Bridging the Translational Gap

A Research-Practice Partnership in Equity-Focused Informal STEM Education

Kristin Michod Gagnier, Megan C. Sharkey, Laura Holian, Melissa Kurman, Maya Heiland, & Steven J. Holochwost

Informal science, technology, engineering, and mathematics (STEM) education includes learning that occurs outside formal school settings, such as in museums, science centers, afterschool programs, and other community initiatives (NRC, 2009).

A growing body of research highlights the critical role these informal experiences play in expanding participation in science, cultivating lifelong engagement, and strengthening scientific literacy (NRC, 2009; Falk & Dierking, 2010; Dierking et al., 2013; Stocklmayer et al., 2010; Sun, 2024). Experiences in these settings support authentic scientific practices, such as questioning, investigating, analyzing data, and considering the broader implications of findings, helping learners build the skills needed to participate in science beyond the classroom (NRC, 2012). Informal education also

KRISTIN MICHOD GAGNIER is a senior research scientist with AnLar LLC and director of preschool development grant in the Division of Early Childhood at the Maryland State Department of Education.
MEGAN C. SHARKEY, research associate, and **LAURA HOLIAN**, research director, are with AnLar LLC.
MELISSA KURMAN is the senior manager of impact and evaluation at the University City Science Center. Previously senior director of STEM education at the University City Science Center.
MAYA HEILAND now serves as the director of apprenticeships and school-to-work at the Community College of Philadelphia.
STEVEN J. HOLOCHWOST is an associate professor in the Department of Psychology at Lehman College and the Graduate Center of the City University of New York.

advances equity by expanding access to meaningful STEM learning opportunities through programs that are culturally relevant, accessible, and designed to meet diverse needs, which supports identity development, motivation, and persistence (Bell et al., 2016; NASEM, 2018). These experiences can also help address inequities by intentionally expanding students' STEM social networks through mentoring, exposure to STEM professionals, and structured connections within informal learning environments, which in turn support broader participation in STEM (Saw, 2020). Partnerships with STEM professionals provide authentic, project-based experiences that mirror scientific practice, helping learners see science in everyday life and envision themselves in future STEM careers (Dabney et al., 2012).

Despite the potential promise of informal experiences, a persistent challenge in STEM education lies in the translational gap: How do we ensure that research-informed, equity-focused programs can effectively scale and sustain themselves within a complex STEM ecosystem? Too often, excellent programs fail to bridge the divide between research findings and on-the-ground practitioner knowledge (Joyce & Cartwright, 2020). The intentional creation of research–practice partnerships (RPPs) is therefore critical for generating and sustaining high-quality, impactful programs. The RPP presented here intentionally designs program structures, recruitment strategies, learning experiences, and research processes to expand access to meaningful participation, support identity development, and reduce barriers within the context of an informal STEM program.

> How do we ensure that research-informed, equity-focused programs can effectively scale and sustain themselves within a complex STEM ecosystem?

We present a case study of the partnership model used in the Experiencing STEM FirstHand project, a collaboration between researchers at AnLar and the City University of New York and practitioners at the Philadelphia-based FirstHand™ program. We begin by situating the program within its local context as a reflection of broader systemic challenges and opportunities. We then describe how our RPP was intentionally structured around the five dimensions of RPP effectiveness (Henrick et al., 2017; 2023) needed for participatory evaluation and the ongoing refinement of program design and implementation. Using examples from our work, we show how these dimensions were translated into concrete actions that strengthened implementation and evaluation. Summative findings from this project will be reported in a future article; our aim here is to offer an example for other teams engaged in partnership work and to illustrate how grounding activities in the five RPP dimensions can generate insights that drive improvement.

Context and Rationale

Despite national efforts to diversify the STEM workforce, Black and Hispanic individuals remain persistently underrepresented (Fry et al., 2021; NASEM, 2023). Increasing diversity in STEM is both an equity issue and an innovation imperative, as a more inclusive workforce can reduce economic disparities while bringing new perspectives that drive creativity and discovery (Hong & Page, 2004; Page, 2025; Taylor, 2024). These inequities stem from long-standing gaps in access to educational opportunities across the STEM pipeline. National data show that schools serving mostly White students are more likely to offer advanced coursework and employ highly qualified teachers than schools serving a majority of Black or Hispanic students (U.S. Department of Education, 2025). As a result, many students are denied access to gatekeeper courses, high-quality instruction, and hands-on learning experiences that support entry into STEM fields. These disparities persist into higher education: Black and Hispanic students enter STEM majors at rates similar to their White peers but are significantly less likely to complete their programs, with introductory "gateway" courses playing a disproportionate role in attrition (Riegle-Crumb et al., 2019; Hatfield et al., 2022). These compounding inequities underscore the need for intentionally designed opportunities earlier in students' educational trajectories.

Philadelphia reflects many of these national patterns. Although research shows that racially and socioeconomically diverse schools benefit all students (The Century Foundation, 2019), segregation in Philadelphia remains high (Slaughter et al., 2016), as it does in many U.S. cities. The School District of Philadelphia serves a predominantly Black and

Hispanic student population, yet persistent inequities remain in access to rigorous coursework and hands-on STEM learning. For example, algebra, a key gateway to advanced mathematics and science, is offered in only 50 of the district's 195 K-8 schools (Mezzacappa, 2023). Achievement gaps further reflect unequal opportunity: In 2024, only 26.1% of economically disadvantaged eighth-grade students scored proficient in science on the Pennsylvania System of School Assessment, compared to 31% of all district eighth graders, and only 16% met proficiency on the statewide high school biology exam required for graduation compared to 21% of all district students. Access to advanced coursework is also uneven: White students are nearly three times more likely than Black and Hispanic students to take advanced placement (AP) courses (Bloomquist & Rosenfeld, 2021), and 27% of high schools offer no AP courses (NCES, 2024). These gaps limit early opportunities for authentic, inquiry-based STEM learning.

