

Project Narrative - Table of Contents

Introduction and Priorities.....	2
A. Significance.....	3
B. Strategy to Scale.....	5
B.1. Unmet Demand.....	6
B.2. Feasibility of Management Plan.....	7
B.3. Relevance and Commitment of Partners.....	10
B.4. Quality of Plan to Deliver Services.....	13
B.5. Broad Dissemination.....	14
C. Quality of Project Design.....	15
C.1. Logic Model and Conceptual Framework.....	15
C.2. Project Goals, Objectives, and Outcomes.....	19
D. Project Evaluation.....	21
D.1. Methods to Generate Evidence That Meets WWC Standards Without Reservations.	23
D2. Guidance for Replication and Broader Implementation.....	28
D3. Measuring Fidelity of Implementation.....	28
D4. Replication-Ready Evidence and Lessons Learned.....	30
E. References for Parts A, B, and C.....	32

Project Title: Pathways to Evidence-Based Literacy: Embedding SRSD in Schools with AI-powered Support (“Pathways”)

Introduction and Priorities

Arizona Educational Foundation, in partnership with SRSD Online and American Institutes for Research, propose an Expansion Grant under the Education Innovation and Research program, administered by the Office of Elementary and Secondary Education of the U.S. Department of Education. The proposed project is titled “Pathways to Evidence-Based Literacy: Embedding SRSD in Schools with AI-powered Support” (hereinafter “Pathways”), and builds on the well-established evidence base for *Self-Regulated Strategy Development* [REDACTED]. SRSD is an instructional approach to writing that has repeatedly demonstrated significant impacts on students’ writing performance and self-regulation. Our project extends SRSD by developing and rigorously testing an innovative AI integration professional development (PD) course that supports coaches and teachers in embedding SRSD into their schools’ existing curricula, content, and schedules. Such integration of an evidence-based practice is one of the most persistent barriers to scale.

By utilizing SRSD as our focal treatment, we address **Absolute Priority 1 – Field-Initiated Innovations: Promoting Evidence-Based Literacy**. We also address **Competitive Preference Priority 1: Returning Education to the States**. As the lead nonprofit agency in the state of Arizona who directly collaborates and is contracted by the Arizona State Department of Education (AZDOE), Arizona Educational Foundation (AEF) will use our network of public schools across the state to develop a model and deliverables for statewide adoption of evidence-based literacy practice. The model will empower states to lead these initiatives.

Together, these commitments ensure the proposed project both fulfills the absolute requirement and provides additional strength under the competitive preference framework.

A. Significance

The Pathways project **addresses one of the most persistent challenges in literacy reform: the gap between proven, evidence-based practices and their sustained, scalable integration into schools.** SRSD is one of the most rigorously tested writing interventions, with over 100 peer-reviewed studies and multiple meta-analyses demonstrating large, positive effects on student writing, motivation, and self-regulation [REDACTED]. Further, according to the Institute for Education Sciences Practice Guide: *Teaching Elementary School Students to Be Effective Writers* [REDACTED], SRSD is supported by “strong evidence” of its effectiveness across student populations, including students performing below grade level in writing.

Despite this strong foundation, adoption of SRSD at scale has been slowed by a persistent integration barrier: educators often struggle to blend SRSD into pacing guides, daily schedules, and across content areas without sacrificing other curriculum commitments. **The proposed project is innovative** in that it directly addresses this “last mile” problem through the development and testing of an AI-supported professional development course for instructional coaches. This new course will provide systematic tools, exemplars, and AI-driven planning supports to help coaches and teachers embed SRSD into diverse instructional contexts with greater ease and fidelity – all based on the years of experience our project team brings to bear in providing professional development and supporting schools with SRSD integration.

The National Need: Writing Achievement Gaps

In addition to the challenges with adopting evidence-based practices like SRSD, this project addresses persistent writing performance gaps for American students. Despite decades of literacy reform, writing remains one of the most under-addressed areas of instruction in U.S. schools. Writing is critical for communication, learning, and future success (Freedman et al., 2016; Lee & Schmidgall, 2020). Yet, writing instruction is often deprioritized in favor of other tested subjects, leaving teachers underprepared and students underachieving. On the most recent National Assessment of Educational Progress (NAEP), nearly two-thirds of 8th and 12th graders scored at or below basic in writing (National Center for Education Statistics, 2012). Teachers themselves often report little to no training in writing pedagogy (Brindle et al., 2016; McKeown et al., 2014). These systemic gaps disproportionately affect students from marginalized backgrounds, widening educational inequities.

The Role of Practice-Based Professional Development (PBPD)

While SRSD has demonstrated strong effects in controlled research, scaling it in authentic school contexts requires robust professional learning models. Traditional professional development has yielded only modest impacts on teaching and student learning (Sims et al., 2023). By contrast, PBPD situates teacher learning in the authentic work of teaching—using classroom materials, focusing on real student needs, and providing opportunities for modeling, practice, feedback, and collaboration (Ball & Cohen, 1999; Darling-Hammond et al., 2017). PBPD has been shown to positively affect teacher practice and student writing outcomes when applied to strategy-based instruction [REDACTED].

Our model combines the strengths of PBPD with SRSD’s evidence base: teachers and coaches engage in authentic, practice experiences, supported by expert modeling and peer collaboration.

Innovative Contributions of Pathways

SRSD Online will lead PBPD efforts for SRSD using our flagship asynchronous online course, which covers the six stages of SRSD and includes lesson plans, videos, and other tools. In addition, we provide coaches with an additional course (asynchronous modules, plus synchronous Zoom sessions) to support them in providing PBPD for their teachers. This innovative use of technology allows for improved access to high-quality PD.

