 2157)	Amendments in 2016 expanded the law by strengthening the PD requirements, increasing the deployment of state-supported literacy coaches, and ensuring districts had structured systems in place for early identification, intervention, and parental communication. These measures further aligned with instructional practices in SoR and provided additional supports for students who were retained (Mississippi DOE, 2024).
MO	2018	Early Literacy Statutes § 167.645	Missouri has passed a series of laws strengthening its statewide literacy framework. 2018 legislation established early literacy supports, setting the stage for later reform, including embedded early literacy supports in statute via §167.645 (Missouri General Assembly, 2025).
	2022	Evidence-Based Reading Instruction Program Fund (SB 681)	In 2022, lawmakers created the Literacy Advisory Council and directed the Department of Elementary and Secondary Education (DESE) to develop a comprehensive statewide literacy plan grounded in evidence-based practices. The law required that struggling readers receive systematic interventions and that DESE publish a list of state-approved instructional materials that meets research-based criteria (Missouri General Assembly, 2025).
	2023	Missouri Revised Statutes § 161.241; § 167.645	In 2023, legislation prohibited the use of the “three-cueing” model in word reading instruction and appropriated funds to provide K – 5 educators with PD in SoR (Missouri General Assembly, 2025).
	2025	Ban on three-cueing	Beginning in 2025, new requirements for teacher preparation programs take effect,

			requiring candidates for K – 3 certification to complete at least six semester hours of coursework on the explicit, systematic teaching of language acquisition, phonological and phonemic awareness, phonics, vocabulary, fluency, and comprehension (Missouri General Assembly, 2025).
MT	N/A	N/A	Montana’s 2023 Early Literacy Targeted Intervention Act encourages districts to adopt research-based early literacy methods but does not mandate statewide structured literacy or SoR-aligned instruction, preparation, or PD.
NE	2018	Reading Improvement Act (Nebraska’s Revised Statutes § 79-2601 to 79-2607	Nebraska’s Reading Improvement Act requires all elementary schools to administer a state-approved reading assessment three times annually to identify students in K – 3 who may have reading difficulties. Students who are not meeting grade-level expectations must receive evidence-based supplemental reading interventions and an individual reading improvement plan. Parents must be notified if their child is identified as struggling in reading, and schools are required to provide progress updates (Nebraska Legislature, 2018).
NV	2015	Read by Grade 3 Act (SB 391)	In 2015, Nevada enacted legislation establishing a third-grade retention policy tied to reading proficiency and requiring all schools to develop literacy improvement plans. The law mandated that each elementary school designate a learning strategist responsible for supporting teachers in delivering interventions and required selected teachers to complete PD in reading instruction (Nevada Legislature, 2019).
	2019	Revisions to the Read by Grade 3 law (AB 289)	In 2019 lawmakers passed SB 391, repealing the mandatory third-grade retention provisions but expanded literacy supports to additional elementary grades. The update also strengthened school-level leadership by requiring every elementary school to designate a literacy specialist to oversee reading instruction and intervention. These provisions collectively emphasized evidence-based instruction, intervention, and schoolwide literacy planning (Nevada Legislature, 2019).