At the same time, Philadelphia is a major hub of STEM innovation, with world-class universities, hospitals, and a rapidly growing life sciences sector (Vey et al., 2017; Burden et al., 2025). Demand for STEM talent continues to grow, yet many of the city's youth, particularly those from historically underserved communities, have limited pathways to these opportunities. This contrast reflects broader national challenges, in which inequities in access persist alongside expanding economic opportunity. Recent research suggests that addressing access alone is insufficient; meaningful participation also depends on whether students experience STEM as relevant, inclusive, and connected to their identities (Zhao et al., 2022; YESTEM, 2021). Constructs such as belonging, identity, and representation, including the relatability of role models, have been shown to play a central role in shaping engagement and persistence (Aish et al., 2018; Edelen et al., 2024). In a city like Philadelphia, where opportunity and inequity coexist, these findings underscore the importance of designing STEM experiences that not only expand access but also support how students see themselves within STEM and their potential pathways into the local innovation economy.

Taken together, this literature suggests that expanding access must be coupled with attention to identity, belonging, and representation to support sustained engagement in STEM. Bridging these gaps requires not only meaningful STEM experiences but also approaches that align research, practice, and local context to support continuous improvement and responsiveness to community needs. RPPs are particularly well suited to this work, as they enable researchers and practitioners to collaboratively design, study, and refine programs in real-world settings. The FirstHand program was developed within this environment to expand access to authentic STEM learning and support sustained engagement among students who have historically had limited opportunities to participate in such experiences (DeWitt et al., 2025). The Experiencing STEM FirstHand project extends this work through an RPP intentionally structured around established dimensions of partnership effectiveness, integrating evaluation with implementation to generate actionable insights and support ongoing program refinement.

Expanding access must be coupled with attention to identity, belonging, and representation to support sustained engagement in STEM.

The FirstHand Program

Launched in 2014, FirstHand was created to address systemic barriers that limit access to meaningful STEM experiences for students in Philadelphia schools by providing no-cost, high-quality opportunities to engage in authentic scientific practice. Its mission is to offer hands-on, mentored STEM experiences that spark curiosity, strengthen identity, and help students envision themselves as future scientists, engineers, and innovators.

The program serves students from historically underresourced Philadelphia schools. School partners are selected on the basis of student need, proximity to the FirstHand lab, and level of resource gaps, with every participating school serving a student population that is 100% economically disadvantaged. Many of these schools face significant structural constraints. According to the district's school profiles and facilities data (https://schoolprofiles.philasd.org), most partner schools are rated as "poor" or "unsatisfactory" in building condition, and only one is rated as "fair" for supporting teaching and learning. These conditions reflect broader resource inequities that also shape access to STEM learning opportunities, including recent district proposals to close or relocate multiple schools, including some FirstHand partners.

The program works closely with each school to develop a recruitment plan best suited to its students. For example, some schools send their entire eighth-grade class, whereas others select a subset of students who may not thrive in the classroom but who the schools believe would benefit from the program. This ensures that the program does not overlook students based solely on academic merit and offers an opportunity for all students to participate at the school's discretion. The program reviews participation data annually to ensure that participants mirror the communities that FirstHand intends to serve. The participant population has historically been 83% Black, 6% Hispanic, and 6% multiracial, with approximately 9% of students speaking a language other than English at home.

To address opportunity gaps, the FirstHand program provides access to high-quality STEM programming and professional-grade lab environments that are not available in many partner schools. The program follows a co-curricular, place-based model that integrates mentoring, inquiry-driven learning, and community engagement. It brings together students from Philadelphia schools, local STEM professionals, and program staff in a professional laboratory space located in uCity Square, an innovation district in West Philadelphia. This local embedding connects students to nearby industries and careers, helping them see STEM pathways as part of their own communities.

The Experiencing STEM FirstHand project focuses on evaluating the middle school programming delivered in partnership with these schools. Students participate in 10-week sessions, attending weekly two-hour experiences in small groups of approximately 14 students, typically accompanied by a teacher. During sessions, students engage in hands-on investigations and labs while receiving mentorship from working STEM professionals who provide career exposure and guidance. The curriculum includes six thematic tracks aligned with school science priorities and is informed by local industry contexts and regional STEM employers. For example, in "DNA Selfie," students explore DNA structure and function using microscopy and gel electrophoresis, tools used in biotechnology settings. In "Polymer Play," students design and test biodegradable materials while engaging in engineering design practices.

> During sessions, students engage in hands-on investigations and labs while receiving mentorship from working STEM professionals who provide career exposure and guidance.

Each week combines guided instruction with collaborative project work, emphasizing inquiry-based learning and real-world relevance. The program culminates in a public showcase in which students present their projects, explain the problems they addressed, and share their work with mentors, educators, families, and caregivers. These experiences support students in developing communication skills, scientific reasoning, and STEM identity by connecting scientific practice to personally and socially meaningful contexts.

Experiencing STEM FirstHand: A Case Study of Research–Practice Partnerships

Early in its journey, the FirstHand program recognized the value of collaborating with researchers to strengthen program design and understand program impact. Over the past 11 years, the program has collected multiple types of data to learn from student experiences; its original research collaborator remains a core partner, providing continuity, shared learning, and deep insight into the program's evolution. In 2021, the Experiencing STEM FirstHand project was awarded an Early-phase Education Innovation and Research (EIR) grant from the U.S. Department of Education, enabling a deeper examination of whether the program is delivered as designed, how it can be improved, and what benefits it offers participants. This kind of in-depth, multifaceted inquiry depends on a strong RPP—a long-term, mutually beneficial collaboration that generates research on problems of practice and promotes the use of rigorous evidence (Coburn et al., 2013). This collaborative approach is aligned with calls for culturally responsive and participatory forms of evaluation that center on participant experiences and local community goals, rather than focusing solely on measuring intervention outcomes (NASEM, 2026).