The proposed AI-supported integration course will extend PBPD by providing adaptive, context-sensitive tools to help coaches learn how to embed SRSD into local curricula, pacing guides, and standards. By conducting a large-scale, three-arm randomized controlled trial (RCT), this project will provide direct evidence of:

1. The effectiveness of SRSD compared to business-as-usual instruction (SRSD vs. waitlist control).
2. The incremental value of the AI integration course for coaches in improving fidelity and outcomes (SRSD + AI vs. SRSD only).
3. The combined effect of SRSD + AI compared to business-as-usual (SRSD + AI vs. waitlist).

This design ensures rigorous causal evidence, addresses a key implementation barrier, and produces actionable knowledge for SEAs and LEAs seeking to embed evidence-based literacy practices.

B. Strategy to Scale

Our scaling strategy is designed to demonstrate the effectiveness of AI-enhanced SRSD implementation while building sustainable infrastructure for continued expansion. The project employs a systematic, evidence-based approach that progresses from pilot testing through large-

scale randomized controlled trial to sustainability analysis and dissemination. This multi-phase strategy ensures rigorous evaluation of the innovation while creating conditions for sustained adoption and broader replication.

The scaling approach addresses five critical dimensions: (1) demonstrating unmet demand for effective writing instruction support, (2) implementing a feasible management plan with clear responsibilities and timelines, (3) securing meaningful partner commitment, (4) delivering services efficiently at scale while maintaining effectiveness, and (5) establishing robust dissemination mechanisms for broader impact.

B.1. Unmet demand for broader implementation

The demand for effective writing instruction support operates at multiple levels of the educational system, from individual teachers seeking practical strategies to district leaders looking for scalable solutions to persistent achievement gaps.

Teacher-Level Demand: Surveys and implementation studies consistently document teachers' desire for more effective writing instruction approaches. A comprehensive review of research on professional development for multicomponent strategy-focused writing instruction found that teachers frequently report needing more targeted support to implement evidence-based writing practices [REDACTED]. Teachers specifically request professional development that helps them integrate new approaches with existing curricula rather than treating them as separate additions (Hodges, 2022).

School-Level Demand: Principals and instructional leaders report significant pressure to improve writing outcomes, particularly for struggling students and those with disabilities. The persistent gaps in NAEP writing scores create accountability pressures that drive demand for proven solutions. Schools serving high-need populations face particular challenges, as they often

lack the resources and expertise needed to implement complex instructional innovations without external support (Cohen & Hill, 2008).

District-Level Demand: Superintendents and curriculum directors seek scalable approaches that can improve outcomes across multiple schools while maintaining implementation fidelity. The challenge of scaling evidence-based practices has been well-documented (e.g. Fixsen et al., 2009), with many districts struggling to move beyond pilot implementations to system-wide adoption. The AI-enhanced integration support addresses this challenge by providing tools that can support consistent implementation across diverse contexts.

State and National Demand: Education agencies at state and federal levels have invested significantly in literacy initiatives while acknowledging that writing instruction remains under-addressed. The Every Student Succeeds Act's emphasis on evidence-based practices creates demand for rigorously evaluated approaches like SRSD, while the integration challenge limits adoption. State education agencies report needing better tools and resources to support districts in implementing evidence-based writing instruction (Carver-Thomas & Darling-Hammond, 2017).

B.2. Feasibility of management plan

The proposed project will be administered by the Arizona Educational Foundation (AEF) as the prime applicant, in close collaboration with SRSD Online and the American Institutes for Research (AIR). Together, these organizations bring complementary expertise in program management, professional development, and independent evaluation.

- AEF will serve as fiscal agent, oversee school recruitment and state-level relationships, coordinate convenings, and lead dissemination efforts.

- SRSD Online will lead course development, deliver professional learning, coordinate communication with school partners, provide technical assistance (TA) and chair the advisory board.
- AIR will serve as the independent evaluator, finalizing the study design, randomization, and all data collection and analysis. They will also oversee fidelity monitoring, and manage scoring.

The project will be governed by an Executive Leadership Team comprising senior representatives from each partner organization. This team will meet regularly to review progress, address challenges, and make strategic decisions. The team includes:

- [REDACTED], Arizona Educational Foundation (Project Director)
- [REDACTED], SRSD Online (Principal Investigator)
- [REDACTED], American Institutes for Research (Evaluation Director)

Day-to-day project management will be coordinated through a Project Management Office (PMO) housed at Arizona Educational Foundation. The PMO maintains project schedules, coordinates partner activities, manages communications, and ensures deliverable quality. Exhibit 1 shows key activities and responsibilities by grant year.

Exhibit 1. Activities and Responsibilities

Year	Activities & Milestones	Partner(s) Responsible
Year 1 (2026)	Develop and pilot AI integration course (5 schools, 10 coaches); recruit 75 additional schools; finalize evaluation design and IRB approval.	SRSD Online (development/pilot), AEF (recruitment), AIR (evaluation prep).

Year	Activities & Milestones	Partner(s) Responsible
Year 2 (2027)	Randomize 75 schools into 3 groups (SRSD + AI, SRSD only, waitlist control); deliver SRSD foundational training; deliver AI integration course to treatment coaches; baseline data collection.	AEF (recruitment logistics), SRSD Online (training), AIR (randomization, baseline).
Year 3 (2028)	Implement SRSD across treatment schools; fidelity monitoring; midline data collection (student writing, surveys, observations); interim findings prepared.	SRSD Online (TA, fidelity), AIR (data collection, interim analysis).
Year 4 (2029)	Train waitlist control group; continue implementation; sustainability analyses initiated.	SRSD Online (training, TA), AEF (support), AIR (analysis).
Year 5 (2030)	All schools continue SRSD; AIR completes outcome and cost-effectiveness analyses; dissemination of findings (policy briefs, guides, presentations, convenings).	AIR (final report), SRSD Online (products, dissemination support), AEF (policy briefs, convenings).

