Your Light and Air: Leveraging Civic Science to Advance High Need, Grade 6-8 Students' Science Learning Through Investigations

Table of Contents – Narrative

Your Light and Air: Leveraging Civic Science to Advance High Need, Grade 6-8 Students' Science Learning Through Investigations

A. Significance

The New York Hall of Science, in collaboration with the American Institutes for Research, New York City Department of Education School District #24 (NYC-24) in Corona, Queens and Chicago Public Schools, proposes a five-year, early phase research and development project to improve the science achievement of middle-grade students, focusing on high-need students who are underrepresented in STEM career pathways. The project partners will develop and rigorously test **Your Light and Air** (**YLA**), an intervention that will bring two promising strategies to bear on the problem of inequitable access to high quality science instruction for these students. It explicitly introduces students to the purpose and practices of **civic scientific investigation**, an inquiry process that invites learners to draw on their curiosity and prior knowledge to explore and improve the conditions of their everyday environments (Condon & Wichowsky, 2018). It also provides **sustained PD experiences** that prepare science teachers in under-resourced schools to build the instructional competencies and the classroom culture that are necessary to effective implementation of high-quality science learning opportunities (Grigg, Kelly, Gamoran, & Borman, 2013; Scheider et al, 2022; Windschitl, Thompson, Braaten, & Stroupe, 2019).

YLA students will use one of two supplemental curricular units to engage in investigations of the qualities of either the light or the air in their local classroom environment at the beginning of the school year. Each of these units will enable grade 6-8 students to investigate how these phenomena impact their daily lives at school, and how they might enhance the conditions of their classroom environment. This early experience will explicitly introduce students to the purpose, processes, and component practices of scientific investigations, which they are unlikely to have experienced in any depth in the elementary grades (NASEM, 2022). Investigations and the component practices used in them are an important part of curricula that are aligned with Next Generation Science Standards (NGSS Lead States, 2013), and increasingly, state science assessments. Through ongoing professional development, teachers will implement one of these introductory units and revise their implementation plans for other investigations that are part of their existing curriculum. Through participation in civically-focused science investigations, students will build their curiosity about science and engagement in science class, which will lead to increased academic achievement in science, civic efficacy, and belief in the social benefits of science.

A.1: Absolute Priorities Addressed

The project will respond to US DOEd's **EIR absolute priorities #1, Demonstrates a Rationale,** and **#3, Field-Initiated Innovations - Promoting Equity in Student Access to Educational Resources and Opportunities: STEM**. The project will involve schools serving large proportions of high-need, underrepresented students in grades 6-8, who live in low-income communities and attend under-resourced schools. Supplemental curricula will help students discover the relevance, purpose, and practices of scientific investigation. Ongoing PD will help their teachers to implement the supplemental curricula, and adapt their implementation of additional investigations to reinforce and build on lessons learned through the supplements.

A.2 Supporting students' three-dimensional science learning: existing strategies.

Scientific investigation is a critical scientific practice emphasized in the NGSS (NGSS Lead States, 2013). These standards have been adopted by 20 states and the District of Columbia; an additional 24 states have developed standards aligned with the NGSS framework (National Science Teachers Association, 2023). The NGSS were adopted by Illinois in 2014, and New

York established its own highly aligned science standards soon after (New York Education Department, 2016). A central concept of these science standards is "three-dimensional science learning," which employs an integrated approach to teaching science by combining 1) content knowledge, 2) cross-cutting concepts, and 3) scientific and engineering practices (NGSS Lead States, 2013). By the 2024-25 school year, all these states, including New York and Illinois, will have implemented three-dimensional assessments of science learning.

Carefully implemented scientific investigations can contribute to increases in science achievement (Schneider et al, 2022). However, most high-need students attend elementary schools that offer few or no opportunities for active scientific investigations (NASEM, 2022). Consequently, they arrive in middle grade science classrooms with little or no preparation for productive participation in scientific inquiry (Boda & Brown, 2020). For example, in New York City, K-5 students receive only 20 minutes of science instruction per week (NASEM, 2022). Even when these students do experience scientific investigations, they are often implemented in overly structured ways that short-change open-ended exploration, engagement, and the interpretation and application of findings, limiting opportunities for impact on learning (Furtak, 2019; Grigg, Kelly, Gamoran, & Borman, 2013; Schneider et al., 2022).

A civic science education approach to scientific investigation will increase high-need students' engagement in the full cycle of scientific investigation and their motivation to participate productively in that process (Condon & Wichowsky, 2018). Levy, Oliveira & Harris (2021) define civic science education as "educational experiences that support individuals' ability to understand, explore, and take informed action on public issues related to science." This approach makes explicit the positive, reinforcing links among personal relevance, engagement with science, and science achievement, which are well documented in research literature

(Flanagan, Gallay, & Pykett, 2022: Levy, Oliveira, & Harris, 2021). A randomized controlled trial led by Condon & Wichowsky (2018) demonstrated that emphasizing the civic relevance of science as a discipline can increase student engagement and motivation as well as science achievement, particularly for students attending under-resourced schools that provide limited access to high quality science learning experiences.

