ENGAGING SCIENCE LEARNING WITH OPENSCIED

CONTENTS

INTRODUCTION

To help science teachers in high-need middle schools foster deep, engaging science learning for all students, BSCS Science Learning, American Institutes for Research® (AIR®), and Southern University, propose a project that exemplifies **Absolute Priorities 1 and 3,** and **Competitive Preference Priority 1.** The *Engaging Science Learning with OpenSciEd* project investigates the efficacy of a widely adopted middle school science program that includes classroom materials and professional learning resources. Designed for the *Next Generation Science Standards* (NGSS; NGSS Lead States, 2013), *OpenSciEd* has received all green ratings from EdReports and is rated as high quality by NextGenScience, and is freely available to anyone.

This field-initiated program was developed by a diverse consortium of educational research and development institutions across the U.S., in collaboration with state and local education agencies, in response to conclusions drawn at a 2017 workshop held by the National Academy of Sciences (NAS, 2018). The workshop concluded that there was an urgent need for high quality instructional materials and associated professional learning (PL) that would allow districts to meet the NGSS that was not being met by commercial publishers, developed by teachers themselves, or created by researchers. The goal of the consortium was to create a program that would achieve the vision of the National Research Council *Framework for K-12 Science Education* (NRC, 2012) and the NGSS in a form that would be practical for teachers to implement with ordinary resources and diverse student populations. The consortium developed a set of Design Specifications to guide the development of a comprehensive middle school science program for grades 6—8, which included a clearly articulated instructional model, curriculum-based PL guidelines, and a focus on equitable science instruction. The program underwent a rigorous 18-month development process with teacher and student voices informing each unit's storyline. The *OpenSciEd* program has been used by over 58,000 teachers in the United States. Teachers have seen their students strengthen their ability to solve problems, become more curious about the world around them, and be excited to ask questions they care about.

In partnership with districts in Louisiana that serve high-need students (see letters of support) and Southern University (an HBCU), this project will determine whether, and in what ways, **the *OpenSciEd* Middle School program (which we define as both the instructional materials and professional learning supports)** can improve achievement for high-need students. The project will include 2 cohorts of Grade 8 science teachers, as shown in Exhibit 1. A pilot cohort will provide opportunities for Southern University to build capacity to deliver the PL program, and the BSCS and AIR teams to refine data collection tools and procedures. The AIR evaluation team will conduct a teacher-level matched quasi-experimental study of one full year of the *OpenSciEd* program, providing rigorous evidence of program effectiveness and meeting What Works Clearinghouse standards with reservations. An implementation study will be conducted throughout the project to provide actionable feedback and refine the program.

Exhibit 1. Number of Teachers in Each Cohort by Year

	Year 1	Year 2	Year 3	Year 4	Year 5
Pilot Cohort (10 teachers)		10 Teachers			
Cohort 1 (30 teachers)			15 treatment teachers and 15 control (delayed treatment) teachers	Delayed treatment with the 15 Cohort 1 control teachers	
Cohort 2 (40 teachers)				20 treatment teachers and 20 control (delayed treatment) teachers	Delayed treatment with the 20 Cohort 2 control teachers

A. SIGNIFICANCE

A.1. Changing Existing Practice to Address Inequitable Achievement in Science

This project addresses science proficiency, a problem for the nation overall and especially for high-need students, and focuses on middle grades where inequalities are particularly pronounced. Policy makers and educators assert that science knowledge is critical to economic productivity and well-being in the U.S., but the National Assessment of Educational Progress (NAEP) shows that 33 percent of eighth grade students fail to reach the “Basic” science achievement level (NAEP, 2019). The problem is most pronounced for Black students and Hispanic students (58% and 47% below Basic, respectively, compared to 20% of white students) and for students from low-income households (48% compared to 19% from higher income households, as measured by eligibility for the National School Lunch Program). **These achievement gaps persist and expand** as students move into high school, where, for example, 69% of black and 56% Hispanic students fail to perform at the Basic achievement level, compared to 28% of white students, pointing to the critical role of an effective science education in early and middle grades. Since the COVID-19 pandemic has resulted in unprecedented decreases in reading and math achievement, and significant widening of achievement gaps (Kuhfeld, Soland, and Lewis, 2022; Fahle et al., 2023) it is expected that the same patterns will be evident in science.

These persistent problems in science education led educators and policy makers to forge a new vision for science teaching and learning—one that addresses past criticism that science instruction in the U.S. lacks focus, connection to student experience, and an authentic view of the scientific enterprise (e.g., NRC, 2007). The development and adoption of research-based teaching practices, represented by the *K-12 Framework for Science Education* and the *Next Generation Science Standards* (NGSS Lead States, 2013; NRC, 2012) is an attempt to transform science instruction and student engagement in science learning. However, most teachers are unprepared to make the dramatic departure from current practices required by contemporary visions of science teaching and learning, and existing curriculum materials and PL fail to support such shifts (Wilson, 2013; Osborne, 2014; Darling-Hammond, Hyler, Gardner, 2017). In response, *OpenSciEd*, launched in 2017, is a collaboration of materials developers, PL experts, educational researchers, classroom educators, and educational leaders. Its goal is to support a transformation of science teaching and learning across the United States by creating a comprehensive science program for grades 6–8 that is distributed for free under an open license. Development was funded by a group of foundations, including the Bill and Melinda Gates Foundation, the Carnegie Corporation of New York, the Charles and Lynn Schusterman Family Foundation, and the William and Flora Hewlett Foundation. The development consortium included BSCS Science Learning, the Dana Center at the University of Texas, Boston College, The Next Generation Science Storylines Project at Northwestern University, and Digital Promise. The consortium collaborates with a wide range of state and local education agencies to support successful implementation. *OpenSciEd* was designed to engage students in active learning of science through sense-making, problem-solving, and decision-making. The teacher's role shifts from the conveyer of information, to being responsible for creating a context for learning – facilitating discussions and creating a classroom that supports learning for all students.

A.2. Building on a Promising Approach to Deep, Effective Science Learning

The notion that curriculum materials truly matter and directly influence the learning process has long been supported in the literature (e.g., Forbes & Davis, 2010; Schmidt, McKnight, & Raizen, 1997). Curriculum materials play a defining role in classrooms, affecting both what and how teachers teach. When teachers have access to high-quality instructional materials, they can focus their time and creativity on bringing lessons to life and finding ways to inspire their students to learn and grow. However, too many teachers are provided with outdated and unaligned materials, sending them online in search of resources. Teachers spend an average of seven hours per week searching for instructional materials and five more hours per week creating materials to accompany their lessons (Goldberg, 2016). Further, in the absence of quality PL opportunities, teachers are unlikely to recognize high-quality curriculum and even less likely to create equitable science learning experiences in line with the vision of the NRC framework.