Finally, a Project Advisory Board provides external guidance and expertise. The board will include representatives from participating districts, state education agencies, and national experts in writing instruction and AI for education. The board will meet quarterly to review progress and provide strategic guidance as needed. We have included **letters of support from several key members in Appendix C:**

- [REDACTED] (SRSD Creator, renowned writing researcher)
- [REDACTED] (Renowned writing researcher)
- [REDACTED] (EIR expansion project director for Scaling SRSD project)
- [REDACTED] (AI education expert)

Notably, project team members have considered potential risks to the project’s success, and have determined mitigation strategies. First, recruitment: the project's success depends on recruiting and retaining 75 schools across the 5-year period. Mitigation strategies include: building on established relationships through Arizona Educational Foundation's networks, providing

meaningful incentives for participation (professional development, materials, technology access), maintaining flexible scheduling and support options to accommodate diverse district needs.

Second, the AI integration course represents an innovative technology application that carries implementation risks. Mitigation strategies include: extensive pilot testing to identify and resolve technical issues before RCT launch, maintaining backup systems and alternative delivery methods, providing comprehensive technical support for users.

Third, multi-partner projects face coordination challenges that can affect timeline and quality. Mitigation strategies include: clear partnership agreements defining roles, responsibilities, and deliverables; regular communication protocols and coordination meetings; shared project management systems and documentation; conflict resolution procedures and escalation processes.

B.3. Relevance and Commitment of Partners

This project is anchored by a strong partnership among the Arizona Educational Foundation (AEF), SRSD Online, and the American Institutes for Research (AIR). Together, these organizations represent the complementary expertise required for both successful implementation and rigorous evaluation of an Education Innovation and Research Expansion grant. AEF provides deep state-level connections and experience in school recruitment, communications, and dissemination, serving as the fiscal agent and convener of partners. SRSD Online brings expertise in SRSD professional development and practice-based implementation, with a proven record of scaling evidence-based practices in classrooms nationwide. AIR contributes methodological expertise and evaluation capacity, ensuring the project's findings will meet the highest evidence standards and inform national policy and practice. Collectively, the partners are committed not only to executing the proposed project with fidelity and precision, but

also to advancing the broader goal of embedding evidence-based literacy practices in schools at scale.

Arizona Educational Foundation (AEF): AEF brings decades of experience to bear in terms of managing grants in support of education efforts, and in playing key roles alongside and for AZDOE. AEF and AZDOE maintain a collaborative partnership rooted in their shared mission to celebrate excellence and strengthen public education across the state. AEF works closely with AZDOE in a variety of ways. For example, AZDOE supports AEF’s *A+ School of Excellence*® program by helping to disseminate application information to districts statewide and allowing ADE employees to serve on adjudication teams for school site visits. Similarly, through initiatives like the *Principal Leadership Academy of Arizona* and *teachSTEM*, AEF aligns its professional development programs with ADE’s goals for educator retention and innovation—ensuring that both organizations’ efforts reinforce statewide improvement priorities. AEF’s Teacher of the Year and A+ Schools events, for instance, often include ADE representation, reflecting a mutual commitment to elevating Arizona’s education narrative.

Beyond programmatic alignment, AZDOE has contracted with AEF to help draft the AI Guidance for Arizona K-12 Schools and the Career Literacy Standards for K-8. AEF staff serve on AZDOE committees and task forces related to workforce development, Career and Technical Education, teacher retention, and principal effectiveness. This strong public–nonprofit partnership allows AEF to serve as an agile, community-connected complement to AZDOE’s statewide mandate—amplifying impact through innovation, storytelling, and recognition that inspire educators and strengthen confidence in Arizona’s public schools.

SRSD Online: SRSD Online is a leading provider of comprehensive professional development for SRSD. Led by [REDACTED], who will serve as Principal Investigator for

this project, SRSD Online has successfully scaled training across multiple states, districts, and schools through a blended model that combines asynchronous courses, synchronous coaching, and practice-based professional development (PBPD). The organization’s expertise is grounded in nearly four decades of SRSD research and is strengthened by a team of leading scholars and practitioners, including faculty and consultants with deep experience in literacy, special education, and instructional design.

SRSD Online brings extensive experience managing large-scale research–practice partnerships, including its role as professional development provider for an ongoing EIR expansion grant that trains teachers across multiple states. Through this work, SRSD Online has established systems for fidelity monitoring, implementation coaching, and integration of writing into broader literacy initiatives. Our proprietary suite of courses is designed for coaches, teachers, and leaders across K–12. By co-developing the AI integration course, SRSD Online will extend this track record of innovation—responding directly to the challenge of embedding evidence-based practices within competing instructional demands. The organization’s dual capacity as a research-informed provider and an agile business uniquely positions it to deliver scalable, sustainable, and high-quality training at the heart of this project.

AIR: The American Institutes for Research (AIR) is one of the world’s largest behavioral and social science research organizations, with a longstanding reputation for producing rigorous, policy-relevant education research. AIR brings deep expertise in the design and implementation of large-scale randomized controlled trials, longitudinal studies, and multi-method evaluations. Its teams include leading experts in literacy, teacher professional development, and education technology, supported by robust infrastructure for large-sample data collection, analysis, and dissemination.

AIR’s track record includes dozens of federally funded studies that meet What Works Clearinghouse standards without reservations, across areas such as literacy, STEM, and educator workforce development. As the independent evaluator for this project, AIR will finalize the study design, secure IRB approval, conduct school randomization, and lead all aspects of data collection and analysis—including scoring and analysis of student writing outcomes, teacher and coach surveys, classroom observations, and cost-effectiveness analyses.

Beyond methodological rigor, AIR is committed to producing actionable knowledge for policymakers and practitioners. Their dissemination strategy includes technical reports, policy briefs, and practitioner-friendly summaries, ensuring findings contribute to the national evidence base while informing state and local decision-making. With its reputation for independence, rigor, and practical relevance, AIR’s involvement bolsters the project’s credibility and ensures that results will meaningfully influence research, policy, and practice.