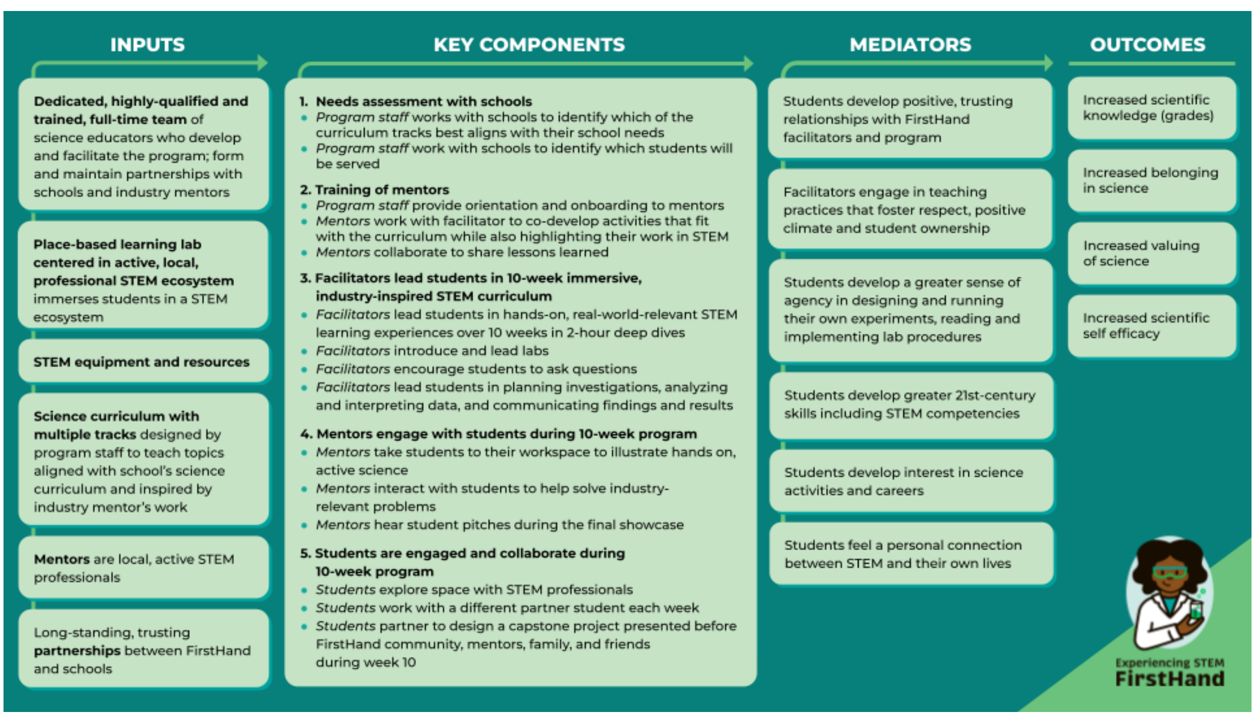
In forming our RPP, we invested in building a genuine partnership through trust building, shared decision making, and co-design of every stage of the research process, from refining a logic model to agreeing on and revising data collection protocols. Our work was guided by the five key dimensions of an effective RPP (Henrick et al., 2017; 2023): (1) cultivating trust and relationships; (2) engaging in inclusive research to address local needs; (3) supporting practice in making progress toward program goals; (4) engaging with the broader field to improve practices, systems, and inquiry; and (5) fostering ongoing learning and developing infrastructure for partnering. In this article, we use specific examples to describe how we adhered to each dimension to strengthen the program, inform continuous improvement, and advance equity-focused STEM learning.

Dimension 1: Cultivating Trust and Relationships

We built trust and strong relationships through three interconnected commitments. First, we developed a *shared vision* of providing equitable science experiences for students, beginning with the collaborative creation of a logic model (Figure 1) that outlined the program's key components, activities, and intended outcomes. This work clarified our goals and established a foundation for using data to continually refine the program. Second, we committed to a *sustainable partnership* grounded in mutual trust, respect, and the understanding that each team member's expertise is essential to our collective aims. Third, we emphasized *transparency and consistent communication* through regular check-ins and open dialogue to ensure that all voices were heard. Co-developing the logic model allowed us to articulate a shared vision, build understanding across roles, and guide our evaluation work.

The logic model shown in Figure 1 presents the project's vision across four levels. *Inputs* are the resources and partnerships that make the program possible. *Key components* are the core activities driving implementation: (1) needs assessment with schools; (2) mentor training; (3) facilitator-led curriculum; (4) mentor engagement; and (5) student engagement. *Mediators* are the changes expected from program participation, such as relationships, agency, and skill development, showing how key components lead to outcomes. Finally, *outcomes* are the ultimate goals of participation in FirstHand. Each partner's expertise shaped the logic model, reinforcing shared ownership and helping us establish a common language and a culture of transparent communication that continues to support our work.

Figure 1. The Experiencing STEM FirstHand logic model



Dimension 2: Implementing Inclusive and Rigorous Research to Understand Local Needs

Mission and RPP Approach: Within our RPP, research questions, design, and interpretation are co-developed with those directly involved in and affected by the program, including FirstHand staff, school partners, mentors, and students. This inclusive approach centers on participant experiences, reflects diverse perspectives, and produces findings that are meaningful and usable in practice.