Prior to *OpenSciEd*, instructional materials that have been developed in response to the NGSS were either partial implementations or impractical to implement broadly. Teacher- and district-developed programs were highly variable in quality and didn't offer a scalable model. Further, implementing research-based instructional materials requires effective professional learning (NAS, 2015), but most PL efforts seldom provide teachers with the science content knowledge and pedagogical skills necessary to help them teach in ways called for in current reforms (Reiser, 2013; Wilson, 2013). The challenges are especially prevalent for middle-school teachers who have limited training in science-specific pedagogy or in the science disciplines they are expected to teach (Darling-Hammond, Hyler, & Gardner, 2017). ***OpenSciEd* brings together high-quality research-based classroom materials and an aligned set of PL resources** with explicit supports for teachers and students' engagement in phenomenon-driven learning.

The *OpenSciEd* Middle School program is an effective strategy that addresses this critical national problem, the central premise of which being that **instructional materials, along with aligned professional learning, can play an important role in supporting lasting change in classroom practices and educational achievement** (Edelson, et al., 2021). Here we build on this program by studying its efficacy on student outcomes, and, in collaboration with educators at a public historically black land-grant university, we explore how the PL component can be co-adapted to address the needs and embrace the resources of high-need students in Louisiana.

A.3. A Promising, Recognized and Widely Adopted Program

EdReports Review. In 2023, EdReports, an independent nonprofit that conducts evidenced-based reviews of K-12 instructional materials, awarded *OpenSciEd* an all-green rating from

EdReports for meeting all expectations in the three EdReports gateways: Designed for NGSS, Coherence and Scope, and Usability. Their report noted that that the program consistently incorporates and assesses the three NGSS dimensions for students' sensemaking, and the materials leverage science phenomena and students' prior knowledge and experiences in the context of driving learning and student performance. EdReports concluded that the materials are coherent in design, scientifically accurate, and support grade-band endpoints. Finally, they found that the program includes opportunities for teachers to effectively plan and utilize materials with integrity and to further develop their own understanding of the content; provide tools, guidance, and support for teachers to interpret and act on data about student progress; and includes materials designed for all student's regular and active participation (EdReports, 2022).

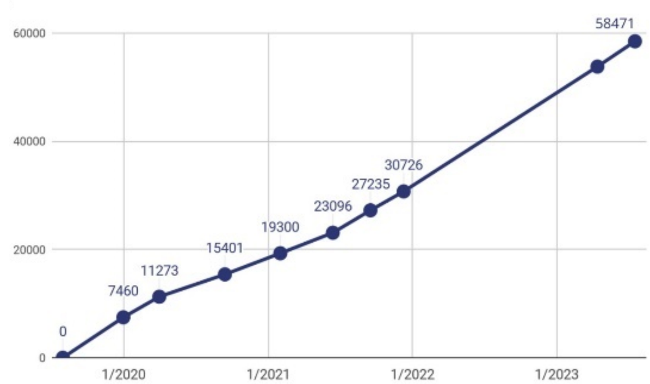
EQuIP Review. NextGenScience (2022) has developed the *EQuIP Rubric for Science*, which provides criteria by which to measure how well lessons and units are designed for the NGSS. The EQuIP review, conducted by an independent review panel, aims to identify lessons and units that best illustrate the cognitive demands of the NGSS. All the middle school *OpenSciEd* units have been rated as high-quality by the EQuIP review process, indicating that they exemplify high-quality design for the NGSS across all three categories of the rubric: I) NGSS 3D Design, II) NGSS Instructional Supports, and III) Monitoring NGSS Student Progress.

Promise of Efficacy from the Field Test. The field test showed that the program can be implemented by a wide variety of teachers in a diverse range of classrooms - the field test included a diverse student population of more than 12,400 students and a diverse teacher population of 341 teachers across 10 states in rural, urban, and suburban districts. In pre and post measures, field-test teachers were shown to develop more confidence with developing NGSS-aligned teaching practices. They also reported that they found the *OpenSciEd* program to be effective with a diverse group of students (95% of teachers), including students historically marginalized in STEM (Deverel-Rico et al., 2023). Through student responses from exit ticket data, students were shown to engage in collaborative sensemaking through reasoning about phenomena (80%), participating in whole-class discussions (77%), building on each other's ideas (71%), and making their thinking visible (56%). Most students also reported that the curriculum was relevant to them or other people in their lives (Deverel-Rico et al., 2023).

While the EdReports and EQuIP reviews point to the program being of high-quality, and the field-test and adoption data show promise of efficacy and potential for wide scale implementation, neither demonstrate that the *OpenSciEd* Middle School program is efficacious when it comes to student achievement. **It is now critical to establish the efficacy of this program on student achievement with rigorous and causal research.**

Wide Adoption. Since its release, *OpenSciEd* has grown to be used by over 58,000 teachers in 37 states (Exhibit 2). In the last three months alone (since May 2023) there have been 41 professional learning events totaling 125 days around the country. The program is also currently being used in the UK, Colombia, Japan, and the Bahamas.

Exhibit 2. Registered Users of OpenSciEd

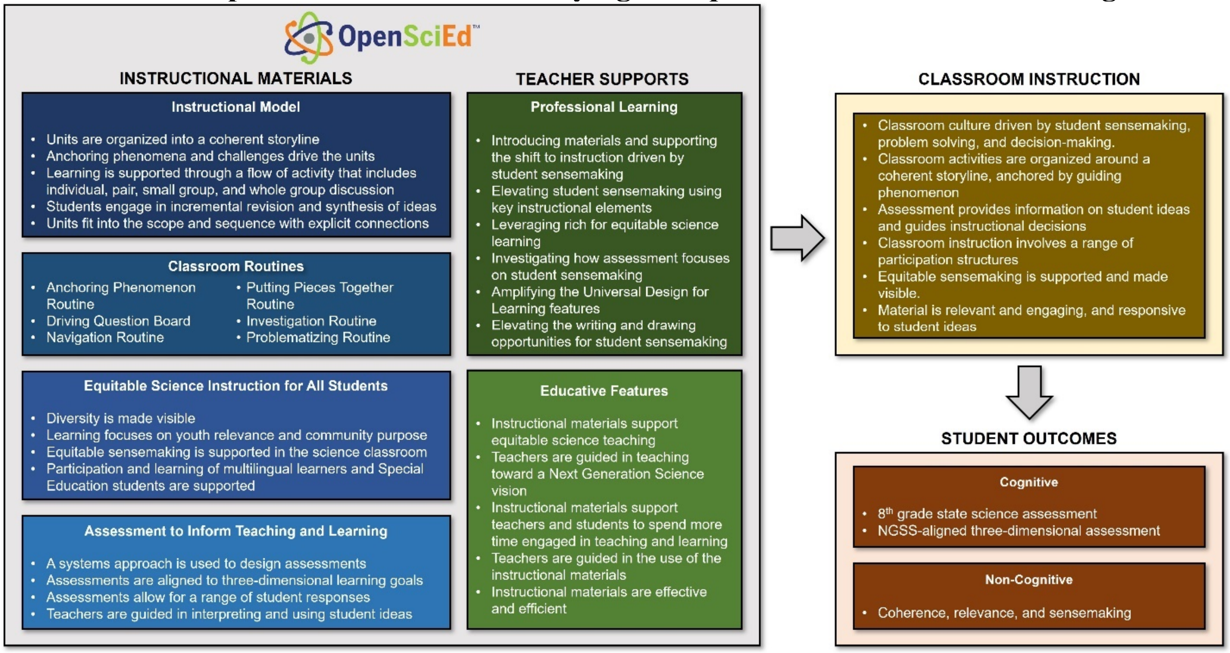


B. QUALITY OF THE PROJECT DESIGN

B.1. A Clearly Articulated Conceptual Framework

The conceptual framework for the *OpenSciEd* program (Exhibit 3) is clearly articulated and grounded in research and practical experience. It describes the two main components of the program: the instructional materials and the teacher learning supports.