When effectively implemented, investigations are an entry point to the full spectrum of NGSS scientific practices and to engagement in three-dimensional science learning (Duschl & Bybee, 2014; Haas, et al, 2021). When students learn how to conduct scientific investigations, they recognize and become familiar with scientific standards for evidence; what kinds of things can be studied scientifically; and how to talk constructively with others about the process of understanding how and why a phenomena works (Redish, 2010).

A.3 Supporting students' three-dimensional science learning: promising strategies

A.3.1. An emphasis on a "civic science education" approach to scientific investigation. The YLA curricular units and PD will emphasize civic relevance by supporting students to link exploration of scientific phenomena to their experience of their everyday environments - that is, students can pose and pursue questions not about "how light behaves" in the abstract, but about the qualities of "the light sources in my classroom" or "the sunlight that heats up our asphalt playground area in the spring." Students complete at least one, 2–3-week YLA unit at the very beginning of the school year. They then complete additional Going Deeper lessons later in the school year, chosen by the teacher, to build on these foundational experiences and connect them to investigations already included in their existing curricula.

YLA provides early, explicit experience with civic scientific investigation, engaging students in the study of their classroom light or air in the first weeks of school. This unit introduces students to the process and the purpose of scientific investigation by focusing on everyday experiences of phenomena and building on students' curiosity about them. This initial experience will create and sustain the processes and context for productive student engagement with scientific investigations throughout the school year. Borman and colleagues have demonstrated that very brief interventions intended to influence future approaches to academic tasks, used early in the school year, can have a lasting impact on middle school students' grades and significantly shrink achievement gaps (Borman, Choi, & Hall, 2021; Pyne & Borman, Rozek, Pyne & Hanselman, 2019). Their work suggests that intervening early to re-frame student expectations can have a lasting impact on performance.

A.3.2. Ongoing professional development. Teachers will participate in two forms of PD: a summer workshop and school year online critical friends groups (CFGs; National School Reform Faculty, 2012). This project draws particularly on the Ambitious Science Teaching (AST) framework to inform our PD approach to supporting reflection on instructional practice (Windschitl, Thompson & Braaten, 2018; Windschitl, Thompson, Braaten, & Stroupe, 2019). This work emphasizes giving teachers the tools to support rigorous student thinking and rich student discourse regardless of the particular curriculum they may be required to follow. Originally designed for middle grade teachers in resource-poor schools, the AST model shows teachers how to use manageable, incremental strategies to link curricular goals to students' everyday lives and encourage productive, collaborative classroom discussions and investigations. We will also use CFGs to support their ongoing implementation of YLA, including adapting supplemental instructional tools to support scientific investigations during the school year (Fahey & Ippolito 2015).

B. Quality of the Project Design

B.1. Conceptual Framework

B.1.1. High-need, underrepresented students receive explicit guidance about how to conduct scientific investigations of their local classroom environments. Each unit will contain six lessons mapped to key aspects of middle school scientific investigation described in the NGSS, NYS standards, and CPS' Skyline curriculum: 1) Develop questions about the local classroom environment that can be solved through investigation; 2) Design the investigation, including identifying variables, selecting the tools needed to gather data, and determining how measurements will be captured and recorded; 3) Conduct the investigation; 4) Assess the accuracy of the data and methods; 5) Analyze data and consider implications for what it means for students and their school environment; 6) Use the data as evidence to support explanations or design solutions. Each lesson will take 2-3 class periods to complete. Students will use digital notebooks to share data, compare findings, and document their thinking over time. For instance, the light unit will start by inviting students to notice the qualities of light in a classroom at different times of the day. Students will connect this to changes in temperature during the day. This may lead them to design a protocol for investigating how light impacts temperature throughout the school. Based on the data they collect, students may notice that some areas of their classrooms have light conditions that are different from others, and begin to consider the pros and cons. Is more sunshine energizing, or does it make that area of the room too hot? Once students develop interpretations of their findings, they can also use their data to advocate for changes, such as introducing shading materials to cool a room that has lots of sunshine. Students could then implement changes and conduct a revised investigation.

The YLA units will be accompanied by six additional Going Deeper lessons that are implemented **later in the school year, at times of the teacher's choosing.** Each additional lesson will support students as they delve deeper on one aspect of the initial YLA unit and apply it to other science investigations. For example, one lesson will focus on data collection, and will encourage students to consider how the collection methods they used in the initial YLA unit might inform the methods they choose to investigate new questions in a different context.