NH	N/A	N/A	NH does not currently have a statewide SoR law. While the state has policies focused on dyslexia screening and related supports, it has not adopted legislation requiring structured literacy or full alignment with SoR.
NJ	2024	Student Literacy Act (P. L. 2024, c.52)	Beginning in the 2025 – 2026 school year, districts must administer literacy screenings to students in grades K – 3 twice annually and provide results to parents and guardians. Schools are required to advise families about intervention options and support services if deficiencies are identified. The law also directs the DOE to design a PD program on foundational literacy for teachers in grades K – 6, which must be made available to districts free of charge. Each district must provide this PD (or comparable training) annually to all teaching staff in grades K – 6 (New Jersey Legislature, 2024).
NM	2019	Dyslexia Screening and Intervention Act / Intervention for Students Displaying Characteristics of Dyslexia	In 2019, New Mexico enacted legislation requiring that students who demonstrated reading or language difficulties receive evidence-based interventions. The law directed school districts to ensure literacy teachers and specialists are trained in structured literacy methods and to embed these practices into school literacy plans (NM Public Education Department, 2023).
	2023	Reading Materials Fund / Structured Literacy Fund (HB 481)	In 2023, the legislature established the Structured Literacy Fund, which provides grants to school districts and charter schools to support the adoption of evidence-based instructional materials and aligned professional learning. Together, these measures signal the state's move towards broader implementation of the SoR (New Mexico Public Education Department, 2023).
NY	N/A	N/A	New York does not currently have any statewide SoR laws.
NC	2021	Excellent Public Schools Act (SB 387)	The Excellent Public Schools Act of 2021 (SB 387) requires all K – 5 teachers to complete a professional learning course in the SoR and mandates that educator preparation programs train elementary teacher candidates in evidence-based literacy instruction. Districts must create

			intervention plans for struggling readers and ensure that their core curricula are aligned to state evidence-based standards (North Carolina General Assembly, 2021).
	2023	Budget Bill Update	In 2023, the state's budget law banned cueing strategies in reading curricula and instruction (North Carolina General Assembly, 2023).
ND	2021	Reading Competency, Teacher PD, & Licensure Act / Science of Reading Training Law (HB 1388)	In 2021, North Dakota enacted legislation requiring all teachers and principals in grades K – 3 to complete PD in scientifically based reading instruction. The law also requires educators in these grades to use scientifically based instructional materials and approaches in literacy instruction. This framework emphasizes alignment with SoR (North Dakota Department of Public Instruction, 2021).
OH	2023	State Budget Act (HB 33)	Ohio's 2023 state budget included a comprehensive literacy package grounded in the SoR. The law bans the use of three-cueing strategies in reading instruction, requires that interventions be aligned with SoR, and mandates that districts select literacy curricula from a state-approved list. The budget also provides funding for literacy coaches and for teacher PD in evidence-based reading instruction. In addition, educator preparation programs must ensure that preservice teachers complete coursework in SoR. These provisions together create a statewide framework for structured literacy implementation (Ohio Legislature, 2023).
OK	2019	Reading Instruction and Teacher Preparation Act (HB 2749)	Oklahoma requires that teacher preparation programs ensure that candidates study the five components of reading (phonemic awareness, phonics, fluency, vocabulary, and comprehension) and learn strategies for delivering explicit, systematic, and sequenced instruction. Within districts, reading instruction in grades K – 12 must be grounded in SoR and districts that receive certain state funds must provide PD in SoR for teachers in grades pre-K – 5. (Oklahoma Legislature, 2021).
	2021	Strong Readers Act (HB 1775)	Students in grades K – 3 must be assessed in foundational reading skills multiple times during the school year. The law also requires the state to publish and maintain a list of

			approved evidence-based literacy programs for districts to adopt (Oklahoma Legislature, 2021).
OR	2023	Early Literacy Success Initiative (HB 3198)	Oregon's 2023 legislation created a grant program to expand early literacy supports in school districts and charter schools. Funds may be used for research-aligned professional development, summer programs, tutoring, hiring literacy specialists, and purchasing instructional materials. To be eligible, districts must develop an Early Literacy Success Plan describing how they will implement evidence-based strategies and report on progress. The law emphasizes structured, research-aligned practices (Oregon Legislature, 2023).
PA	2021	State Board of Education Actions	In 2021, the Pennsylvania State BOE required that all teacher-preparation programs include coursework in structured literacy for certification in early childhood, elementary/middle, special education, English as a second language, and reading specialist programs. The Board also mandated that in-service teachers in these areas receive training in structured literacy, defined as systematic, explicit instruction that integrates listening, speaking, reading, writing, and spelling (Pennsylvania State Board of Education, 2021).
	2024	Evidence-Based Reading Instruction Act (SB 801)	In 2024, SB 801 expanded these requirements statewide by directing the DOE to create PD in structured literacy, requiring school districts to employ literacy coaches, and mandating that schools adopt high-quality instructional materials grounded in evidence-based approaches (Pennsylvania General Assembly, 2024)
RI	2019	Right to Read Act (HB 5887)	Rhode Island requires that elementary licensure candidates demonstrate proficiency in the best practice of scientific reading instruction, while candidates seeking certification in other grade levels must demonstrate an awareness of those practices. Additionally, districts must provide PD in science-based reading instruction for educators. The law establishes a statewide expectation that instruction be grounded in