Exhibit 3. Conceptual Framework Underlying the *OpenSciEd* Middle School Program.



An **Instructional Model** is the heart of any design framework for instructional materials. The *OpenSciEd* developers selected the Next Generation Science Storylines instructional model (Reiser, Novak, & McGill, 2017) because it places phenomenon-driven, three-dimensional learning called for by the NRC Framework and NGSS at the center of teaching and learning. In this model, the flow of lessons builds toward three-dimensional performance expectations in a coherent storyline, while making sense to students from their own perspectives. Learning is motivated by attempting to make sense of anchoring phenomena related to the science learning targets, leading to iterative cycles of investigating, improving explanations with new evidence, and further questioning. Students are positioned as collaborators who work as a community to figure something out about the world. To enable access and participation for all students, lessons provide opportunities for students to work out their thinking in a variety of participation structures, including individual, pair, and small group contexts, prior to whole group discussions. Developing explanatory ideas requires figuring out pieces of ideas and then assembling them into more complex explanations, then providing opportunities for students to revise and improve ideas. Finally, units are designed to support a coherent learning experience, where each unit builds explicitly on the ideas that have been established in earlier units.

The instructional model takes advantage of five **Classroom Routines**: structures that play specific roles in advancing the unit's storyline. The routines follow a pattern as students kick off a unit of study, investigate different questions they have, put the pieces together, and then consider the next set of questions to investigate. **Equitable Science Instruction for All Students.** Recognizing the range of student diversity in today's classrooms, *OpenSciEd* builds on guidelines in *A Framework for K-12 Science Education* to support learners who come from non-dominant communities or are underrepresented in STEM. These features of the program are described below in Section B3, along with the **teacher learning components**.

Assessment to Inform Teaching and Learning. *OpenSciEd* assessments are designed in tandem with the materials so that evidence gathered can inform teaching and learning. The program takes multiple purposes of assessments into account, including formative tasks that support the use of student ideas to inform instruction. Assessments are aligned to the NGSS learning goals, with the multiple assessment opportunities in a unit providing evidence of students' ability with all three dimensions. Assessments anticipate the wide range of experiences, resources, and ideas that students bring to the classroom, and allow students to express their understanding in a language and format in which they are most comfortable. The instructional materials include assessments where student ideas are not considered simply right or wrong, but rather as ideas that can be used to support a progression toward higher levels of understanding.

B.2. Clearly Specified and Measurable Goals, Objectives, and Outcomes

The project's goal is to test the *OpenSciEd* instructional materials and engage in local adaptation and refinement of teacher professional learning supports. Exhibit 4 specifies the objectives, strategies, and outcomes to be achieved, and how each outcome will be measured.

Exhibit 4. Objectives, Strategies, Outcomes, and Measures.

Strategies	Outcomes	Measures
Objective 1: Leverage expertise and perspective of Southern University to co-adapt, pilot, monitor and refine the professional learning program to support implementation of *OpenSciEd* (Years 1 and 2)		
Strategy 1.1. Co-Adapt PL program	BSCS and Southern University co-adapt the *OpenSciEd* PL program for the target population.	Measure 1.1. Monthly memos documenting changes to the PL program.
Strategy 1.2. Recruit pilot teachers	Recruit 10 8th grade teachers to participate in the Year 2 pilot.	Measure 1.2. Districts, principals, and teachers sign MOU.
Strategy 1.3. Pilot PL program	Pilot PL with 8th grade teachers in high-need schools.	Measure 1.3. Documentation of teacher participation in PL activities.
Strategy 1.4. Implement curriculum materials	Implement *OpenSciEd* materials with 10 teachers in 8th grade in high-need schools.	Measure 1.4. Weekly teacher implementation logs. Periodic interviews and focus groups.
Strategy 1.5. Collect implementation data	Project partners collect implementation data on the extent to which all project activities were implemented as planned.	Measure 1.5. Weekly teacher implementation logs. Periodic interviews and focus groups.
Strategy 1.6. Analyze implementation and share findings	AIR analyzes all implementation data and shares findings with project partners.	Measure 1.6. AIR provides quarterly implementation briefs for project partners.
Strategy 1.7. Refine the program	BSCS and Southern University revise the PL program for the target population.	Measure 1.7. Biannual memos documenting changes made to the PL.
Objective 2: Examine the impact of *OpenSciEd* Middle School on state assessments, non-cognitive outcomes, and equitable learning (2 cohorts of a matched quasi-experiment in Years 3 and 4)		
Strategy 2.1. Recruit 35 treatment teachers	Eligible 8th grade teachers agree to participate in the study for one year.	Measure 2.1. Districts, principals, and teachers sign MOU.
Strategy 2.2. Recruit 35 matched control teachers	Eligible 8th grade teachers agree to participate in the study for one year.	Measure 2.2. Districts, principals, and teachers sign MOU.
Strategy 2.3. Implement revised *OpenSciEd* PL program	BSCS and Southern University provide summer PL for treatment teachers.	Measure 2.3. BSCS and Southern University provide documentation of enactment of planned activities and teacher participation.
Strategy 2.4. Implement *OpenSciEd* curriculum materials	Treatment teachers implement the full set of 8th grade *OpenSciEd* units.	Measure 2.4. Treatment teachers complete online logs documenting use of the materials.
Strategy 2.5. Collect implementation and impact data	Project partners collect implementation data on the extent to which all project activities were implemented as planned.	Measure 2.5. AIR collects implementation logs and data.
Strategy 2.6. Analyze implementation data	AIR analyzes all implementation data and shares findings with project partners.	Measure 2.6. AIR provides quarterly implementation briefs for project partners.
Strategy 2.7. Assess the impact of OpenSciEd on teacher and student outcomes	AIR conducts impact analyses and produces impact findings that meet WWC standards with reservations.	Measure 2.7. AIR documents impact analyses and findings in an impact findings memo.
Strategy 2.8. Examine mediation and moderation of student effects	AIR conducts analyses to assess the mediation and moderation of student effects by implementation and student demographic variables respectively.	Measure 2.8. AIR documents impact analyses and findings in an impact findings memo.
Strategy 2.9. Disseminate findings	Project partners enact a comprehensive dissemination plan to reach a broad range of stakeholders and decision-makers.	Measure 2.9. Dissemination products are produced and shared with stakeholders and decision-makers.
Objective 3: Explore how variations in implementation of *OpenSciEd* support engaging, relevant and coherent student learning, and what factors are needed to support coherent implementation (throughout the project in Years 1-5)		
Strategy 3.1. Implement the *OpenSciEd* program	Pilot, Cohort 1, Cohort 2, and delayed treatment (control) teachers implement the *OpenSciEd* program.	Measure 3.1. Project partners collect implementation logs and data.
Strategy 3.2. Collect Implementation data	Project partners collect data from a range of sources to examine variation in program implementation.	Measure 3.2. Documentation of data sources and classroom artifacts.
Strategy 3.3. Collect non-cognitive student outcome data	Collect data on how the program is perceived as engaging, relevant and coherent by students.	Measure 3.3. Student surveys, interviews, and exit tickets,
Strategy 3.4. Analyze implementation impacts on non-cognitive student outcomes	Use a range of qualitative and quantitative research methods to explore Objective 3.	Measure 3.4. Project partners document all analyses and findings in a findings memo.
Strategy 3.5. Develop implementation scaffolds and recommendations	BSCS and Southern University develop a set of recommendations to increase the impact of the program on non-cognitive outcomes.	Measure 3.5.Project report documenting implementation recommendations and guidelines.