B.4. Quality of Plan to Deliver Services

The project’s plan to deliver services is designed to be both **scalable and sustainable**, leveraging SRSD Online’s professional development infrastructure and AEF’s statewide networks. SRSD Online’s suite of asynchronous and synchronous courses—including modules for teachers, coaches, and leaders—has already demonstrated the ability to train large numbers of educators in cost-effective, flexible formats. (See Appendix J12 for scope and sequence documents, course module screenshots, and a sample lesson plan.) These courses, hosted on a robust online platform, can be accessed by participants across diverse geographies, ensuring equity of access and minimizing barriers related to time, location, or school resources.

The proposed project extends this infrastructure by developing an **AI integration PD course** that specifically addresses one of the most persistent barriers to scaling evidence-based

practices: embedding them into existing curricula, pacing guides, and daily schedules. This add-on course will provide practical, context-specific strategies, supported by AI tools, that help coaches and teachers map SRSD into their school’s instructional framework. By situating this within the same digital platform as SRSD Online’s existing PD, delivery of services is streamlined, allowing participants to move seamlessly from foundational training into specialized integration supports.

To ensure fidelity and quality, services will be delivered through a combination of **asynchronous modules, synchronous coaching calls, practice-based professional development (PBPD) cycles in accordance with our existing PD scope and sequence, plus ongoing technical assistance**. SRSD Online will monitor engagement and progress within the course platform, provide feedback to coaches through PBPD structures, and ensure consistent pacing across treatment schools. AEF will support delivery by managing communications with districts, coordinating logistics, and ensuring alignment with state-level priorities. The design allows services to be delivered efficiently at scale while maintaining effectiveness, providing a replicable model for other states and districts.

B.5. Broad Dissemination

The proposed project includes a comprehensive plan for **broad dissemination of products, findings, and lessons learned**. Consistent with EIR requirements, all grant-funded deliverables will be made publicly available at no cost, ensuring that the knowledge generated through this project can be adapted and replicated nationally. Anticipated deliverables include:

- The **AI integration professional development course** (open-access version).
- An **integration toolkit** with templates, pacing guides, and exemplars for embedding SRSD into existing instructional schedules.

- **Fidelity and pacing exemplars** documenting successful implementation models across diverse school contexts.
- **Policy briefs** highlighting key findings, implications for literacy policy, and recommendations for state and district leaders.
- Practitioner-friendly summaries of evaluation findings developed by AIR.

Dissemination will occur through multiple, complementary channels. AEF will use its **statewide network in Arizona**, including relationships with the Arizona Department of Education, principals' associations, and teacher leadership networks, to share products and findings directly with schools and districts. SRSD Online will use its **national PD platform**, newsletters, webinars, and conference presentations (e.g., International Literacy Association, Council for Exceptional Children, American Educational Research Association) to reach educators, researchers, and policymakers beyond Arizona. AIR will contribute through **peer-reviewed publications, technical reports, and national briefings**, ensuring rigorous evidence reaches both academic and policy audiences.

Finally, the project team will host **annual convenings** to showcase lessons learned, highlight practitioner voices, and discuss implications for scaling evidence-based literacy practices. These convenings will invite state leaders, district administrators, and practitioners, positioning the project as a national resource for embedding SRSD and similar evidence-based models into schools at scale.

C. Quality of the Project Design

C.1 Logic Model and Conceptual Framework

The project's logic model is grounded in a comprehensive theoretical framework that integrates research on effective instruction, professional development, technology-enhanced

learning, and organizational change. This multi-theoretical approach ensures that the intervention addresses the complex, interconnected factors that influence educational innovation and implementation.

Self-Regulated Strategy Development (SRSD)

The instructional foundation rests on SRSD's integration of sociocultural and cognitive theories, and theories of self-control, metacognition, and strategy instruction [REDACTED]. This theoretical triangulation creates a robust framework for supporting both student learning and teacher development. The six-stage SRSD model (develop background knowledge, discuss it, model it, memorize it, support it, independent performance) provides a systematic yet flexible approach to instruction that can be adapted across diverse contexts while maintaining fidelity to core principles.

Practice-Based Professional Development (PBPD)

The professional development approach draws on situated learning theory and research on effective teacher learning. PBPD emphasizes learning in authentic classroom contexts, emphasizing the importance of analyzing student work, examining teaching practice, and connecting new learning to existing knowledge and beliefs [REDACTED]. This approach recognizes that sustainable change requires deep engagement with practice rather than superficial exposure to new ideas.

Technology Acceptance and Integration Theory

The AI enhancement component will be informed by the Technology Acceptance Model (TAM) and the Substitution, Augmentation, Modification, Redefinition (SAMR) frameworks. These theories emphasize that successful technology integration requires attention to perceived usefulness, ease of use, and the degree to which technology transforms rather than merely

substitutes for existing practices (Davis, 1989). The AI integration course is designed to move beyond simple substitution to transformation of professional learning experiences.

Organizational Change Theory

The scaling strategy incorporates principles from diffusion of innovation theory and organizational learning theory. These frameworks emphasize the importance of early adopters, social networks, organizational culture, and systematic support for change (Rogers, 2003). The focus on instructional coaches as change agents reflects research showing that coaches can serve as critical bridges between external innovations and internal organizational contexts (Coburn & Russel, 2008).

Logic Model (see Appendix G)

The model delineates how foundational training in Self-Regulated Strategy Development (SRSD), augmented by a new AI-powered integration course for coaches, leads to both immediate changes in educator knowledge and practice and long-term systemic improvements in literacy outcomes and sustainability.

Inputs and Core Activities. The foundational SRSD courses for teachers and coaches, unit plans, and synchronous practice-based professional development (PBPD) serve as the project's baseline inputs. These resources provide a consistent, evidence-based framework for writing instruction and have demonstrated effectiveness in prior large-scale randomized controlled trials. To address a persistent barrier to scale—difficulty integrating SRSD into existing curricula and schedules—the project introduces a new AI-powered integration course for coaches. This targeted intervention equips coaches to use AI tools to generate prompts, exemplar essays, and differentiated supports aligned with local standards and curricula.