Identifying Local Needs: Local needs were identified through ongoing engagement with partners, including early conversations with teachers, administrators, and local businesses; analysis of participation and outcome data; and continuous feedback from educators, mentors, and students. Across these inputs, three key needs emerged:

1. Expanding equitable access to high-quality STEM learning opportunities, particularly for students historically underrepresented in STEM
2. Strengthening STEM confidence, belonging, and identity
3. Ensuring that STEM experiences are relevant to students' lives and local contexts.

Evaluation Design and Framing: The evaluation is designed to reflect these locally identified needs and priorities. Consistent with our RPP approach, research questions, measures, and methods were co-developed to ensure relevance and inclusivity. This includes attention to both academic outcomes and equity-centered constructs such as STEM identity, belonging, and engagement. The evaluation integrates quantitative and qualitative methods to capture both measurable outcomes and the lived experiences of participants.

Consistent with our RPP approach, research questions, measures, and methods were co-developed to ensure relevance and inclusivity.

Rigorous Impact Evaluation: To generate credible evidence of program effectiveness, the evaluation includes a multi-cohort randomized controlled trial (RCT) with a delayed-treatment design. This approach allows the partnership to estimate the causal impact of participation on scientific knowledge, science self-efficacy, valuing of science, and belonging. Scientific knowledge is measured using district science grades. A validated suite of measures is used to assess science self-efficacy, valuing of science, and belonging in science (Panorama Education, 2024).

While the summative evaluation is ongoing, formative insights and lessons learned from implementation are shared publicly through the project website (https://stemfirsthand.anlar.com/), providing timely access to emerging learning for practitioners, partners, and the broader informal STEM education field.

Qualitative and Embedded Approaches: Qualitative methods, including student focus groups, interviews, and observations, complement the quantitative measures by capturing how students experience the program and how participation influences their confidence, interest, and identity. These approaches center student voices by allowing participants to describe their experiences in their own words, providing critical context for understanding program impact. To support high-quality, contextually grounded data collection, local research team members are embedded in the program and attend sessions regularly. Their consistent presence enables relationship building with students, staff, and partners, fostering trust and encouraging authentic, in-depth student sharing. Integrating qualitative and quantitative data strengthens the depth, validity, and contextual relevance of the findings while ensuring that participant perspectives directly inform program refinement.

Dimension 3: Supporting Progress Toward Program Goals

Whereas Dimension 2 describes how the evaluation was designed, Dimension 3 focuses on how the resulting data are used. The partnership organized this work around three questions:

1. Implementation fidelity: Is the FirstHand program being delivered as designed?

The FirstHand program was intentionally designed to meet locally identified needs. Evaluating implementation fidelity ensures that program activities remain aligned with these goals. The research team worked with FirstHand staff to define indicators for each core component and establish thresholds for adequate implementation.

- *Program Observations:* Structured observations are conducted using the STEM Program Quality Assessment (STEM PQA; Smith, 2012), an adaptation of the Youth Program Quality Assessment for informal STEM settings.
- *School Leader Conversation:* Semistructured interviews with principals and school liaisons examine how schools align program participation with local priorities and select students, while also providing insight into broader contextual factors such as curriculum shifts and district expectations.
- *Mentor Preparation and Development:* Mentor training is assessed through training observations, staff discussions, and an annual mentor survey, capturing how mentors are prepared to engage students and connect learning to authentic STEM practices.

To support partners and strengthen inclusivity, the team shares customized participation snapshots with each school after every program cycle.

Fidelity data are tracked across program iterations and schools, summarized, and shared with the FirstHand team. Regular review of these data supports alignment with program goals and informs ongoing refinement.

2. Continuous improvement: How can the program be improved?

Formative evaluation provides actionable insights to support continuous improvement and ensure that the program remains responsive to local needs. Data collection strategies were collaboratively developed with input from staff, school partners, and participants. Multiple data sources have been used to identify strengths, challenges, and opportunities for refinement:

- *Program Quality Assessment:* In addition to supporting fidelity, the STEM PQA assesses program quality across four domains: Supportive Environment, Interaction, Engagement, and STEM Skills. Findings are shared annually to highlight strengths and guide improvements.
- *Stakeholder Feedback:*
  - Student focus groups explore experiences of belonging, confidence, and interest, as well as suggestions for improvement.
  - Facilitator interviews examine onboarding, coaching, and implementation challenges.
  - Mentor surveys and interviews capture engagement strategies and perceived benefits.
  - Industry partner interviews assess participation value and conditions for sustained engagement.
  - Informal conversations with teachers, caregivers, and principals provide additional perspectives on program quality and classroom relevance.

To support partners and strengthen inclusivity, the team shares customized participation snapshots with each school after every program cycle. These summaries include student quotes, participation metrics, and engagement data, making findings meaningful and relevant to each school's context and supporting shared decision making.

3. What are the benefits of participation?

The RCT evaluation is designed to measure the impact of the program on student science grades as well as student beliefs and attitudes. The RPP is also working to continue data collection to observe students into high school and postsecondary experiences. The short-term impact study shows the effect of the program on immediate changes in students. In addition to studying the quantitative data, the evaluation team talks to students and teachers to understand their perspectives on how they see the program changing student experiences.

Findings from both the impact evaluation and qualitative data are used to inform program refinement, strengthen mentorship, and align implementation with FirstHand's goals. Quantitative results provide evidence of program effectiveness, whereas qualitative insights help explain how and why outcomes occur.