B.3. A Project Design for Addressing the Needs of the Target Population

The *OpenSciEd* project is specifically designed to address the needs of the target populations – high-need students, middle school science students, and science teachers. In Louisiana, where this project will take place, 58% of students are non-white, 60% are economically disadvantaged, and 45% are below the Basic level in science (NCES, 2022). In East Baton Rouge Parish (see letters of commitment) the second largest school district in Louisiana that serves over 40,000 students, 89% of students are non-white, 81% are eligible for free/reduced price lunch, and all middle schools qualify for the schoolwide Title 1 program (Lousiana Dept. of Education, 2023).

High-Need Students. The *OpenSciEd* Middle School units address the needs of sciences teachers and students alike, and guide teachers in implementing equitable science instruction for all students, with particular attention to student groups who have historically been underserved in science. Equitable instructional practices are central to the material's design, not add-on strategies that need to be deployed in the presence of certain students. Instructional materials are flexible enough to be adapted to fit teachers' and students' local circumstances. Individuals, teams, and communities from all nations and cultures have contributed to science and to advances in engineering—across differences of race, ethnicity, gender, and abilities. The *OpenSciEd* materials make this diversity visible by including a broad range of images and stories of who does and has done STEM endeavors in our society, and prioritize the interests of underserved communities. Further, the instructional materials relate to the interests, identities, and experiences of students and the goals and needs of their communities. The materials support equitable sense-making by scaffolding multiple forms of engagement, and leverage students' sensemaking repertoires. To facilitate participation and learning of multilingual learners, the materials build on the lived experiences and linguistic resources that all students bring to the science classroom. Activities are included that create opportunities for teachers to leverage what they know about specific students' experiences to help students make personal connections to science content knowledge. Finally, the materials are designed to follow principles of the *Universal Design for Learning (UDL)* to support participation and learning of Special Education students. Findings from *OpenSciEd* field testing show promising impacts across student demographic groups. More than 90% of students report instruction to be relevant to them, and this was true for students from all racial backgrounds. Students from different racial backgrounds, genders, and linguistic backgrounds reported they contributed at high levels to classroom discussion, and say their ideas are taken seriously by others.

Middle School Grades. The *OpenSciEd* program is designed to address low achievement and widening gaps in middle school grades, as described in Section A.1. To address this, most U.S. states have now either adopted or adapted standards based on the NGSS, which sets up a need for instructional support aligned to the vision of these new standards. One shift unique to middle school grades is in how the domains that were previously siloed by grade are now integrated – meaning, each middle school grade addresses standards from all Disciplinary Core Idea domains. The *OpenSciEd* program is designed to support educators in making this shift.

Teacher Professional Learning (PL). As described in Section A.1., PL focused on the implementation of high-quality materials that is aligned with local teacher needs, can have a significant impact on teaching and learning (Darling-Hammond, 2017). When teachers adopt research-based curriculum materials, it is essential that they learn about key features of the materials as well as the rationale behind them (Lin & Fishman, 2006). Teacher PL has too often been treated as a separate path to improvement, independent of the use of high-quality instructional materials. In contrast, the work on the *OpenSciEd* Middle School program has focused on design frameworks for both classroom instruction for students and PL for teachers.

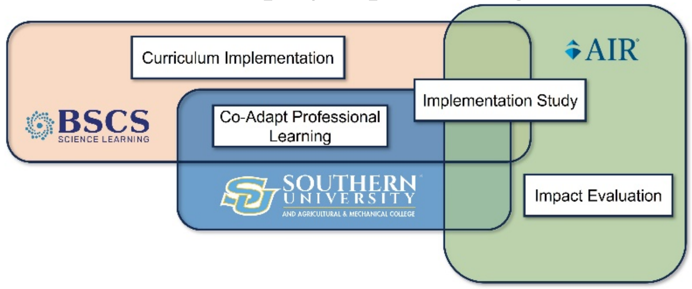
This attention to how instructional materials and PL can work together to support shifts in teacher practice is an important characteristic of *OpenSciEd*. Teachers develop the tools needed to switch from a lecturer to a learning facilitator that helps students own and drive their learning. There are six different PL sessions that support the classroom materials, and a corresponding six sessions to train PL facilitators (Appendix J7). Like the teacher and student materials, all *OpenSciEd* PL materials are Open Educational Resources. **An important component of this project is the co-adaptation of existing *OpenSciEd* PL resources for the local context.** This work will be a collaborative effort between BSCS Science Learning, who have in-depth knowledge of the program, and Southern University, who bring in-depth knowledge of the local context. Together, these teams will adapt the existing PL modules to develop a Summer Institute that will support local teachers in enacting the units with the full range of student goals in mind.

Educative Features are the elements that are added to materials that are explicitly intended to promote teacher learning (Davis & Krajcik, 2005). *OpenSciEd's* educative features are designed to support the wide range of teachers who use the materials, including attending to issues of equity. Embedded educative features include information on the instructional model, additional scientific background, alternative understandings students may have, and scaffolds to support classroom culture conducive to spending more time engaged in teaching and learning.

C. QUALITY OF THE PROJECT PERSONNEL

As shown in Exhibit 5, three organizations will lead the work of the proposed project, bringing expertise in implementing research-based curriculum materials, professional learning, implementation research, and impact evaluation (see Appendix B for résumés). Each of the three organizations is committed to equity, diversity, and inclusion, has explicit policies to encourage recruitment of diverse staff, and will commit to staffing the project with a diverse group of team members with experience working with the high-need populations and communities the project will serve.