The YLA units and Going Deeper lessons will be shaped by feedback from partner teachers. In an initial design cycle during project Year 1, we will recruit 3 teachers each from Chicago and NYC, who will test lessons created by the program team and use ongoing reflective activities on the Participate platform to provide critiques and suggest modifications. In a second design cycle, NYSCI will recruit an additional six teachers, who will help to develop and then pilot the YLA PD. All activities in Cycles 1 and 2 will take place virtually, allowing for cross-city collaboration. In a third design cycle, 18 teachers will field test the full intervention.

B.1.2. Middle grade science teachers learn how to make high-impact changes in their instructional practices while continuing to work as required by their existing school curricula and policies. Grade 6-8 science teachers from NYC-24 and CPS will participate in a 3.5 day, in-person PD in their city, prior to the start of the school year. The PD will draw on the Scale Immersion Model for Professional Learning (SIMPL), which employs an iterative learning process in which teachers first experience the light and air units and investigation practices as learners, carrying out the complete investigations themselves, and then reflect on the experience as educators (Grig, et al, 2013; Kelley, et al 2012; Lauffer, 2010). They will collaboratively analyze the instructional materials and unpack the underlying approach to supporting investigations. They will select which of the two units (light or air) they will implement at the start of the year. They will also begin analyzing and revising their existing curriculum to identify where authentic investigation can be strongly supported. Teachers will return to their classrooms to implement one of the YLA units with their students at the very beginning of the school year. They will use Participate, a dedicated platform for Communities of Practice, (i.e., the Critical Friends Groups) to share their observations and reflections during implementation.

B.1.3. PD supports will guide teachers as they help students get better at defining, designing and carrying out scientific investigations of their local classroom environment throughout the school year, leading to improved science achievement and increased understanding of the civic relevance of science. Over the course of the academic year, teachers will be supported through Critical Friends Groups (CFGs) and supplemental instructional resources as they implement student-led investigations in their classrooms. CFGs will meet virtually once per month during the school year and interact asynchronously on Participate, sharing learning artifacts and observations about student engagement and curiosity during the investigations. They will include 5-7 teachers per group, structured by grade level. A foundation for the work of the CFGs will be the six supplemental Going Deeper lessons. CFGs will work together to select which additional lessons to implement and will modify these materials to meet their local needs. In their CFGs, teachers will critique the lessons, making decisions about where they might fit well into the curricula, how they might be improved, and if the instructional and assessment strategies are appropriate for use in other investigations.

The logic model (see Appendix G) identifies YLA elements and their anticipated impact on student outcomes - increased science achievement, civic efficacy, and beliefs in the social benefits of science - and teacher outcomes - increased focus on NGSS practices, self-efficacy for teaching science, and support for students' civic engagement in science.

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

The goals of the proposed project are 1) to develop and evaluate the YLA intervention, and to 2) maximize the impact of the project by sharing its findings broadly and managing the project responsibly. These goals are tied to objectives and measurable outcomes as follows (see **Evidence Form** for further detail):

Goal 1: Develop and evaluate the YLA intervention

1. Objective 1: Develop, pilot test and field test YLA units and PD
    - 1.1. Iteratively create units and PD for pilot test, in collaboration with teachers (NYSCI). Develop pilot and field test data collection plan, and instruments (AIR)
    - 1.2. Conduct pilot test; analyze data; report findings (AIR, NYSCI)
    - 1.3. Apply findings to revise and critique units and PD for field test (NYSCI, AIR)
    - 1.4. Conduct field test; analyze data; report findings to participants (NYSCI, AIR)

Objective 1 measurable outcomes: ⅔ of teachers recruited for implementation design cycles persist throughout their R+D cycles; Pilot and field test are conducted on schedule and with 5/6 of participating teachers (N=6, 12, and 18 for the three cycles) implementing.

2. Objective 2: Conduct an experiment assessing YLA implementation and impact
    - 2.1. Recruit teachers and schools (AIR)
    - 2.2. Apply field test findings to revise YLA units and PD for experiment (NYSCI)
    - 2.3. Develop RCT data collection and analysis plan and revise instruments (AIR)
    - 2.4. Randomly assign teachers to YLA and control conditions (AIR)
    - 2.5. Collect and analyze data; report to participants and EIR (AIR, NYSCI)

Objective 2 measurable outcomes: Recruitment is completed on schedule; 90% of planned summer workshop and CFG activities are provided; teachers and students meet evaluation thresholds for implementation; data collection and analysis are completed on time and on budget.

Goal 2: Maximize the impact of the project by sharing its findings broadly and managing the project responsibly

3. Objective 3. Develop and implement knowledge mobilization plan

    3.1. Iteratively develop and revise knowledge mobilization plan (NYSCI, AIR)

    3.2. Develop criteria for publicizing and providing access to findings and intervention materials (NYSCI, AIR)

    3.3. Implement knowledge mobilization plan for field test findings and intervention materials (NYSCI, AIR)

    3.4. Implement knowledge mobilization plan for experiment findings and intervention materials (NYSCI, AIR)

Objective 3 measurable outcomes: 2 conference presentations on the field test findings and/or the R+D process; 2 conference presentations, 2 articles published in practitioner journals and 1 in a peer-reviewed research journal on the results of the experiment.