Mediators and Short-Term Outcomes. These inputs are designed to shift critical mediators at the coach and teacher levels. Coaches will increase their pedagogical knowledge, self-efficacy, and capacity to integrate SRSD using AI supports. Teachers, in turn, will experience fewer barriers related to time, pacing, and curricular fit, enabling more faithful and sustained delivery of SRSD writing instruction. Early student outcomes include increased access to explicit strategy instruction in writing, embedded across subject areas, which enhances engagement and confidence.

Medium-Term Outcomes. As these mediators strengthen, schools in the treatment conditions will demonstrate higher fidelity of SRSD implementation compared to schools receiving SRSD-only training or those in the waitlist control group. Coaches and teachers will develop broader capacity to embed SRSD into science, social studies, and other content areas, supporting cross-disciplinary literacy. This is expected to yield measurable improvements in student writing performance, particularly for high-need subgroups, while simultaneously generating cost-effectiveness evidence for district and state leaders.

Long-Term Outcomes and Impacts. The long-term outcomes emphasize sustainability and scalability at both the school and system levels. Schools will continue to implement SRSD with strong fidelity beyond the study period, aided by the availability of the integration course, exemplars, and toolkits. States and districts will gain access to tested models for embedding evidence-based writing instruction within local curricula, supported by policy briefs and open-access resources. Collectively, these outcomes contribute to narrowing literacy achievement gaps, strengthening state capacity for literacy-focused professional development, and advancing

the national evidence base on practice-based PD, AI integration, and scaling evidence-based literacy practices.

Contextual Moderators. The model also accounts for contextual moderators—such as grade level, student demographics, and teacher experience—that may shape implementation. Including these variables strengthens the evaluation design by ensuring that findings are generalizable across diverse educational contexts and provide actionable guidance for replication.

C.2. Goals, Objectives, and Measurable Outcomes

The project’s goals, objectives, and outcomes are ambitious yet achievable within the grant period and align directly with the purposes of the EIR Expansion program. These are laid out in Exhibit 2 below.

Exhibit 2. Project Goals, Objectives, Measures and Outcomes by Year

Goal	Objective (Year)	Performance Measure	Target Outcome
1. Develop and refine supports for integrating SRSD into schools	Obj. 1 (Year 1): Develop and pilot the AI integration course	Number of pilot participants completing course	10 coaches, 50 teachers
		Percentage reporting increased confidence in embedding SRSD	≥ 80%
2. Rigorously test SRSD with and without AI supports	Obj. 2 (Years 2–3): Conduct RCT with 3 groups (SRSD+AI, SRSD-only, waitlist control)	Number of schools randomized	75 schools (≈150 coaches, 750 teachers)
		Course completion rates for treatment groups	≥ 90%
		Fidelity of SRSD implementation	≥ 75% of teachers meeting fidelity threshold
		Student writing outcomes	Statistically significant gains;

Goal	Objective (Year)	Performance Measure	Target Outcome
			effect size ≥ 0.25 SD
3. Document sustainability and integration practices	Obj. 3 (Year 3): Follow-up data collection	Response rate to follow-up surveys	$\geq 80\%$
		Sustained implementation in SRSD-only schools	$\geq 60\%$
		Sustained integration in SRSD+AI schools	$\geq 60\%$
4. Expand access and disseminate results broadly	Obj. 4 (Years 4–5): Train waitlist group and scale dissemination	Number of waitlist participants completing training	50 coaches, 250 teachers
		Fidelity of waitlist-trained teachers	$\geq 70\%$
		Dissemination reach (open-access course, toolkits, briefs, convenings)	$\geq 5,000$ stakeholders

All objectives directly support the EIR program's emphasis on developing, implementing, replicating, taking to scale, and rigorous evaluation of innovative and evidence-based practices, programs, or strategies. The project's focus on scaling SRSD through AI-enhanced professional development addresses the program's priority on innovations that can be widely adopted and sustained.

The measurable outcomes align with GPRA (Government Performance and Results Act) requirements and provide clear benchmarks for assessing project success. The combination of implementation metrics (training completion, fidelity scores) and impact measures (student achievement, sustainability indicators) ensures comprehensive evaluation of both process and outcomes.

The ambitious yet achievable targets reflect the project's commitment to generating meaningful evidence while acknowledging the challenges inherent in large-scale educational

innovation. The specific timelines and success criteria provide clear accountability measures while maintaining flexibility for adaptive implementation based on emerging evidence and stakeholder feedback.

D. Project Evaluation

AIR will conduct a rigorous external evaluation to provide the Arizona Educational Foundation (AEF) and SRSD Online with both formative and summative evidence regarding the implementation and impact of SRSD and the SRSD AI-supported integration professional development (PD). The implementation evaluation will examine whether the key components of SRSD PD and AI-supported integration PD are delivered as intended—and explore the reasons behind any deviations. It will supply project partners with ongoing feedback to guide the refinement of SRSD and its AI-supported integration PD, as well as insights into barriers and effective supports for scaling SRSD.

The impact study is designed not only to replicate previous research on SRSD’s effectiveness in improving teacher and student outcomes, but also to assess the added value of a new AI-supported integration PD in enhancing teachers’ implementation of SRSD at scale. Our comprehensive evaluation plan is structured to generate evidence of SRSD’s effectiveness that meets the **What Works Clearinghouse (WWC) evidence standards without reservations**. The evaluation will be conducted in 80 elementary schools across Arizona, focusing on teachers and students in Grades K–5. During the 2026–27 school year, five schools will participate in a pilot study. In the following year (2027–28), the remaining 75 schools will take part in a cluster randomized controlled trial (RCT). SRSD Online and AEF will provide ongoing support to all participating schools to sustain SRSD implementation into subsequent school years. Throughout 2027–28 and 2028–29, AIR will collect data to inform strategies for scaling SRSD and ensuring

its long-term sustainability. Exhibit 3 summarizes the evaluation’s research questions, data sources, and measures.