Exhibit 5. Primary responsibilities of the three project partner organizations



[redacted] **PhD**, Director of Research & Innovation at BSCS Science Learning, will serve as the Project Director and will be responsible for overall project leadership. [redacted] has over 20 years of experience in science education research and has managed multiple large-scale NSF and DoE projects. [redacted] is currently Project Director of a Mid-phase EIR award exploring a science PL program for rural teachers and high-need students in Tennessee and Kentucky (U411B190029). He has also served as PI of NSF research projects examining the impact of interventions in science teaching and learning, including the STeLLA High School project, exploring the efficacy of a video-analysis based approach to science teacher professional learning (DUE #1503280) and The Model-Based Education Resource project (DUE #1813538) exploring impacts from a high school biology program focused on engaging students in developing and using models. Through these and other projects, [redacted] has developed expertise in project management, including managing timelines, staffing, budgets, human subjects, data management, and dissemination. Implementation of the *OpenSciEd* materials will be overseen by [redacted] **PhD**, BSCS Director for Design for Social and Environmental Justice Outcomes, and *OpenSciEd* Middle School unit development lead. [redacted] work focuses on creating science education programs that provide equitable and meaningful opportunities for students from marginalized communities to learn science. [redacted] **PhD**, Senior Science Educator at BSCS Science Learning, will manage the PL aspects of the project. [redacted]

[redacted] work focuses on establishing equitable access for all students to rigorous, inquiry-based science instruction, and has extensive expertise in effective PL models. [redacted] **PhD**, BSCS Science Educator, will lead the BSCS side of the PL program co-adaptation and associated exploratory research. [redacted] work focuses on designing research-driven, professional learning approaches to support meaningful science learning, and was a lead designer of the *OpenSciEd* middle school professional learning model. [redacted] BSCS project manager, will be responsible for managing deliverables, timelines, and the project budget. [redacted] is an experienced project manager and has served in this role on a previous EIR project.

On the evaluation team, [redacted] is a principal researcher at the American Institutes for Research (AIR) and Assistant Professor in the Department of Leadership, Research, and Foundations at the University of Colorado, Colorado Springs. [redacted] directs intervention impact studies, syntheses of the effects of science education interventions, and contributes to AIR's work with the What Works Clearinghouse (WWC). [redacted] is currently the Principal Investigator (PI) for an evaluation of an EIR award that is studying the effects of a promising professional development program for elementary science teachers (U411B190029). [redacted] will be the PI for the evaluation of the proposed project, responsible for managing all aspects of the project evaluation. [redacted] a doctoral candidate, is a senior researcher at AIR, with extensive experience with educator effectiveness projects and student outcome analysis at both state and district levels. She works with [redacted] on the current EIR evaluation as the Project Director and will serve as the Project Director and co-PI for the proposed evaluation. [redacted] a senior researcher at AIR, will serve as the Impact Evaluation Lead, overseeing impact data collection and data analysis at the task level and contributing to reporting and dissemination activities. All three AIR senior personnel are WWC certified in group design studies.

At Southern University, [redacted] is Assistant Professor in the Dept. of Curriculum and Instruction at Southern University and A&M College. Her research focuses on attracting and retaining underrepresented teachers of color in STEM education and will lead the co-adaptation of the PL program. [redacted] is Associate Professor in the Southern University Dept. of Curriculum and Instruction, where she conducts STEM education research and teaches graduate-level courses. [redacted] will work alongside BSCS in adapting the *OpenSciEd* PL program to help the Louisiana teachers meet the needs of local students. [redacted] is Assistant Professor at Southern University, with expertise in conducting professional learning for emerging educational researchers. [redacted] will be responsible for monitoring and evaluating the local adaptations to the *OpenSciEd* PL model.

D. QUALITY OF THE MANAGEMENT PLAN

D.1. A Management Plan Defining Responsibilities, Timelines, and Milestones

Our management plan ensures that the objectives will be achieved on time and within budget, and draws on the expertise of the personnel described in Section C. An organizational chart can be found in Appendix J6. Each project objective and strategy described in Exhibit 4 has a timeline and associated milestones, and will be led by one of the three teams in collaboration with the other project partners. The timeline chart presented in Appendix J5 indicates the lead staff and milestones for each strategy. **BSCS Science Learning** is the lead organization for the study responsible to the U.S. Department of Education for grant performance. BSCS Science Learning's primary project roles will be to a) support the implementation of the *OpenSciEd* materials by treatment teachers, b) adapt the PL program with the Southern University partner c) support the implementation of the *OpenSciEd* PL program by Southern University, d) continuously improve project materials and resources, e) support evaluation activities and data collection, and f) coordinate the work of the partner organizations. BSCS is in its 64th year as a leader in science education, with a commitment to transforming teaching and learning for all students. BSCS has demonstrated its capacity for developing, implementing, and studying innovative science curriculum and PL through numerous IES and NSF awards, with recent publications from those studies in the *Journal of Research on Educational Effectiveness*, the *American Educational Research Journal*, and the *Journal for Research on Science Teaching*. BSCS staff have the expertise required to develop and implement research-based curriculum materials and effective PL programs, as well as the experience managing complex projects required to anticipate and prevent issues before they arise. BSCS was central to the development of the *OpenSciEd* program, and will ensure that implementation is of high quality and consistent.

To ensure the independence of the evaluation, the **American Institutes for Research (AIR)** will conduct all aspects of the evaluation and will have no role in the development or the implementation of the *OpenSciEd* intervention except to share implementation analyses as feedback. This structure ensures the independence of key evaluation activities including random assignment, outcome data collection, analysis, and reporting, and is consistent with OII guidance (U.S. Dept. of Education, Office of Innovation and Improvement, n.d.; Abt Associates, 2015). AIR is uniquely qualified for this role, having successfully led four projects for IES in the last decade focused on teacher PD interventions and curriculum evaluation. These projects involved coordinating across subcontracted organizations, including an intervention provider and several school districts, recruiting schools and teachers, and conducting an independent evaluation. AIR's experience monitoring intervention providers and providing feedback on fidelity for continuous improvement also helps ensure relevant, actionable feedback from the evaluation team, which will draw on instruments and methods that AIR has refined across several studies.

Southern University and A&M College is a public historically black land-grant university located in Baton Rouge, Louisiana. Their mission is to provide opportunities for a diverse student population to achieve a high-quality global educational experience, to engage in scholarly research, and creative activities, and to give meaningful public service to the so that Southern University graduates are competent, informed, and productive citizens. The School of Education provides intensive educational and field experience opportunities for teacher candidates and professional educators, and provides professional service and leadership to school systems within the region and state. Southern University has extensive experience working with high-need schools in Louisiana, and is uniquely positioned to work alongside BSCS in adapting and implementing the *OpenSciEd* professional learning program for this population of teachers.