4. Manage project effectively and efficiently

    4.1. Identify teacher R+D partners in New York and Chicago (NYSCI)

    4.2. Hold project kick off meeting with teachers (NYSCI)

    4.3. Convene biweekly project director meetings (NSYCI)

    4.4. Convene performance feedback/progress assessment meetings (AIR)

    4.5. Participate in meetings convened by IES and EIR TA providers (NYSCI, AIR)

Objective 4 measurable outcomes: Team participation in 80% of offered technical assistance (TA) events and EIR-sponsored workshops; 100% participation in EIR PI meetings.

B.3. Addressing Needs of Target Population.

This project responds to Absolute Priority #3 by creating opportunities for high-need, underrepresented students to improve their science learning by increasing their ability to use NGSS practices through sustained scientific investigations of their local environments. The project targets students residing in low-income communities and attending under-resourced schools. These overlapping populations of youth have historically been under-represented in STEM education pathways (NASEM, 2019), are likely to have been marginalized as civic actors in their communities (Flanagan, Gallay & Pykett, 2022). Moreover, they are likely to have attended elementary schools that offered limited access to high quality science learning experiences (NASEM, 2022). These students are also a major proportion of those served by the schools participating in the proposed project, as documented in Table 2, below.

Table 1: Demographic characteristics of students in NYC-24 and CPS

	Total number of students	**Economically disadvantaged students**	**English Learners**	**Students with Disabilities**	**Students under-represented in STEM*****
NYC-24*	49,149	38,529 (78%)	14,409 (29%)	9,220 (19%)	33,615 (68%)
6-8	12,104	9,732 (80%)	2,972 (25%)	2,390 (20%)	8,243 (68%)
CPS**	322,106	234,283 (73%)	72,019 (22%)	49,347 (15%)	265,785 (83%)
6-8	71,585	56,437 (79%)	15,714 (22%)	11,537 (16%)	59,762 (84%)

* https://data.nysed.gov/enrollment.php?year=2022&instid=800000043407

** https://www.cps.edu/about/district-data/demographics/

*** Black/African American, Hispanic, and Native American/Alaskan

These often-overlapping student populations consistently perform at lower levels than their peers on measures of science achievement (National Assessment of Educational Progress [NAEP], 2019). For example, the average performance of eighth grade students who are eligible for the National School Lunch Program (NSLP) on the NAEP science test is consistently significantly lower than that of students who are not eligible for NSLP. Across multiple NAEP administrations since 2005, non-eligible students' average scores are close to the "proficient" score of 170, while NSLP-eligible students' scores averaged at or near the "basic" score of 140.

C. Project Personnel

C.1 Recruitment and staff diversity at NYSCI and AIR

C.1.1 Recruitment at the New York Hall of Science. NYSCI is an equal opportunity employer and has standardized practices to guide hiring that ensure recruitment through diverse networks, consistent interviewing, and objective rating systems. NYSCI recently completed a major benchmarking study of staff job titles and compensation and made significant investments in recommended adjustments that have created greater consistency and transparency in job titles and responsibilities and established equitable compensation standards across departments and job categories. NYSCI is currently engaged in a restructuring of our performance management and review process, to ensure processes for review and opportunities for advancement are broadly accessible and supervision is conducted in an equitable and inclusive manner for all staff.

C.1.2 Recruitment at AIR. American Institutes for Research is an equal employment opportunity/affirmative action employer committed to excellence through diversity. Minorities, women, individuals with disabilities and veterans are encouraged to apply for open positions. All qualified applicants receive consideration for employment without discrimination on the basis of age, race, color, religion, sex, gender, gender identity/expression, sexual orientation, national origin, protected veteran status, or disability. AIR's commitment to diversity is fully integrated into its strategy, operations, and work environment. AIR believes that embracing diverse perspectives, identities and life stories drives innovation and employee engagement.

C.2 Qualifications and relevant training of key personnel (See Appendix B for resumes)

C.2.1 NYSCI Lead Staff: [redacted] Project Director, is NYSCI's Director of Strategic Initiatives and an expert in both standards-based science teacher PD and environmental education. She holds a B.S. in Cognitive Neuroscience from Brown University, and an M.S. in environmental policy, SUNY College of Environmental Science and Forestry. [redacted] research advisor, is the Chief Learning Officer at NYSCI, leading the R+D agenda and educational service programs for schools and teachers. She is a former research director of USDOE/IES-funded Regional Education Laboratory for the Northeast and Islands. She holds a Ph.D. in Developmental Psychology from Teachers College, Columbia University (1999). [redacted] senior project advisor, is President and CEO of NYSCI. She is a member of the National Academy of Education and sets NYSCI's long-term strategy for R+D of innovative tools for STEM learning. She holds a Ph.D. in Developmental Psychology, Teachers College, Columbia University (1988).