Exhibit 3. The evaluation research questions, data sources, and measures

Research questions	Data sources and measures
Program impact	
<p>RQ1. What are the impacts of SRSD on students’ writing quality, literacy achievement, academic dispositions, and writing behavior, compared to business as usual? (confirmatory)</p> <p>1a. Does adding AI-supported integration PD lead to greater improvements in outcomes compared to SRSD PD-only? (confirmatory)</p>	<p>Study-administered writing assessment (a holistic measure of student’s writing quality)</p> <p>District student records (state and district ELA test scores)</p> <p>Student survey, Grades 3–5 (attitudes toward writing, self-efficacy for writing, self-regulated learning, approaches to writing)</p>
<p>RQ2. What are the impacts of SRSD on teacher self-efficacy and writing instruction, compared to business as usual? (confirmatory)</p> <p>2a. Does adding AI-supported integration PD lead to greater improvements in teacher self-efficacy and writing instruction, compared to SRSD PD-only? (confirmatory)</p> <p>2b. Does adding AI-supported integration PD improve teachers’ fidelity of SRSD instruction in classrooms compared to SRSD PD-only? (confirmatory)</p>	<p>Teacher survey (efficacy for teaching writing, self-reported writing instructional practices)</p> <p>Classroom video observations (quality of writing instruction, fidelity of SRSD instruction)</p>
<p>RQ3. To what extent are the impacts of SRSD on student outcomes moderated by student and school characteristics? (exploratory)</p>	<p>Same as RQ1</p>
<p>RQ4. To what extent are the impacts of SRSD on student outcomes mediated by teachers’ practice and their fidelity of SRSD instruction? (exploratory)</p>	<p>Same as RQs 1 and 2</p>
Program implementation	
<p>RQ5. To what extent was SRSD implemented with fidelity in the SRSD AI-supported integration group and the SRSD PD-only group?</p>	<p>SRSD Online program data, including PD attendance, course products, coach logs</p> <p>Teacher and coach survey</p> <p>Classroom observations</p>
<p>RQ6. How do teachers and coaches perceive the feasibility and sustainability of SRSD implementation in both treatment groups, and how do these perceptions differ over time?</p>	<p>Teacher and coach survey</p> <p>Teacher focus groups</p> <p>Coach interviews</p> <p>Site visits</p>
<p>RQ7. What obstacles and opportunities do teachers and coaches perceive regarding scaling and sustaining SRSD, and how do these perceptions differ between groups and over time?</p>	<p>Teacher and coach survey</p> <p>Teacher focus groups</p> <p>Coach interviews</p> <p>Site visits</p>
Cost-effectiveness	

RQ8. What is the cost-effectiveness of scaling SRSD?

Teacher and coach survey
Teacher focus groups
Coach interviews

D1. Methods to Generate Evidence That Meets WWC Standards Without Reservations

Impact Evaluation Design. The impact evaluation will employ a three-armed cluster RCT design to rigorously assess the effects of SRSD on teacher and student outcomes, ensuring the evaluation will generate evidence that **will meet the WWC standards without reservations**. The impact study will involve 75 schools, randomly assigned to one of three groups (arms) in spring 2027: those receiving SRSD PD-only in 2027–28; those receiving SRSD PD and AI-supported integration PD in 2027–28; and those continuing with business as usual in 2027–28 (the control group). Control schools and SRSD-only schools will receive AI-supported integration PD in 2028–29, after the completion of the evaluation data collection. Randomization will be conducted within blocks, with blocks being districts or groups of schools with similar characteristics. Confirmatory analyses will focus on the impact of SRSD on student and teacher outcomes after one year of SRSD implementation, while additional exploratory moderator and mediation analysis will help identify for whom SRSD works, and under what conditions. See Appendix J1 for a detailed evaluation timeline.

Outcome Measures and Data Sources. The impact study will examine teacher and student outcomes using validated and reliable measures that are directly related to the intended program outcomes and meet the WWC’s validity and reliability requirements (WWC, 2022).

Student outcomes (RQ1). The primary measures of *literacy achievement* are students’ scores on state assessments in ELA (Grades 3–5) and district benchmark tests (Grades K–2). Student scores will be standardized by grade and test. Additionally, writing samples will be gathered for students in Grades K–5 in fall 2027 and spring 2028 to measure *writing quality*. Students will be

asked to write an expository essay using source texts (informational passage(s) appropriate to grade level). We will holistically score writing quality using an existing rubric that has demonstrated 90% inter-rater agreement (McKeown et al., 2019). Raters will be trained extensively in the rubric using anchor papers and will need to pass a reliability test before scoring for the study. We will explore using automated essay scoring (AES) scores (Tate et al., 2024), with the expectation that AES scores match human raters exactly 60% of the time and within one score point 90% of the time. Students' academic dispositions and writing behavior will be measured using a short survey administered online to students in Grades 3–5 in fall and spring of the intervention year. The survey will include measures of *attitudes toward writing* [REDACTED], *self-efficacy for writing* [REDACTED], *self-regulated learning* (Usher & Pajares, 2008; $\alpha = 0.83$), and *approaches to writing* (Lavelle, Smith, & O’Ryan, 2002; [REDACTED]; $\alpha = 0.96$). Baseline measures will be prior-year scores on state assessments or district benchmark tests, scores from writing samples gathered and scored in the fall, and survey scale score from the fall.