Dimension 4: Engaging the Broader Field to Strengthen Practices, Systems, and Inquiry

To extend the impact of our work beyond the Experiencing STEM FirstHand project, we prioritize sharing knowledge and processes—including findings, strategies, frameworks, and tools—with external audiences. This engagement is reciprocal, involving listening, learning, and adapting approaches to strengthen practice across settings. We actively engage the Philadelphia community through shared events, conferences, and partnerships. For example, the project team organized a symposium at the American Educational Research Association's annual meeting to explore how Philadelphia programs aimed at diversifying the STEM workforce pipeline evaluate their work. We also connect with informal science education networks, workforce collaboratives, afterschool alliances, educators, parents, and caregivers.

Locally, FirstHand contributes to Philadelphia's STEM ecosystem by collaborating with organizations to expand high-quality STEM opportunities. The School District of Philadelphia's (SDP's) Office of Strategic Partnerships hosted student, staff, and partner workgroups to co-develop an Enrichment Theory of Change, aligning goals, measures, and strategies across district initiatives. FirstHand staff worked with representatives from 19 organizations to inform this process and shared the STEM PQA as a formative tool. The resulting theory of change (SDP, 2024) recommends the PQA to strengthen quality improvement across district programs, illustrating how partnership and knowledge exchange can drive broader systems change. Insights from this RPP also inform classroom practice. Teachers who attend FirstHand with their students report increased confidence with lab-based instruction and regularly apply FirstHand strategies, such as end-of-class reflections, to enhance engagement and deepen learning. These adaptations help integrate FirstHand's approach into the city's educational ecosystem.

FirstHand contributes to Philadelphia's STEM ecosystem by collaborating with organizations to expand high-quality STEM opportunities.

Beyond Philadelphia, FirstHand's program model informs national practice. The program advised on a replication at Charlotte, North Carolina's innovation district, The Pearl. FirstHand's industry-informed, lab-based STEM programming was adapted to Charlotte's ecosystem using lessons from the Experiencing STEM FirstHand project, drawing particularly on the logic model. Because The Pearl and uCity Square share a common developer, Wexford Science & Technology, this collaboration also offers potential pathways for scaling the model to other innovation districts.

Dimension 5: Fostering Ongoing Learning and Developing Infrastructure for Partnership

Our aim is for this RPP to continue beyond the EIR grant; we have dedicated time and resources to fostering learning among all team members and building a strong partnership infrastructure. A strong RPP focuses not only on generating research findings but also on empowering practitioners to sustain improvement and evidence-informed practice after the research team steps back.

At the *individual level*, we build program staff capacity and support ongoing professional development with a focus on sustainability. Facilitators are trained on the STEM PQA to set facilitation goals and strengthen practice, equipping them to continue cycles of inquiry independently and fostering a culture of reflective decision making beyond the funded project. At the *program level*, collaboration and formative data have driven tangible improvements. The student lab notebook was redesigned with embedded reflection prompts to better capture student thinking and feedback, and facilitators now more explicitly connect lab activities to real-world STEM applications and career pathways.

At the *partnership level*, we have established routines and systems to support collaboration and continuous improvement, including structured weekly meetings, standard data processes, clearly defined roles, and expectations for communication and participation. Annual check-ins focus on partnership health, revisiting the shared vision, reflecting on successes, and identifying growth opportunities. These discussions emphasize sustainability and scaling, exploring how program practices, research findings, and partnership models can be maintained and adapted to other contexts. The team also regularly evaluates future opportunities for alignment with program and research priorities.

From Partnership to Practice: Lessons Learned from Our RPP Journey

Our collaboration has generated a wealth of practical insights about what it takes to build, sustain, and learn from an RPP. These lessons extend beyond our specific context and speak to the everyday realities of conducting inclusive, improvement-oriented research that meets local needs and advances broader educational goals. Several key lessons have shaped how we think about partnership, evaluation, and the role of research in driving meaningful change.

Lesson 1: Strong partnerships start before research begins. A successful RPP is built on trust, shared purpose, and collaboration, which must be cultivated before data collection begins. Our partnership's strength came from relationships and mutual understanding established early. Investing time to learn program priorities, understand contextual pressures, and collaboratively shape research questions ensured that the research was done *with* schools, not *to* them. Clarifying roles, aligning goals, and establishing communication routines—such as regular check-ins, shared planning documents, and transparent decision making—created processes that respected practitioners' time and expertise.

Investing time to learn program priorities, understand contextual pressures, and collaboratively shape research questions ensured that the research was done with schools, not to them.

Lesson 2: Formative evaluation fuels continuous improvement. Formative evaluation is valuable at every stage of a program's life cycle. Studying implementation through an improvement-oriented lens (Marwell & Mosley, 2025) provides essential insights for growth, whether a program is new or established. In the Experiencing STEM FirstHand project, formative evaluation fostered ongoing relationships between researchers and practitioners, creating immediate value even before summative results were available. It also strengthened the program's capacity for continuous quality improvement, cultivating a culture of reflective practice and data-informed decision making. Over time, these habits become self sustaining, enabling staff to conduct their own inquiry cycles, test small changes, and refine practice based on emerging evidence.

Lesson 3: Enter RCTs with a learning orientation and clear expectations. RCTs can provide powerful evidence, but they require significant capacity, stability, and clarity of purpose. Programs should assess whether this approach fits their development stage and organizational readiness by asking the following:

1. *Does your organization have the capacity to conduct an RCT and to learn from it?* RCTs are expensive, complex, and time intensive (Marwell and Mosley, 2025); therefore, only organizations with strong infrastructure, staffing, and support systems can typically manage the demands of a full trial. Programs must be prepared to implement the model consistently across participants and sites, with core components clearly defined and operationalized in the protocol. Reliable delivery requires systems to monitor fidelity, track deviations, and provide feedback, supported by staff training, rehearsal of procedures, and iterative testing to ensure consistent implementation under real-world conditions. RCT readiness also requires a learning orientation. Entering an RCT means assuming that positive effects may not emerge and focusing instead on whether the study is measuring the right outcomes at the right time and whether participants interpret items as intended. Embedding early cognitive interviews and pilot testing helps ensure that instruments are valid, meaningful, and aligned with the program's logic model. A flexible, learning-focused stance, rather than a rigid "RCT or bust" approach, ensures that, regardless of effect size, the RCT generates actionable insights that strengthen the program.
2. *How will the program adapt and maintain fidelity during the RCT and manage service capacity*? Programs evolve naturally, but in an RCT, mid-course changes can undermine the ability to attribute outcomes to the intervention. In this project, we reduced burden through clear, proactive communication, including concise infographics outlining data collection timelines, instruments, and time commitments. We also remained responsive to partner feedback. When pre-testing proved too disruptive, we shifted to a post-test-only design, preserving research goals while reducing administrative strain.