C.2.2 AIR Lead Staff: [redacted] (PhD, University of Michigan), principal researcher, will be the Co-Project Director, lead the AIR team and serve as the evaluation's principal investigator, providing intellectual guidance and support for the research and development phase. [redacted] (MA, Georgetown), researcher, will serve as the project manager and lead qualitative researcher for the evaluation, managing project scope, timeline, and personnel and recruitment, data collection, and qualitative analysis. [redacted] (PhD University of Pittsburgh), researcher, will serve as the lead quantitative analyst , managing the design and execution of the RCT and the collection and analysis of outcomes data. [redacted] (PhD, Capella University), senior technical assistance consultant, will serve as the progress assessment lead during the R+D cycles, managing the development and implementation of data collection, analysis, reporting, and findings utilization. [redacted] (PhD, Stanford University), senior director, will serve as the project's measurement lead, managing development and use of implementation and outcomes measures. [redacted] (PhD, University of Illinois Urbana-Champaign), vice president, will serve as the evaluation's senior advisor, providing strategic guidance and quality assurance.

D. Management Plan

D.1 Project partners and personnel: roles and responsibilities

The management of the proposed project builds on NYSCI's and AIR's past partnership on an i3 development study (2012-2017) of *Playground Physics*' impact (Friedman, Margolin, Swanlund, Dhillon, & Liu, 2017), a mid-phase EIR scale-up of *Playground Physics* in New York State, and a 2019 early-phase study of *The Pack* as a middle grade science curriculum intervention.

D.1.1 New York Hall of Science. NYSCI will provide leadership and oversight of all program activities, manage communications among partners and advisors, and lead curricular development and professional development activities. NYSCI is a leading science and technology center and provides STEM education through design-based learning products and services to 300,000 K-12 students per year, and professional learning to approximately 2,000 teachers each year. Our PD offerings range from half-day workshops to intensive, year-long coaching programs and extended institutes during school break weeks and summer terms.

D.1.2 AIR. AIR will lead the independent evaluation. AIR will oversee the study design and execution, monitor quality assurance, and provide methodological direction. AIR has over

75 years of experience evaluating education implementations for local and state education agencies, the U.S. Department of Education, and private sector entities.

D.1.3 Advisors. An **interdisciplinary team of advisors** will guide and support the work of the project team. Advisors include [redacted] Director of Science for the NYC Department of Education; [redacted] 6-12 grade Science Manager for the Chicago Public Schools; [redacted] Professor of Teaching, Learning and Curriculum at the University of Washington; and [redacted] Professor of Education in Science Education at the Pennsylvania State University. Advisors will meet with project staff yearly to guide planning, provide feedback and help to define mid-course modifications.

D.2 Managing to project timelines and budgets

Staff will be organized into a development team (NYSCI) and a study team (AIR), with each team holding responsibility for specific project objectives and tasks (see above). All tasks will be monitored by the project PI. The project director and co-director will hold monthly meetings to monitor progress, recommend direction, and make adjustments. The project director, co-director, and coordinator will meet bi-weekly to discuss implementation. The evaluation team will join check-ins bi-monthly. The development and research teams will hold weekly meetings to focus on teacher collaborations, R+D phases, and teacher implementation, drawing on AIR analyses.

The project will proceed in three phases of work. Phase 1 (Jan 2024-May 2027) will focus on the R+D process (NYSCI), including a field test of the intervention (AIR); monitoring and feedback during the R+D process (AIR); and development and testing of research instruments for the experiment (AIR). Weekly meetings will be used to share and reflect on AIR feedback, connecting evidence to possible adjustments in communications with teachers, design of intervention materials, or other changes. Bi-weekly leadership meetings will focus on implications of changes for timeline and budget. The Project Director and co-Director will also work closely with EIR TA providers and the project program officer, keeping them informed of progress and bringing their recommendations back to the team.

Phase 2 (June 2027-May 2028) will focus on the experiment to test the efficacy of YLA (AIR and NYSCI). NYSCI will support implementation, provide the summer workshops and lead the school year Critical Friends Groups (CFGs). AIR will conduct random assignment, collect and analyze data and report out on findings. Weekly meetings will focus on responding to any challenges at the building or classroom level that are impeding implementation, and ensuring timely and complete data collection. In Phase 3 (June, 2028-December, 2028), the team will implement a plan for knowledge mobilization that will reflect lessons learned through the R+D process and the experiment. Throughout the life of the project, the team will participate actively in TA and dissemination opportunities provided by the Institute of Education Sciences.

E. Evaluation of YLA

The American Institutes for Research® (AIR®) will conduct an independent evaluation of YLA. The evaluation will include an implementation study designed to provide feedback on program components and support continuous program improvement and a one-year randomized controlled trial (RCT) designed to meet What Works Clearinghouse (WWC) standards without reservations that will provide causal evidence of YLA's impact on students' science achievement, civic efficacy, and their belief in the social benefits of science. The research questions and data sources are shown in Exhibit 1 below (See Appendix J6 and J7 for timelines).