Teacher outcomes (RQ2). *Teacher self-efficacy* will be measured using a teacher survey. The Teacher Self-Efficacy for Teaching Writing measure [REDACTED] will be administered in spring of 2027 as a baseline measure and in spring 2028 as an outcome measure to all participating teachers in grades K–5 in the 75 study schools. The survey also will include items asking teachers to report their *writing instruction practice*, such as how often they taught specific aspects of writing [REDACTED]. To measure the *quality of teachers' writing instruction*, we will randomly select four teachers per school for observations. For each teacher, we will video-record two lessons using IRIS Connect technology in early fall of the intervention year (as baseline) and two lessons in spring (as outcomes). The video-

recorded lessons also will be coded by trained observers, who will be blind to treatment status, using a protocol that was developed and is currently being tested in another large-scale study of SRSD (See Appendix J2). The protocol draws from recommendations from the WWC practice guide, *Teaching Elementary School Students to Be Effective Writers* [REDACTED], and focuses on observable aspects of writing instruction, such as opportunities for collaborative writing, peer editing, and explicit goal setting. A subset (20%) of lessons will be double-coded by independent coders to check reliability. After initial scoring for all videos is complete, for both treatment groups, observers will score the video-recorded lessons using a fidelity of SRSD instruction checklist, which was developed by SRSD experts and measures the fidelity of teachers' implementation of SRSD lessons in classrooms (See Appendix J3).

The same student and teacher outcome measures as those described previously will be included in the differential impact (moderator) analysis for addressing **RQ3**, and the mediator analysis for addressing **RQ4**. Additional information on the student and teacher outcome measures can be found in Appendix J4, along with a specified strategy for dealing with multiple comparisons in Appendix J5.

Sample and Statistical Power. The sample for the proposed RCT will include 75 schools and 900 teachers, assuming 12 elementary school teachers per school (i.e., two per grade). To limit the risk of bias due to joiners (individuals who enter the cluster after the time of random assignment), teachers who were joiners and their students will not be included in the analytic samples. For student outcomes, the sample will consist of students taught by teachers participating in the study. Assuming an average of 20 students per teacher (i.e., 240 students per school), the anticipated sample sizes will provide an estimated minimum detectable effect size (MDES) of 0.12 for writing and ELA achievement, 0.18 for SEL outcomes, 0.35 for classroom

instruction, and 0.27 for teacher self-efficacy for our main impact analysis (SRSD PD and SRSD AI-supported integration PD vs. control). The estimated MDESs for student outcomes are consistently smaller than the effect sizes reported in previous studies of SRSD that contributed to the rating of strong evidence for recommendations in the WWC practice guide, *Teaching Elementary School Students to Be Effective Writers* [REDACTED], indicating the study will have sufficient power to detect effect sizes that are of statistical and practical significance. See Appendix J6 for MDESs for all comparisons and assumptions for each power calculation.

Strategies for Dealing with Attrition. For the proposed RCT, the main threat to internal validity is potential selection bias resulting from sample attrition during the study. AIR assumed 20% attrition at the school level in power calculation. We expect the actual school-level attrition to be lower than 20%. First, the delayed treatment design will allow control schools to receive the SRSD AI-supported integration PD after the experiment year, therefore providing motivation for remaining in the study and reducing differential attrition. Second, AIR will collect student administrative data at the district level so that schools that have dropped out of the study could still be included in the intent-to-treat analysis for student achievement outcomes. Third, to reduce the risk of attrition, AIR will communicate the data collection requirements and the importance of retention to potential participating schools to ensure an understanding about the effect of participation on the evaluation. Student-level attrition due to missing administrative data is expected to be relatively low (20% assumed in power calculation) for student achievement outcomes because data for these outcomes are based on district administrative records. For other outcome data (writing assessment, student surveys, teacher surveys, observations), AIR's goal is to achieve an 80% response or completion rate. AIR has demonstrated success in achieving this

goal in other similar projects, and has established procedures to help achieve the target response rate. For example, AIR attained a 93% survey response rate in the New York City Single Shepherd evaluation for 2 years in a row.

Impact Analysis. Our main impact analyses will be fixed-effects intent-to-treat analyses that estimate the impact of SRSD on students (**RQs 1 and 1a**) and teachers (**RQs 2, 2a, and 2b**) in the study schools. The study will use hierarchical linear modeling to account for the nesting of students and teachers within schools. The impact models will incorporate fixed effects for randomization blocks and, as appropriate, covariates at the student, teacher, and school level that are expected to be correlated with the outcomes. Inclusion of covariates will improve the precision of the impact estimates and guard against any bias due to imbalance in baseline covariates that arises due to random chance. We will adapt the main student impact model to assess the differential impact of SRSD (**RQ3**) by incorporating a treatment-by-moderator interaction term, where the moderator is a student or school characteristic of interest (e.g., student grade level, school size). A significant interaction term would indicate that the estimated impact varies by the student or school characteristics of interest and allows for estimation of a separate effect for each subgroup. To address **RQ4**, we will conduct exploratory mediation analysis to assess whether the impact of SRSD on student outcomes is mediated by the quality of classroom instruction and SRSD implementation fidelity. Appendix J7 provides more information on the analytic models and covariates; see Appendix J8 for approaches to handling missing data.

Because the impact evaluation is based on a school-level RCT free of confounding factors and is expected to have low attrition at all levels, it will produce strong evidence about SRSD's

effectiveness at scale that will meet the **What Works Clearinghouse (WWC) group design standards without reservations.**

D2. Guidance for Replication and Broader Implementation

The findings will inform strategies to scale SRSD and identify the settings and populations where it has the greatest impact. The three-arm RCT enables us to assess the replication of SRSD's effects in Arizona and the added value of an AI-supported approach to scaling. In addition, as articulated in **RQs 5–7**, we will collect comprehensive implementation data, including site visits at schools demonstrating the highest levels of SRSD implementation and student achievement in literacy. In addition, we will conduct a video analysis that will examine how teacher instruction mediates student outcomes. Collectively, the evaluation's rich data will inform guidance about effective strategies suitable for replication and their implementation.

The evaluation will involve a large and representative sample of students from 80 elementary schools (five pilot and 75 RCT schools) across Arizona, representing a range of geographic regions, school sizes, student demographics (e.g., grade level, economic disadvantage, special needs, English learners), and teacher characteristics (e.g., years of experience, certification status). To guide implementation in other settings, we will test for differential effects across these characteristics. This type of information will provide guidance on the contexts and populations for which SRSD is most effective. Coupled with the rich implementation data, these insights will inform tailored strategies that align with specific student needs, school settings, and instructional contexts.