			evidence-based literacy practices (Rhode Island General Assembly, 2019b).
SC	2014	Read to Succeed Act (Act 284, S.516)	South Carolina's Read to Succeed Act established a statewide literacy framework. The law requires that all teachers deliver evidence-based reading instruction and intervention and that students with reading difficulties are identified as early as possible, with parents notified of progress. It mandates both preservice and in-service teacher training to equip educators to help all students comprehend grade-level texts. It also establishes a third-grade retention policy with intensive interventions provided to those retained (South Carolina General Assembly, 2014).
SD	N/A	N/A	South Dakota does not currently have a statewide SoR law.
TN	2021	Literacy Success Act (SB 7003)	Tennessee's Literacy Success Act requires the DOE to develop professional learning for teachers in evidence-based approaches to reading instruction. Teacher preparation programs must include coursework in foundational reading skills, ensuring new educators are trained in structured literacy methods. Schools must provide explicit instruction in these skills, deliver interventions and supports for struggling readers, and administer a universal reading screener in grades K – 3. Collectively the law establishes a statewide framework for aligning curriculum, teacher training, and student supports with SoR (Tennessee General Assembly, 2021).
TX	2019	Reading Academies and Phonics Law (HB 3)	In 2019 Texas passed a law that required all elementary schools to adopt systematic phonics programs and mandated that K – 3 teachers attend state-provided reading academies focused on evidence-based literacy instruction. The law also introduced a teacher certification requirement, making the passage of a proficiency exam in evidence-based reading instruction mandatory certification for teacher candidates seeking certification (Texas Education Agency

	2023	Ban of Three-Cueing & Strengthened Foundational Reading Standards (HB 1605)	In 2023, Texas also banned the use of cueing strategies in state-approved instructional materials (Texas Education Agency, 2023b).
UT	2022	Early Literacy Outcomes Improvement (SB127)	Utah's SB 127 established a statewide framework for structured literacy. The law requires the State BOE to provide and train literacy coaches for K – 3, offer statewide professional learning in SoR, and approve a statewide benchmark assessment for early literacy. It also mandates that educator preparation programs train all preservice candidates in evidence-based reading instruction and that current teachers complete an approved PD program in SoR. The law further directed the state to assemble a literacy advisory board to consult with higher education and to create a bank of instructional materials to support districts (Utah State Legislature, 2022).
VT	N/A	N/A	Vermont does not currently have a statewide SoR law.
VA	2022	Virginia Literacy Act (HB 319; SB 616)	The Virginia Literacy Act of 2022 requires higher education institutions to prepare preservice teachers in evidence-based reading instruction and mandates that candidates pass a reading instruction certification exam for licensure. Local school boards must adopt districtwide literacy plans, employ reading specialists, administer assessments for reading difficulties, and provide interventions for students who are behind. Principals and certain teachers must also complete PD in evidence-based literacy practices (Virginia General Assembly, 2022b)
	2023	Virginia Literacy Act Expansion (SB 1175)	In 2023 the law was expanded to extend the state's structured literacy and intervention requirements from grades K – 5 to grades K – 8, broadening the scope of implementation (Virginia General Assembly, 2023).
WA	N/A	N/A	Washington does not currently have a statewide SoR law.
WV	2019	Initial Literacy Law	In 2019, West Virginia enacted legislation to strengthen early literacy instruction and professional learning in evidence-based reading. The law mandates that elementary teacher candidates complete at least nine credit