Exhibit 1. Research Questions and Data Sources for the Evaluation of YLA

Research Questions	Data Sources
Impact Research Questions (Confirmatory)	

<table>
<tr><td>RQ1. What is the impact of participation in YLA’s on students’ (a) science achievement; (b) civic efficacy; and (c) beliefs in social benefits of science?</td><td rowspan="5">- NGSS assessment
- State science assessment
- Science grades
- Teacher and student pre-/post-surveys
- Teacher lesson log
- Student and teacher science learning profile
- YLA extant data[1]
- District administrative data[2]</td></tr>
<tr><td>RQ2. What is the impact of participation in YLA’s on teachers’ (a) focus on and use of NGSS practices; (b) self-efficacy as providers of high-quality science learning opportunities; and (c) support for students’ civic engagement in science?</td></tr>
<tr><td>Moderator and Mediator Research Questions (Exploratory)</td></tr>
<tr><td>RQ3. To what extent are the impacts of YLA on student and teacher outcomes moderated by student, teacher, and school characteristics?</td></tr>
<tr><td>RQ4. To what extent are program impacts on student (a) science achievement; (b) civic efficacy; and (c) beliefs in social benefits of science mediated by teacher outcomes, student engagement, and students’ curiosity and interest?</td></tr>
<tr><td colspan="2">Implementation Research Questions</td></tr>
<tr><td>RQ5. To what extent and in what respects is YLA implemented with fidelity?</td><td rowspan="2">- Teacher R+D log
- Teacher interviews
- Teacher and student focus groups</td></tr>
<tr><td>RQ6. What are educators’ and students’ perceptions and experiences of usability and value of YLA units and educators’ perceptions of the useability and value of professional development?</td></tr>
</table>

RQ7. What are barriers and facilitators for students and teachers to participating in the YLA program?	- Teacher and student pre-/post-surveys - Teacher lesson log - Student and teacher science learning profile - YLA extant data[1] - District administrative data[2]

E1. Evidence That Meets WWC Standards Without Reservations

AIR will produce strong evidence regarding the effectiveness of YLA on students' science achievement, civic efficacy, and belief in the social benefits of science using an experimental design designed to meet WWC evidence standards without reservations. AIR will address RQs 1 and 2 using a school-level randomized controlled trial (RCT) with blocking at the district-level. AIR and NYSCI will recruit 44 schools with grades 6-8 across two large, urban school districts (CPS and NYC-24) to participate in the experiment in the 2027-2028 school year (see Appendix C for letters of support). Both districts enroll at least 70% of students eligible for free or reduced-price lunch, 20% English Learners, and at least 65% of students from racial and ethnic groups historically under-represented in science and engineering (e.g., Black, Hispanic, and American Indian or Alaskan Native students; see NCSES, 2023). Schools will be randomly assigned within each district with equal probability to participate either in YLA (the treatment group) or in a control group before the 2026-2027 school year.

[1] YLA use generates extant data, such as logins, documents, spreadsheets, and student and teacher notebooks.
[2] Districts will provide demographic and performance data on students, teachers, their schools.

Mitigating potential threats to internal validity. Due to school-level assignment, this study design poses little risk of contamination or bias due to "joiners" (in this instance, students entering the treatment classrooms after random assignment) per recent WWC guidance (What Works Clearinghouse, 2022). There is little reason to believe that YLA activities would influence students to join study schools because the information about which schools are participating in the study and in which condition will not be publicly available. Further, if joiner quantities and patterns warrant, we will drop joiners from the analysis. There is a potential risk of bias due to attrition at the school or individual (teacher or student) level. To mitigate this risk, AIR and NYSCI will collaborate to employ several strategies including providing incentives for teacher and district assistance with data collection, planning consent and survey completion events with participating schools, coordinating follow-ups, and collecting initial study data within a few months of randomization. School-level attrition from changes in school leadership or instructional priorities will be minimized by having schools participate in the evaluation for only one school year, rather than in a multiyear intervention. To examine whether bias has been introduced via attrition, the team will compute overall and differential attrition at the school and individual (teacher and student) level and test for baseline equivalence in school, teacher, and student characteristics for the impact analysis sample. Missing covariate data will be addressed using multiple imputation by chained equations (Royston, 2009). See Appendix J2 for details on the treatment of missing covariate and outcome data, and strategies for receiving a rating of *WWC Standards With Reservations* in the event of high overall or differential attrition).