E. QUALITY OF THE PROJECT EVALUATION

AIR will conduct a rigorous and independent evaluation of *OpenSciEd* middle school science program, designed to meet What Works Clearinghouse (WWC) standards with reservations for science teaching and science achievement outcomes. The evaluation will provide BSCS Science Learning with timely and actionable formative feedback essential for ongoing monitoring and improvement of program implementation. The evaluation will also examine the impact of the associated professional learning (PL) program on teachers' classroom instruction. BSCS Science Learning will support the implementation of *OpenSciEd* materials by recruiting 70 teachers (35 treatment and 35 comparison then delayed treatment) within seven school districts, across two cohorts. The implementation and impact evaluation will occur in academic years 2025–26 and 2026–27 (Project Years 3 and 4). In Project Year 5 we will conduct data analysis and disseminate findings (See Appendix J1. for the evaluation timeline). Our impact evaluation will address research questions (RQs) 1–4 in Exhibit 6 using a propensity score matched quasi-experimental design, with assignment at the teacher level and blocking at the district level. The impact of *OpenSciEd* on state science achievement scores will be a confirmatory contrast. All other outcomes in the impact evaluation are considered exploratory.

Exhibit 6. Research Questions and Data Sources

Research questions (RQs)	Primary data source(s)
Impact questions[1]	
1. What is the impact of *OpenSciEd* materials on eighth-grade students' science achievement? Confirmatory: Impact on state science test score Exploratory: Impact on project specific measure of achievement	State science test – Grade 8; Project-specific measure of science achievement (NGSS-aligned)
2. What is the impact of *OpenSciEd* materials on eighth-grade students' perceptions of their curriculum materials as relevant, coherent, and supportive of sensemaking? (Exploratory)	*Student Survey of Classroom Instruction* (project specific) 2)
3. To what extent does the impact of the *OpenSciEd* materials on student outcomes vary by demographic characteristics (e.g., race, gender, poverty levels)? (Exploratory)	State science test – Grade 8; Project-specific measure of science achievement (NGSS-aligned); Student administrative records
4. What is the impact of the *OpenSciEd* PL on teachers' classroom instruction? (Exploratory)	Teachers' classroom practice indicators from *Student Survey of Classroom Instruction*
5. To what extent is the impact of *OpenSciEd* curriculum materials mediated by its impact on teacher instruction? (Exploratory)	Teachers' classroom practice indicators from *Student Survey of Classroom Instruction;* State science test – Grade 8; Project-specific measure of science achievement (NGSS-aligned)
Implementation questions[2]	
6. To what extent are the *OpenSciEd* PL program and instructional materials being implemented with fidelity across schools?	Program records: Teachers' weekly log of unit delivery, PL session attendance.
7. What are teachers' experiences with *OpenSciEd* PL and instructional materials? To what extent do teachers and students perceive that *OpenSciEd* materials support equitable instruction and learning?	Biannual teacher survey and annual student survey; teacher interviews
8. How do BSCS and Southern University co-adapt the PL supports to support implementation of *OpenSciEd*? To what extent do teachers perceive that the supports are relevant and responsive to their local contexts?	Program records: Biannual teacher survey; teacher interviews; BSCS and Southern University staff interviews

Notes. [1]For impact questions, baseline data will be collected in Spring 2025 (Cohort 1) and Spring 2026 (Cohort 2) while outcomes data will be collected in Spring 2026 (Cohort 1) and Spring 2027 (Cohort 2). [2]For implementation questions, data will be collected during 2024–25 through 2026–27 (Pilot cohort and Impact Cohorts 1 and 2 treatment teachers).

E.1. Evidence That Meets WWC Standards

To estimate the impacts of *OpenSciEd* materials and PL on student achievement, perceptions, and teacher practice, AIR will use a propensity score matching design that **will meet WWC standards with reservations**. This design allows for direct comparison of student outcomes from routine users of *OpenSciEd* materials with that of routine users of business-as-usual materials and instruction. Empirical within-study comparisons demonstrate that studies using propensity score methods can reproduce the results of randomized controlled trials (Shadish et al., 2008). The proposed model will match the propensity scores of *OpenSciEd* teachers with those of teachers from the same district who do not use *OpenSciEd* materials. Propensity scores will be based on teacher characteristics such as educational background and teaching experience, as well as prior student achievement levels and student demographics.

We will assess impact for Cohort 1 and Cohort 2 during the implementation years—between fall 2025 and spring 2026 for Cohort 1 and between fall 2026 and spring 2027 for Cohort 2. AIR will conduct separate within-district matching for the teachers in Cohort 1 in fall 2025 and for Cohort 2 in fall 2026 to ensure valid matches for each cohort. The design will generate contrasts of comparison group teachers with *OpenSciEd* teachers who are equivalent on key characteristics specified in the WWC's *Study Review Protocol*, *Version 5.0*, thereby greatly increasing the likelihood of satisfying the WWC's baseline equivalence standard and allowing the study to receive the research rating *Meets WWC Standards with Reservations* under the *WWC* Group Design Standards, Version 5.0. To achieve this, AIR will repeat the matching process and re-assess baseline equivalence until the baseline variables, particularly prior year's student achievement scores and at least two baseline demographic characteristics, have a standard mean difference (SMD) across treatment conditions that is less than 0.25 standard deviations. Any baseline variables that have SMD between two groups to be between 0.05 and 0.25 standard deviations will be added into impact models for covariate adjustment.

Analytic Plan. After constructing the matched groups and testing for baseline equivalence, AIR will measure the **effect of *OpenSciEd* materials on student achievement (RQ1)** using a multilevel regression analysis that will estimate differences in student outcomes between *OpenSciEd* teachers and matched comparison teachers. Models will control for student covariates such as prior-year achievement and demographic characteristics, as well as the teacher-level covariates (e.g., experience and educational background) used in the matching process. A power analysis estimated that the minimum detectable effect size (MDES) is 0.14 standard deviations for students' science outcomes, post anticipated attrition. Thus, this design is well powered given that a recent large-scale meta-analysis of science education intervention effects (Taylor et al., 2018) observed an average treatment effect of 0.49 standard deviations with 95% confidence interval [0.37, 0.61]. Details of the analytic model for our confirmatory **RQ 1** and corresponding power analyses are provided in Appendix J2. Exploratory **RQ 2** (impact on students' perceptions of the curriculum) will be modeled in a similar fashion. For **RQ 3**, the binary treatment indicator will be interacted with student-level demographic indicators and the corresponding interaction effect will be an estimate of any differential impact of *OpenSciEd* program across student subgroups.