			hours of reading instruction focused on the five essential components of reading (West Virginia Legislature, 2023).
	2023	Third Grade Success Act (HB 3035)	This act expanded prior requirements by directing the DOE to implement a systematic approach to closing reading gaps and to create a list of approved screening tools for grades K – 3. The law requires training for all K – 3 educators in SoR and prohibits the use of instructional materials that rely on cueing methods (West Virginia Legislature, 2023).
WI	2023	Right to Read Act (Act 20 / AB 321)	Wisconsin's 2023 Act 20 establishes a statewide framework for literacy reform grounded in the SoR. The law prohibits the use of three-cueing in K – 3 instruction, requires districts to adopt curricula from a DPI-approved list, and mandates the creation of a state council to recommend science-based literacy practices. The state will hire up to 64 full-time literacy coaches, provide them with ongoing training, and support districts in implementation. Teacher-preparation programs must show they teach science-based early reading instruction and avoid three-cueing. To be eligible for K – 6 teaching licensure, candidates must have instruction in phonics. All K – 3 teachers are required to complete training (e.g. LETRS or an approved equivalent), and all K – 3 students must be assessed for reading readiness. Districts must develop personal reading plans for students identified as “at risk” (Wisconsin Legislature, 2023).
WY	N/A	N/A	Wyoming has statutes requiring K – 3 reading assessments and individualized reading plans, but does not yet mandate structured literacy, so it is classified as not having a SoR law.

Note. The table summarizes state legislation meeting the criteria for Science of Reading (SoR) alignment, including the year passed and key policy provisions.

Appendix Table A2. Coding of state SoR Legislation.

 Binary Column (SoR Law 0/1)

- ☐ 1 = Yes if the state has passed any legislation or statewide regulation that references the SoR. SoR is defined as grounded in the five essential components of reading: phonemic awareness, phonics, fluency, vocabulary, and comprehension
- ☐ 0 = No if the state has not passed any such legislation (even if other literacy policies exist, e.g. dyslexia screening only).

Number of Categories 0 – 6

- ☐ Count of how many of the following six categories are explicitly address in state law.

Categories

1. Teacher Preparation
 - a. Code “yes” if the law requires an exam, endorsement, or renewal tied to evidence-based literacy instruction.
 - b. Code “no” if only in-service teachers are affected or EPPs are not mentioned.
 2. Teacher Certification / Licensure
 - a. Code “yes” if the law requires an exam, endorsement, or renewal tied to evidence-based literacy instruction.
 - b. Code “no” if professional learning is mandated but not connected to licensure or certification.
 3. Professional Development / Coaching
 - a. Code “yes” if the law mandates PD, coaching, or state-provided training for current teachers.
 - b. Code “no” if PD is voluntary or left entirely to local discretion.
 4. Assessment
 - a. Code “yes” if the law requires universal screening, diagnostic assessment, or use of state-approved assessments for reading.
 - b. Code “no” if assessment is optional or not included.
 5. Materials
 - a. Code “yes” if the law requires adoption of state-approved curricula or restricts the use of certain approaches (e.g., bans of three-cueing, mandates phonics-based materials).
 - b. Code “no” if materials are not regulated by law.
 6. Instruction / Intervention
 - a. Code “yes” if the law requires:
 - i. Explicit/systematic instruction in foundational reading skills **or**
 - ii. Intervention supports such as individualized reading plans, summer school, MTSS structures, or tutoring.
 - b. Code “no” if instruction is only implied or left entirely to local discretion.
-

Note. The table presents the coding criteria used to classify state legislation as SoR-aligned and to code the six policy components used to construct the policy scope measure.