Lesson 4: Rethink what it means to scale. Scaling is often framed as replicating a model in new contexts, but it can also involve extending a program's influence through practices, tools, or insights that inform other initiatives. Programs benefit from planning for scale early by documenting core components, identifying adaptable elements, and considering how the work might extend beyond its original setting. Framing scaling as the diffusion of learning rather than duplication of form allows programs to pursue broader impact while remaining responsive to local needs.

Drawing on Coburn (2003), scale can be understood as multidimensional: spread into new contexts without requiring exact replication; depth of change in practice among adopters; a shift in ownership from external to local actors; and sustainability beyond initial funding. For example, the resulting program of the Charlotte replication efforts (described in Dimension 4) focuses on healthcare, whereas FirstHand in Philadelphia emphasizes biotechnology. In both cases, core components remain visible but are tailored to local priorities and opportunities. Together, these dimensions emphasize that scaling is not just about growth, but also about how programs are taken up, adapted, and sustained across contexts.

Framing scaling as the diffusion of learning rather than duplication of form allows programs to pursue broader impact while remaining responsive to local needs.

Lesson 5: Immersed local researchers are foundational for trust and actionable data. Recruiting, developing, and retaining qualified staff for long-term, embedded research roles can be difficult, particularly in out-of-school time contexts where staffing instability is well documented (Brooks, 2025). These challenges were present in our work as well, requiring ongoing attention to hiring, support, and continuity.

Our experience suggests that one critical strategy for navigating these challenges is to invest deeply in partnerships and relationship building. The decision to hire local research team members who were physically present and immersed in the program's daily ecosystem was foundational. By showing up consistently at every program session, researchers moved beyond the role of external observers to become known and trusted members of the community. This relational presence not only strengthened data collection but also contributed to staff retention. This aligns with research showing that afterschool educators are motivated to stay when they experience strong relationships and a sense of impact (Moroney et al., 2025). In our context, investing in relationships with students, staff, and partners helped foster that sense of purpose and belonging, supporting stability within the research team and data quality.

Conclusion

The Experiencing STEM FirstHand project illustrates the potential of an RPP designed to advance equitable learning in informal STEM education while also extending what RPPs can contribute methodologically and structurally. Unlike many prior RPP case studies in informal STEM, this partnership integrates a multi-cohort RCT in a responsive, practice-embedded design, pairing causal inference with continuous formative feedback. By combining these elements, the project expands students' access to meaningful STEM experiences while demonstrating how RPPs can simultaneously support rigorous impact evaluation, embedded practice-based learning, and adaptive implementation at scale.

As the project expands within Philadelphia and adapts to new contexts, the principles guiding this RPP continue to shape its evolution. We hope these lessons offer a model for how partnerships can integrate embedded research roles, causal evaluation designs, and context-responsive scaling to both improve practice and contribute new methodological approaches for the field of informal STEM education.

Acknowledgments

This work was funded by a grant from the U.S. Department of Education, S411C210055. The contents were developed under a grant from the U.S. Department of Education, Education Innovation and Research (EIR) Program. However, those contents do not necessarily represent the policy of the U.S. Department of Education; the reader should not assume endorsement by the federal government.

This work would not have been possible without the extraordinary dedication of countless individuals across Philadelphia. Our deepest gratitude goes first and foremost to the students and families. Your participation is at the heart of this effort, and none of this would have been possible without you. We extend sincere thanks to the School District of Philadelphia for its commitment to providing high-quality, hands-on educational experiences and to our partner schools for their thoughtful collaboration. We are especially grateful to the FirstHand facilitators and leadership, as well as to the many mentors and industry professionals who generously volunteered their time and expertise. The collective dedication of each of these partners to supporting Philadelphia's youth has been essential to this work.

References

Aish, N., Asare, P., & Miskioğlu, E. E. (2018). People like me: Providing relatable and realistic role models for underrepresented minorities in STEM to increase their motivation and likelihood of success. In *2018 IEEE Integrated STEM Education Conference (ISEC)*, pp. 83–89. IEEE. https://digitalcommons.bucknell.edu/cgi/viewcontent.cgi?article=1057&context=fac_conf.

Bell, J., Falk, J., Hughes, R., Hunt, G., Parrish, J., Ruffin, M., Sacco, K., & Troxel, G. (2016). Informal STEM education: Resources for outreach, engagement and broader impacts. Center for Advancement of Informal Science Education (CAISE). https://informalscience.org/wp-content/uploads/2019/02/CAISE_Broader_Impacts_Report_2016_0.pdf.

Bloomquist, S., & Rosenfeld, M. (2021, September 7). Philadelphia area schools see racial disparities in AP enrollment. 6 ABC (Action News). https://6abc.com/post/data-investigation-schools-educational-inequity-advanced-coursework/10990411.

Brooks, L. (2025, September 29). New report explores how we can support and grow the afterschool STEM workforce. Afterschool Alliance. https://www.afterschoolalliance.org/afterschoolSnack/New-report-explores-how-we-can-support-and-grow-the-afterschool_09-29-2025.cfm.