For the effect of *OpenSciEd* on teachers' classroom instruction **(RQ 4),** we will conduct a regression analysis to estimate mean differences in classroom instruction scores between *OpenSciEd* and comparison group teachers. Classroom instruction will be measured by selected indicators from the project-specific Student Survey of Classroom Instruction. The analysis will control for teacher characteristics (e.g., experience, education levels, pre-intervention classroom instruction scores), as well as those related to school demographics and context. For **RQ 5**, mediation will be tested on students' achievement scores combining the impact estimate from RQ 4 (effect of *OpenSciEd* on classroom instruction) with a separate estimate of the effect of classroom instruction on student achievement, controlling for treatment condition. Statistical significance of the *indirect/mediation* effect will be tested with the Sobel Test (Sobel, 1982).

Outcomes: For RQ 1 (impact on student science achievement), AIR will collect test scores and demographic/developmental information for all treatment and matched comparison students. The eighth-grade state science test in Louisiana is presumed reliable by the WWC and will meet WWC standards for outcome measures. The project-specific science assessment will be pilot tested in Project Year 1 to improve its psychometric properties prior to being used in the impact evaluation. All reported impacts on the project-specific science measure will be accompanied by corresponding psychometric information based on the impact sample (e.g., reliability/internal consistency coefficient). With sufficient reliability, the project-specific science measure will meet WWC standards for outcome measures. **For RQ 2 and RQ 4 (impact on teacher practice and students' perceptions of the curriculum materials)**, AIR will use a project-specific *Student Survey of Classroom Instruction* to collect data from treatment and control group students about teachers' classroom practice and students' perceptions of the curriculum materials (i.e., coherence, relevance, sensemaking). AIR will administer the survey post-intervention as an outcome measure and pre-intervention to demonstrate baseline equivalence. The project-specific student survey will be pilot tested in Project Year 1 to improve its psychometric properties prior to being used in the impact evaluation. All reported impacts on the project-specific student survey will be accompanied by corresponding psychometric information (e.g., reliability) based on the impact sample.

E.2. Performance Feedback and Periodic Assessment of Progress

To provide BSCS and Southern University with timely and actionable formative feedback, and to permit periodic assessment of progress toward achieving the intended program outcomes, AIR will collect and analyze quantitative and qualitative data collected from the pilot cohort and two impact cohorts. Data will include program records (e.g., teacher and facilitator PL session attendance), teacher and student surveys, artifacts and participant feedback, and teacher and program staff interviews. Feedback will center on three areas: (a) fidelity of implementation, (b) experiences and perceptions of the program, and (c) program adaptation in response to local contexts. To examine fidelity of implementation **(RQ 6),** AIR will work with BSCS to determine a set of fidelity indicators (e.g., teacher attends all PL sessions) and thresholds indicating the program was implemented as intended. AIR will analyze program records, including teacher and facilitator PL attendance records and the weekly teacher log of unit delivery, to determine the extent to which activities took place as planned. Examples of AIR's approach to measuring fidelity of implementation for two components (PL and classroom routines) are provided in Appendix J3.

To understand participants' experiences and perceptions of the *OpenSciEd* PL and instructional materials **(RQ 7)**, AIR will administer online surveys to teachers twice a year, once at the midpoint and once at the end of each implementation year from 2024–25 and 2026–27. Surveys will ask about teachers' experience with PL sessions and classroom instruction, particularly around opportunities for students to work out their thinking and participating in various participation structures, classroom routines, and assessment to inform teaching and learning. Surveys will also ask about teachers' perceptions of the instructional materials and learning supports with regard to the coherence among units, utility in promoting teacher learning and transforming instructional practice, and usefulness in promoting equitable science instruction. In addition, AIR will survey students at the end of each implementation year to collect information on students' classroom experience (i.e., participation, routines, and assessment) and perceptions about the coherence among units and the connection between instructional materials and students' interests, identities, and experiences.

To facilitate an in-depth understanding of implementation, AIR will also conduct 45-minute virtual interviews with 10 teachers each year from 2024–25 and 2026–27 to collect rich information about challenges experienced during implementation, resources needed to ensure successful implementation, and progress toward intended outcomes (e.g., equitable teaching and learning). To ensure that we capture a diversity of perspectives, AIR will purposively sample interviewees to provide representation across teaching experience, regions, schools, and classroom demographic characteristics. To analyze the survey data collected from teachers and students, AIR will use the Rasch model (Andrich, 1978; Rasch, 1980; Wright & Masters, 1982; Wright & Stone, 1979) to examine the psychometric properties of constructs and create valid scale scores for each construct (e.g., perceived coherence among units), which is made up of multiple items that connect conceptually. The analysis of interview data will rely on a systematic approach connected to the RQs (Denzin & Lincoln, 2003; Dey, 1993; LeCompte, 2000). AIR will triangulate the interview data with survey data to enable a systematic review across sources. See Appendix J4 for more details about the analytic approach. To understand how BSCS and Southern University co-adapt the PL supports to support implementation of *OpenSciEd* and the extent to which teachers perceive the supports to be relevant and responsive to their local contexts **(RQ 8)**, AIR will collect and analyze data via teacher surveys and interviews from the pilot cohort and two impact cohorts. For each of these cohorts, AIR will also conduct virtual interviews with 5 program staff from BSCS and Southern University to facilitate an in-depth understanding of how they collaboratively adapt PL supports in response to local contexts. Findings from this RQ will inform BSCS and Southern University on best practices and lessons learned in the process of tailoring *OpenSciEd* PL to the local needs of teachers and students.

To provide timeline feedback and support continuous improvement, AIR will meet regularly with BSCS and Southern University to provide timely performance feedback and assessment of the progress toward project objectives share interim findings, and produce a final summative report on the implementation and impact of *OpenSciEd* on student and teacher outcomes. To further **provide periodic assessment toward intended outcomes**, AIR will provide quarterly briefs from Year 1 to Year 4 to share preliminary results on implementation and annual interim findings memos in Years 4 and 5 to share preliminary results on impact. The periodic assessment findings will enable BSCS and Southern University staff to (a) identify factors that facilitate and impede successful program implementation, (b) continue program refinement, and (c) further tailor the approach to regional contexts.

E.3. Key Project Components, and Acceptable Implementation Thresholds

The proposed evaluation is informed by clearly articulated key components, mediators, and outcomes of the *OpenSciEd* program. As the logic model shows, *OpenSciEd* is a set of six program components: instructional model, classroom routines, equitable science instruction for all students, assessment to inform teaching and learning, as well as PL and educative curriculum features. AIR will analyze program data (see Section E2) to determine the degree to which program components are implemented with fidelity against a fidelity rubric, working with BSCS to identify clear and measurable thresholds. The thresholds for acceptable implementation will focus on the following key measures: teachers' delivery of *OpenSciEd* instructional materials and participation in *OpenSciEd* PL (see Appendix J3 for examples of thresholds). **Mediators and Outcomes**. The *OpenSciEd* Conceptual Framework **(Exhibit 3)** specifies that the *OpenSciEd* PL will improve teachers' classroom instruction and those improvements are at least partially responsible for improvements in students' science achievement and perceptions of their teachers' science instruction. Teachers' classroom instruction is conceptualized as a partial mediator of student outcomes of science achievement and perceptions of the relevance, coherence, and sensemaking supports of the *OpenSciEd* curriculum materials.