Sample and power. With a sample of 11,660 students and 132 teachers (assuming a harmonic mean (WWC, 2022) of 265 students and three Grade 6-8 science teachers per school), the minimum detectable effect size (MDES) is 0.22 for student level outcomes and 0.40 for teacher level outcomes (see Appendix J1 for power analysis assumptions). Accounting for 20% attrition at the school and student level, the study is still powered to detect an MDES of 0.27 for student-level outcomes and 0.50 for teacher-level outcomes. This level of estimated individual attrition is appropriate given our approaches to mitigate attrition and reviews of educational RCTs that show an average attrition of 19% (Weidmann & Miratrix, 2021). We will explicitly test for bias introduced from attrition or missing outcome data and use multiple imputation or matching methods to mitigate these risks if found to be above the conservative WWC threshold (What Works Clearinghouse, 2023; see Appendix J2 for additional details). These estimated effect sizes are appropriate based on the literature about effects of science interventions on middle school outcomes (Taylor et al., 2018; Thomas & Larwin, 2023), and effects of interventions of teachers' instructional practices (Garrett et al., 2019).

Outcome measures.

Student science achievement outcomes. Student science achievement will be measured through students' prior year and end-of-course science grades using a five-point ordinal scale (0.0 = 0-59 or "F", 1.0 = 60-69 of "D", 2.0 = 70-79 or "C", 3.0 = 80-89 of "B", and 4.0 = 90-100 or "A"). When districts or schools use different or fractional grading scales, the evaluation team will standardize these measures to ensure that outcomes are comparable prior to analysis (What Works Clearinghouse, 2022; 2023). We also will collect state standardized science tests from New York and Illinois to measure science achievement of students in Grade 8.[3] Both tests measure performance relative to state standards aligned with NGSS. Finally, to measure student understanding of science content, students of treatment and control teachers will complete a 20-

[3] State achievement tests in science are only available for 8th grade students in NYC and CPS, and therefore these analyses will have a different minimally detectable effect size (MDES) of 0.191. See Appendix J.4 for additional details.

item assessment aligned with NGSS practices and with evidence of validity with students in Grade 4 through 8 (Yang, He & Liu, 2017). This measure has shown Rasch person reliability of 0.62, which is comparable with Cronbach's alpha but usually represents a lower (conservative) bound of reliability (see Bond & Fox, 2007).

Student survey outcomes. In Fall 2027 and Spring 2028, students of treatment and control teachers will complete an online survey to measure four constructs. A six-item measure of *belief in social benefits of science* (Belanger et al., 2020; α=.66) captures students' perceptions about how science and scientists can improve society and foster connections. A three-item measure of *civic efficacy* (Syvertsen et al., 2015); α=.84 gauges students' beliefs that they can make a difference in their communities. Two measures of potential mediating variables are: an eight-item measure of *science curiosity and interest* (Chung et al, 2016a; α= .86) and a four-item measure of *engagement in science learning activities* (Chung et al, 2016b; α= .80).

Teacher survey outcomes. In Fall 2027 and Spring 2028, treatment and control teachers will complete an online, 47-item survey of three scales – *focus on and use of NGSS practices* (24 items, Hayes et al., 2016; α=.80-.88); *self-efficacy as providers of high-quality science learning opportunities* (13 items, Rigg & Enochs, 1990; α=.90); and, *support for student civic engagement in science* (10 items adapted from Cole et al., 2016; α=.95). See Appendix J4 for descriptions of student and teacher survey measures.

Analyses of impact. Our primary approach to determining the impact of YLA on student outcomes (RQ1) is an intent-to-treat (ITT) analyses, including students and teachers from randomized schools regardless of whether the schools fully participate. Analyses of student outcomes will use a three-level model that nests students within teachers and schools, includes district (randomization block) fixed effects, and controls for school (e.g., proportion of students eligible for free and reduced-price lunch, proportion of EL students, proportion of students from underserved race/ethnicities) teacher credentials (e.g., years of teaching experience, degree earned), and student level covariates (e.g., race/ethnicity, gender, English proficiency, eligibility for free and reduced-price lunch) to improve precision of impact estimates (see Appendix J2 for details on impact models). Controlling for randomization blocks ensures that only classrooms within the same school districts are compared with one another, which will reduce alternative explanations for estimated impacts (e.g., other initiatives in those districts). For determining the impact of YLA on teacher–level outcomes (RQ2), we will conduct ITT analyses using a two-level model that nests teachers within school, includes randomization block fixed effects, and controls for school and teacher characteristics. We will also conduct a series of exploratory analyses to examine the extent to which student, teacher, and school-level characteristics moderate the effect of *YLA* on student and teacher outcomes (RQ3), and exploratory mediation analyses to examine the extent to which the effects of *YLA* on students' academic outcomes are mediated by teacher outcomes, students' curiosity and interest in science, and science engagement (RQ4). Moderation analyses will modify the main impact analyses by including an interaction term between treatment status and each moderator variable. Mediation analyses will use the main impact models and models predicting the mediator variable to estimate indirect effects and calculate the percentage of the effect mediated as the total effect divided by the indirect effect (see Appendix J4 for additional details).