D3. Measuring Fidelity of Implementation

Implementation Study. The evaluation will establish clear thresholds for acceptable implementation of SRSD and the AI-supported integration PD as a scaling strategy to assess the

fidelity of implementation (**RQ5**). First, we will develop an implementation rubric (and refine during the pilot) by collaborating with SRSD Online to identify quantifiable indicators for the key activities in the logic model and establish clear thresholds for acceptable implementation. See Appendix J9 for an example fidelity matrix that specifies examples of fidelity indicators, how we will measure each, and what thresholds we will use for determining fidelity at each level (i.e., teacher/coach, school, and sample level). As an example, implementation fidelity of the coach participation in the AI-supported integration PD will be measured by the quality of the AI-supported products that coaches create, demonstrating SRSD alignment with district curricular resources. The threshold for full implementation specifies that a school-level coach must submit at least one AI-supported product and that at least one of those products must be of high quality according to a rubric (see Appendix J10). These indicators and thresholds will ensure a shared understanding of program components and expectations for the quality of implementation. During the pilot, AIR will test and finalize these proposed implementation thresholds. During the RCT, we will generate school-level implementation statistics based on the implementation rubric, which will be included in models to inform how implementation is associated with outcomes.

To generate insights on how to scale and sustain SRSD efficiently and effectively, we will collect teacher surveys and qualitative data. During the pilot, we will conduct two focus groups with 10 teachers and interviews with five coaches. During the RCT, we will survey teachers about implementation and conduct 10 focus groups with 50 teachers (10 teachers per grade level) and interviews with a stratified random sample of 25 coaches from both treatment groups. After the RCT, to explore sustained implementation and its association with outcomes, we will conduct four site visits each year (eight site visits total) at the schools with the highest levels of

SRSD implementation and student achievement in literacy. These visits will include observations of staff and PLC meetings, collection of pacing guides and lesson plans, focus groups with teachers, and interviews with coaches and administrators. We will analyze the qualitative data to identify factors that support or hinder sustained implementation and effective strategies for affecting long-term student outcomes.

D4. Replication-Ready Evidence and Lessons Learned

This evaluation will guide future replication and adaptation in other contexts and help the field understand how to apply SRSD and the AI-supported integration to similar scaling efforts. We will do this by implementing a rigorous, mixed-methods RCT evaluation design, followed by a descriptive study focused on sustainability. The RCT is designed to meet **WWC standards without reservations**, ensuring the generation of valid and reliable evidence on the effectiveness of both SRSD and the SRSD AI-supported integration. To guide replication of the program, the evaluation data will inform continuous improvement of the SRSD Online teacher manuals.

Cost Analysis. To assess cost-effectiveness (**RQ8**), we will conduct a comprehensive cost analysis using the Resource Cost Model (RCM), a method widely used by AIR. The analysis will be conducted using the *CostOut* tool developed by the Center for Benefit-Cost Studies in Education at Teachers College, Columbia University. This tool leverages administrative data on unit costs for labor and non-personnel resources, along with data collected on the teacher survey. These cost estimates will be paired with impact findings to generate cost-effectiveness ratios, providing actionable insights for decisionmakers considering replication or adaptation in other settings.

Together, the evaluation design, implementation documentation, and cost-effectiveness analysis will offer a robust foundation for guiding replication, adaptation, and scaling of SRSD

and its AI-supported integration in a range of educational contexts. The evaluation team's resumes are included in Appendix B3. References for this Part D Evaluation plan are available in Appendix J11.

References – Parts A, B, C (References for Part D in Appendix J11)

Ball, D., & Cohen, D. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession* (pp. 3–32). Jossey-Bass.

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

Carver-Thomas, D., & Darling-Hammond, L. (2017). Teacher turnover: Why it matters and what we can do about it. *Learning Policy Institute*.

Coburn, C. E., & Russell, J. L. (2008). District policy and teachers' social networks. *Educational evaluation and policy analysis*, 30(3), 203-235.

Cohen, D. K., & Hill, H. C. (2008). *Learning policy: When state education reform works*. Yale University Press.

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340.

Fixsen, D. L., Blase, K. A., Naoom, S. F., Van Dyke, M., & Wallace, F. (2009). Implementation: The missing link between research and practice. *NIRN implementation brief*, 1(1), 218-227.

Freedman, S. W., Hull, G. A., Higgs, J. M., & Booten, K. P. (2016). Teaching writing in a digital

and global age: Toward access, learning, and development for all. In D. H. Gitomer & C. A. Bell (Eds.), *Handbook of Research on Teaching* (5th ed., pp. 1389–1450).

Washington, DC: American Educational Research Association.

[REDACTED]

Hodges, T. S. (2022). The WRITE principles for developing a research-to-practice cycle

- professional development. *The Reading Teacher*, 76(2), 158-167.
- Lee, S., & Schmidgall, J. (2020). The importance of English writing skills in the international workplace. Research Memorandum. ETS RM-20-07. *Educational Testing Service*.
- McKeown, D., FitzPatrick, E., & Sandmel, K. (2014). SRSD in practice: Creating a professional development experience for teachers to meet the writing needs of students with EBD. *Behavioral Disorders*, 40(1), 15–25.
- National Center for Education Statistics. (2012). *The nation's report card: Writing 2011*. Institute of Education Sciences, U.S. Department of Education.
<https://nces.ed.gov/nationsreportcard/pdf/main2011/2012470.pdf>
- Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.).
- Sims, S., Fletcher-Wood, H., O'Mara-Eves, A., Cottingham, S., Stansfield, C., Goodrich, J., . . . Anders, J. (2023). Effective teacher professional development: New theory and a meta-analytic test. *Review of Educational Research*, 00346543231217480.