Burden, D., Cole, C., Fetterman, J., Hatfield, C., Holland, S., & Myers, J. (2024). Life sciences report 2024. Colliers. https://www.colliers.com/en/research/nrep-life-sciences-report-2024.

The Century Foundation. (2019). The benefits of socioeconomically and racially integrated schools and classrooms. https://production-tcf.imgix.net/app/uploads/2016/02/26171529/Factsheet_Benefits_FinalPDF.pdf.

Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher, 32*(6), 3–12. https://doi.org/10.3102/0013189X032006003.

Coburn, C. E., Penuel, W. R., & Geil, K. E. (2013). Research–practice partnerships: A strategy for leveraging research for educational improvement in school districts. William T. Grant Foundation. https://wtgrantfoundation.org/wp-content/uploads/2014/02/Research-Practice-Partnerships-at-the-District-Level.pdf.

Dabney, K. P., Tai, R. H., Almarode, J. T., Miller-Friedmann, J. L., Sonnert, G., Sadler, P. M., & Hazari, Z. (2012). Out-of-school time science activities and their association with career interest in STEM. *International Journal of Science Education, Part B, 2*(1), 63–79. https://doi.org/10.1080/21548455.2011.629455.

DeWitt, J., Godec, S., Watson, E., Archer, L., & Moote, J. (2025). Who participates in informal science learning experiences? Analysis of data from six surveys of a cohort in England from age 10–22 years. *International Journal of Science Education, Part B, 15*(4), 612–629. https://doi.org/10.1080/21548455.2025.2462798.

Dierking, L. D., Falk, J. H., & Storksdieck, M. (2013). Learning from neighboring fields: Conceptualizing outcomes of environmental education within the framework of free-choice learning experiences. In *International handbook of research on environmental education* (R.B. Stevenson, M. Brody, J. Dillon, A. E. J. Wals, eds.). pp. 359–366. Routledge.

Edelen, D., Cook, K., Tripp, L. O., Jackson, C., Bush, S. B., Mohr-Schroeder, M. J., Schroeder, D. C., Roberts, T., Maiorca, C., Ivy, J., Burton, M., & Perrin, A. (2024). "No, this is not my boyfriend's computer": Elevating the voices of youth in STEM education research leveraging photo-elicitation. *Journal for STEM Education Research, 7*(3), 444–462. https://doi.org/10.1007/s41979-024-00118-z.

Falk, J., & Dierking, L. (2010). The 95 percent solution. *American Scientist, 98*(6), 486. https://doi.org/10.1511/2010.87.486.

Fetterman, D. M. (2001). The transformation of evaluation into a collaboration: A vision of evaluation in the 21st century. *American Journal* *of Evaluation, 22*(3), 381–385. https://doi.org/10.1177/109821400102200315.

Fry, R., Kennedy, B., & Funk, C. (2021). STEM jobs see uneven progress in increasing gender, racial and ethnic diversity. Pew Research Center. https://www.pewresearch.org/social-trends/2021/04/01/stem-jobs-see-uneven-progress-in-increasing-gender-racial-and-ethnic-diversity/?gad_source=1&gad_campaignid=22378837192&gbraid=0AAAAA-ddO9HF_99Doi_2FxKDinyhoef3R&gclid=CjwKCAjwwpDQBhAuEiwAa-4WowDLnmKxxng1fZJn_3l1basQQdZhzxX4MvV3s0uch8Q_rz9uwMpLwhoC6P4QAvD_BwE.

Hatfield, N., Brown, N., & Topaz, C. M. (2022). Do introductory courses disproportionately drive minoritized students out of STEM pathways?. *PNAS Nexus, 1*(4), pgac167. https://doi.org/10.1093/pnasnexus/pgac167.

Henrick, E. C., Cobb, P., Penuel, W. R., Jackson, K., & Clark, T. (2017). Assessing research-practice partnerships: Five dimensions of effectiveness. William T. Grant Foundation. https://rpp.wtgrantfoundation.org/wp-content/uploads/2019/09/Assessing-Research-Practice-Partnerships.pdf.

Henrick, E., Farrell, C. C., Singleton, C., Resnick, A. F., Penuel, W. R., Arce-Trigatti, P., Schmidt, D., Sexton, S., Stamatis, K., & Wellberg, S. (2023). Indicators of research-practice partnership health and effectiveness: Updating the five dimensions framework. National Center for Research in Policy and Practice and National Network of Education Research-Practice Partnerships. https://doi.org/10.17605/OSF.IO/YVGCN.

Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences, 101*(46), 16385–16389. https://doi.org/10.1073/pnas.0403723101.

Joyce, K. E., & Cartwright, N. (2020). Bridging the gap between research and practice: Predicting what will work locally. *American Educational Research Journal, 57*(3), 1045–1082. https://doi.org/10.3102/0002831219866687.

Marwell, N. P., & Mosley, J. E. (2025). The nonprofit sector has an RCT problem. *Stanford Social Innovation Review, 23*, 5867. https://doi.org/10.48558/Y8Z0-8264.

Mezzacappa, D. (2023, November 13). Philly's premier high school requires 8th grade algebra for admission. Many kids can't take it. Chalkbeat (Philadelphia). https://www.chalkbeat.org/philadelphia/2023/11/13/eighth-grade-algebraaccess-equity-masterman.

Moroney, D., Brown, M. E. M., Diffenderffer, A. E., Vasudevan, D. S., Richter, J. Y., Carroll, R., Syvertsen, A. K., & Tasayco Prado, G. N. (2025). Power of Us: The youth fields workforce findings. American Institutes for Research; Wallace Foundation. https://doi.org/10.59656/YD-G3634.001.