E2. Performance Feedback and Periodic Assessment of Progress

AIR will facilitate performance feedback and progress assessment meetings throughout the project using implementation data to address RQs 5 (implementation fidelity), 6 (teachers' and students' perceptions and experiences of YLA), and 7 (barriers and facilitators to YLA participation). During the three R+D cycles preceding the experiment, the meetings with NYSCI and R+D partner teachers will occur monthly during summer development periods and bi-monthly during the school year. During the experiment, AIR will meet with NYSCI to facilitate its use of implementation data for performance feedback and progress monitoring.

Data Sources. Data sources addressing RQ 5 (implementation fidelity) for the R+D cycles and the experiment are: (a) *YLA extant data* generated by treatment teachers' participation in in-person and virtual PD and their students' use of a YLA unit's digital resources and (b) *teacher lesson logs* completed for each lesson in the YLA unit (Kurz et al., 2014; Glennie, Charles, & Rice, 2017). Additional implementation data sources addressing RQs 6 (teacher and student perceptions and experience) and 7 (barriers and facilitators) in the R+D Cycles and looking deeper into YLA feasibility and usability are: (c) weekly online *R+D log entries* by teachers and NYSCI team members that document and assess the strengths and weaknesses of their YLA development work; (d) virtual *interviews* at the end of each R+D cycle with six teachers[4] about their assessments of YLA core components and outcomes; (e) two virtual, end-of-year *focus groups of teachers, one for teachers in each of the two cities,* in the second and third R+D cycles about the feasibility and value of YLA PD and (f) six virtual, grade level-specific end-of-year, student *focus groups, one in each city for each grade,* evaluating the feasibility, engagement, and value of YLA units.

Analyses. AIR will conduct ongoing descriptive analyses of YLA extant data and teacher lesson logs. The analyses will measure progress toward acceptable implementation thresholds for PD, such as the number of workshops or CFG sessions provided and the number of participants

---

[4] The six will be the six first R+D cycle teachers, a random three from each location of the 12 second cycle teachers and a random three from each location of the 18 third cycle teachers.

generating more NGSS-focused lesson plans. Implementation thresholds for units may include the number of YLA lessons provided and the number of students entering light quality data into a shared spreadsheet. (See Section E3 and Appendix J3 for more information on implementation fidelity indicators and thresholds.) AIR will conduct thematic and sentiment analyses of the weekly R+D log entries, interviews, and focus groups. (See Appendix J5 for details.) These analyses will identify the YLA components (themes) that teachers and students think do and do not work well (sentiment) and their reasons for thinking so (themes).

During performance feedback and progress assessment meetings at the beginning of summer, the NYSCI team and the R+D teachers, facilitated by AIR, will use findings from the interview and focus group analysis to identify YLA components and component parts for further development and improvement and findings from the R+D log analyses to strengthen YLA development. They will use findings from the analyses of the YLA extant data and teacher lesson logs to strengthen CFG activities in support of reaching YLA implementation thresholds.

E3. Articulates Key Project Components, Mediators, Outcomes, and Acceptable Implementation Thresholds

Project components. The evaluation design is aligned to clear YLA components, mediators, and outcomes, as depicted in the logic model (see Appendix G) as well as in the research questions and data sources to address them (see Exhibit 1). YLA has two key components: professional development for participating teachers consisting of the summer workshop and the school year Critical Friend Groups and the YLA unit for their students, consisting of the six lessons at the beginning of the school year and six Going Deeper lessons thereafter. Together, these components are expected to improve teachers' use of NGSS practices, self-efficacy, and support for students' civic engagement in science. These teacher outcomes are hypothesized to mediate the impact of YLA on student outcomes (RQ4). Similarly, two proximal student outcomes, curiosity about science and engagement in science class and stronger science investigations, are hypothesized to mediate the impact of the YLA program on three distal student outcomes: science achievement, civic efficacy, and beliefs in societal benefits of science (RQ4). (See Appendix J2 for details about mediation analyses.)

Acceptable implementation thresholds. The YLA extant and teacher lesson log data will be the basis for implementation fidelity measures and thresholds for the core components of YLA PD—summer workshop and CFGs—and the YLA unit—unit lessons and Going Deeper lessons. NYSCI and AIR have established the following core set of implementation indicators and thresholds, which we will refine with partner teachers during the R+D cycles.

Exhibit 2. Initial, Core Implementation Indicators and Thresholds

YLA PD	YLA Unit
- 90% of workshop activities provided - 90% of CFG activities provided - On average, teachers participate in 80% of workshops and 80% of CFG activities - On average, a teacher produces 80% of workshop and CFG activity artifacts	- Unit resources provided to 100% of teachers - On average, teachers teach 80% of unit lessons and 66% of Going Deeper lessons - On average, a teacher's students complete 75% of lesson assignments

As part of the three R+D cycles, AIR will facilitate the team's iterative comparison of YLA indicators and thresholds to similar middle school science interventions meeting the WWC standards with and without reservations (Hill et al, 2023; see Appendix J5 for details).