REFERENCES

Andrich, D. (1978). A rating scale formulation for ordered response categories. *Psychometrika, 43*, 561–573.

Darling-Hammond, L. (1997). *Doing what matters most: Investing in quality teaching*. New York, NY: National Commission on Teaching and America's Future.

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

Davis, E., & Krajcik, J. (2005). Designing educative curriculum materials to promote teacher learning. *Educational Researcher*, 34(3), 3–14.

Denzin, N., & Lincoln, Y. L. (2003). *Collecting and interpreting qualitative materials (2nd ed.)*. Sage.

Deverel-Rico, C., Penuel, W. R., Singleton, C., Allen, A., Krumm, A., & Pazera, C. (2023, In Press). *Lessons Learned from the OpenSciEd Middle School Field Test*.

Dey, I. (1993). *Qualitative data analysis: A user friendly guide for social science*. Routledge. https://doi.org/10.4324/9780203412497

Edelson, D. C., Reiser, B. J., McNeill, K. L., Mohan, A., Novak, M., Mohan, L., Affolter, R., McGill, T. A. W., Buck Bracey, Z. E., Deutch Noll, J., Kowalski, S. M., Novak, D., Lo, A. S., Landel, C., Krumm, A., Penuel, W. R., Van Horne, K., González-Howard, M., & Suárez, E. (2021). Developing research-based instructional materials to support large-scale transformation of science teaching and learning: The approach of the OpenSciEd middle school program. *Journal of Science Teacher Education*, 32(7), 780–804.

EdReports (2022). OpenSciEd Middle School Curriculum Review: https://www.edreports.org/reports/detail/openscied-2023/sixth-to-eighth/gateway-one

Fahle, Kane, Patterson, Reardon, Staiger and Stuart, (2023). School district and community factors associated with learning loss during the COVID-19 pandemic: https://cepr.harvard.edu/sites/hwpi.harvard.edu/files/cepr/files/explaining_covid_losses_5.23.pdf

Forbes, C. T., & Davis, E. A. (2010). Curriculum design for inquiry: Preservice elementary teachers' mobilization and adaptation of science curriculum materials. *Journal of Research in Science Teaching*, 47(7), 820–839.

Goldberg, M. (2016). Classroom Trends: Teachers as Buyers of Instructional Materials and Users of Technology (State of the K-12 Market 2016): https://mdreducation.com/wp-content/uploads/2020/12/StateofK12Market2016_ClassroomTrends.pdf

Kuhfeld, M., Soland, J., & Karyn Lewis, K. (2022). Test Score Patterns Across Three COVID-19-impacted School Years. (EdWorkingPaper: 22-521). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/ga82-6v47

LeCompte, M. D. (2000) Analyzing qualitative data. *Theory Into Practice, 39*(3), 146–154. https://doi.org/10.1207/s15430421tip3903_5

Lin, H.-T., & Fishman, B. J. (2006). Exploring the relationship between teachers' curriculum enactment experience and their understanding of underlying unit structures. In S. A. Barab, K. E. Hay, & D. T. Hickey (Eds.), Proceedings of the 7th International Conference of the Learning Sciences (pp. 432–438). Mahwah, NJ: Lawrence Erlbaum.

Lousiana Department of Education (2023). East Baton Rouge Parish District Snapshot: https://www.louisianabelieves.com/docs/call-to-action-toolbox/district-snapshot---east-baton-rouge-parish.pdf?sfvrsn=2

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP, 2022). 2022 Louisiana Overview: https://www.nationsreportcard.gov/profiles/stateprofile/overview/LA?cti=PgTab_OT&chort=2&sub=SCI&sj=LA&fs=Grade&st=MN&year=2009R3&sg=Gender%3A%20Male%20vs.%20Female&sgv=Difference&ts=Single%20Year&tss=2009R3&sfj=NP

National Academies of Sciences, Engineering, and Medicine. 2015. *Science Teachers' Learning: Enhancing Opportunities, Creating Supportive Contexts*. Washington, DC: The National Academies Press. https://doi.org/10.17226/21836.

National Academies of Sciences, Engineering, and Medicine. (2018). Design, Selection, and Implementation of Instructional Materials for the Next Generation Science Standards: Proceedings of a Workshop. Washington, DC: The National Academies Press. doi: https://doi.org/10.17226/25001.

National Research Council. (2012). A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/13165.

NextGenScience (2022) EQuIP Peer Review Panel for Science (PRP). OpenSciEd Middle School Review: https://www.nextgenscience.org/resources/examples-quality-ngss-design?field_exemplar_tags_target_id%5B371%5D=371&field_exemplar_tags_target_id_1%5B386%5D=386&field_exemplar_developers__target_id%5B501%5D=501

NGSS Lead States. 2013. *Next Generation Science Standards: For States, By State*s. Washington, DC: The National Academies Press.

National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: The National Academies Press.

Osborne, J., (2014). Teaching Scientific Practices: Meeting the Challenge of Change. *Journal of Science Teacher Education, 25*, 177-196.

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests.* Chicago: University of Chicago Press.

Reiser, B. J. (2013). *What professional development strategies are needed for successful implementation of the Next Generation Science Standards.* Invitational Research Symposium on Science Assessment: K-12 Center at ETS.

Reiser, B. J., Novak, M., McGill, T. A. W., & Penuel, W. R. (2021). Storyline units: An instructional model to support coherence from the students' perspective. *Journal of Science Teacher Education*, 32(7), 805–829. https://doi.org/10.1080/1046560X.2021.1884784

Schmidt, W., McKnight, C., & Raizen, S. (1997). *A splintered vision: An investigation of U.S. science and mathematics*. Boston, MA: Kluwer.

Shadish, W. R., Clark, M. H., & Steiner, P. M. (2008). Can nonrandomized experiments yield accurate answers? A randomized experiment comparing random and nonrandom assignments. *Journal of the American Statistical Association*, 103(484), 1334-1343.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, *13*, 290–312.

Taylor, J. A., Kowalski, S. M., Polanin, J. R., Askinas, K., Stuhlsatz, M. A., Wilson, C. D., & Wilson, S. J. (2018). Investigating science education effect sizes: Implications for power analyses and programmatic decisions. *AERA Open*, *4*(3).

U.S. Department of Education, Office of Innovation and Improvement. (n.d.). Evaluation activities and responsibilities (January 2015). Washington, DC: Author.

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2019 Science Assessment.

Wilson, S. M. (2013). Professional development for science teachers. *Science, 340*, 310-313. Washington, DC: American Association for the Advancement of Science.

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. Mesa Press.

Wright, B. D., & Stone, M. H. (1979). *Best test design*. Mesa Press.