National Academies of Sciences, Engineering, and Medicine. (2023). Advancing anti-racism, diversity, equity, and inclusion in STEMM organizations: Beyond broadening participation. Washington, DC: The National Academies Press. https://doi.org/10.17226/26803.

National Academies of Sciences, Engineering, and Medicine. (2018). Learning through citizen science: Enhancing opportunities by design, p. 25183. Washington, DC: The National Academies Press. https://doi.org/10.17226/25183.

National Academies of Sciences, Engineering, and Medicine. (2026). The status of informal science and engineering education: Proceedings of a convocation. Washington, DC: The National Academies Press. https://www.nationalacademies.org/read/29291.

National Center for Education Statistics. (2024, March 19). *A majority of grade 9-12 public schools rate themselves favorably on preparing students for college*. Institute of Education Sciences, U.S. Department of Education. https://ies.ed.gov/learn/press-release/majority-grade-9-12-public-schools-rate-themselves-favorably-preparing-students-college.

National Research Council. (2009). Learning science in informal environments: People, places, and pursuits. Washington, DC: National Academies Press. https://doi.org/10.17226/12190.

National Research Council. (2012). A framework for K-12 science education: Practices, crosscutting concepts, and core ideas. Washington, DC: National Academies Press. https://doi.org/10.17226/13165.

Page, S. (2025). *The difference: How the power of diversity creates better groups, firms, schools, and societies*. Princeton, NJ: Princeton University Press. https://press.princeton.edu/books/paperback/9780691271743/the-difference?srsltid=AfmBOorC8hLs6qf9hHUseHfjwjTfTNa42nS76h-lpIAMAv0vMns6ggTm.

Panorama Education. (2024). Reliability and validity of Panorama's survey topics for students: 2024 update with addendum for versioned content. https://go.panoramaed.com/hubfs/Validity-Report-Student-Topics-2020.pdf.

Riegle-Crumb, C., King, B., & Irizarry, Y. (2019). Does STEM stand out? Examining racial/ethnic gaps in persistence across postsecondary fields. *Educational Researcher, 48*(3), 133–144. https://doi.org/10.3102/0013189X19831006.

Saw, G. K. (2020). Leveraging social capital to broaden participation in STEM. *Policy Insights from the Behavioral and Brain Sciences, 7*(1), 35–43. https://doi.org/10.1177/2372732219895997.

The School District of Philadelphia. (2024). Pathways to community, joy, and fulfillment for School District of Philadelphia students. https://www.philasd.org/strategicpartnerships/#partnerguide.

Slaughter, A., Moran, D., Lapp, D., & Lin, J. (2016). Racial disparities in educational opportunities in Pennsylvania: A first look at new civil rights data. Research for Action (RFA). https://www.researchforaction.org/wp-content/uploads/2021/07/RFA-Civil-Rights-Data-PACER-Brief-Nov-1-2016.pdf.

Smith, C. (2012). Development and early validation evidence for an observational measure of high quality instructional practice for science, technology, engineering and mathematics in out-of-school time settings: The STEM supplement to the Youth Program Quality Assessment. The Forum for Youth Investment and Providence Afterschool Alliance. https://www.researchgate.net/publication/324089942_Development_and_early_validation_evidence_for_an_observational_measure_of_high_quality_instructional_practice_for_science_technology_engineering_and_mathematics_in_out-of-school_time_settings_The_STEM_

Smith, C., Akiva, T., Sugar, S., Devaney, T., Lo, Y.-J., Frank, K., Peck, S. C., & Cortina, K. S. (2012). Continuous quality improvement in afterschool settings: Impact findings from the Youth Program Quality Intervention study. Forum for Youth Investment. https://forumfyi.org/knowledge-center/continuous-quality-improvement-in-afterschool-settings-impact-findings-from-the-youth-program-quality-intervention-study/.

Stocklmayer, S. M., Rennie, L. J., & Gilbert, J. K. (2010). The roles of the formal and informal sectors in the provision of effective science education. *Studies in Science Education, 46*(1), 1–44. https://doi.org/10.1080/03057260903562284.

Sun, L. (2024). Embracing informal learning experiences to support science and mathematics education. *School Science & Mathematics, 124*(5), 289–292. https://doi.org/10.1111/ssm.18307.

Taylor, D. (2024). The STEM labor force: Scientists, engineers, and technical workers (No. NSB-2024-5). National Science Board, National Science Foundation. https://ncses.nsf.gov/pubs/nsb20245/.

U.S. Department of Education, Office for Civil Rights. (2025). 2021–22 civil rights data collection, a first look: Students' access to educational opportunities in U.S. public schools. https://www.ed.gov/media/document/2021-22-crdc-first-look-report-109194.pdf.

Vey, J. S., Andes, S., Hachadorian, J., & Katz, B. (2017). Connect to compete: How the University City-Center City innovation district can help Philadelphia excel globally and serve locally. The Brookings Institution. https://www.brookings.edu/wp-content/uploads/2017/05/csi_20170511_philadelphia_innovationdistrict_report1.pdf.

YESTEM Project Team. (2021). YESTEM Insight 3.1: Equitable Youth Outcomes Model for informal STEM learning. https://discovery.ucl.ac.uk/id/eprint/10138557/1/2021%20YESTEM%20Insight%203.1%20Equitable%20Youth%20Outcomes%20Model.pdf.

Zhao, M., Mathews, C. J., Mulvey, K. L., & Hartstone-Rose, A. (2022). Promoting diverse youth's career development through informal science learning: The role of inclusivity and belonging. *Journal of Youth and Adolescence, 52*(2), 331–343. https://doi.org/10.1007/s10964-022-01694-2.