

Data and Partnerships Drive Continuous Improvement in EIR Projects

January 2023



Acknowledgments

The contents of this document were developed under a contract from the U.S. Department of Education through the Office of Innovation and Early Learning Programs (IELP) within the Office of Elementary and Secondary Education (OESE), by AnLar under Award # GS00F229DA. This contains resources that are provided for the reader's convenience. These materials may contain the views and recommendations of various subject matter experts as well as hypertext links, contact addresses, and websites to information created and maintained by other public and private organizations. The U.S. Department of Education does not control or guarantee the accuracy, relevance, timeliness, or completeness of any outside information included in these materials. The views expressed herein do not necessarily represent the positions or policies of the U.S. Department of Education. No official endorsement by the U.S. Department of Education of any product, commodity, service, enterprise, curriculum, or program of instruction mentioned in this document is intended or should be inferred.

Suggested Citation

Holian, L. (2023). Data and Partnerships Drive Continuous Improvement in EIR Projects. Cross-project summary produced by AnLar and the Office of Elementary and Secondary Education; Education Innovation and Research Program (EIR).

Introduction

Continuous improvement is an ongoing process that is systematic and cultural. It includes performance assessment, data-driven planning, support for implementation, and reassessment and tracking. While educational programs and educators can and often do engage in continuous improvement cycles on their own, partnership with researchers or external evaluators can enhance the continuous improvement process. The recent focus on researcher-practitioner partnerships in education has pushed informal relationships to formal or required involvement of researchers with program developers.

Many grant-making agencies and offices have focused on use of research evidence. As such, they often require applicants to consult with an external evaluator. Some require an outcomes focused study using methods such as quasi-experimental designs or randomized controlled trials. While rigorous impact studies are conducted on a program as a whole, and provide evidence about program effectiveness, it occurs after a program has been implemented. The information is important for future actions and programs, but program staff and teachers need more timely feedback. Program staff and teachers rely on continuous formative feedback while implementing a program. Therefore, collaborating with an external evaluator should also involve the analyzing of early or preliminary results; supporting the applicant's understanding of what works, what does not work, and for whom; and contributing expertise and support for program planning and improvement. The Education Innovation and Research (EIR) program requires grantees to work with an external evaluator and encourages use of external evaluators to support each of those purposes.

Education Innovation and Research Notice Inviting Applications for FY 2022

“Early-phase grants provide funding for the development, implementation, and feasibility testing of a program, which prior research suggests has promise, for the purpose of determining whether the program can successfully improve student achievement and attainment for high-need students. Early-phase grants are not intended simply to implement established practices in additional locations or address needs that are unique to one particular context. The goal is to determine whether and in what ways relatively new practices can improve student achievement and attainment for high-need students.”

While all tiers of EIR grantees can benefit from ongoing formative feedback from evaluators, Early-phase grantees are in a particularly strong position to make changes or improvements as they develop their programs based on data and evaluator feedback. Early-phase grant recipients are encouraged to work with their evaluators and to use feedback to iteratively improve their programs. The relationship between a grantee and their evaluator can be conceptualized as a research-practitioner partnership (see [Adam, Gangier, and Jones-Manson 2022](#)). The grantee is responsible for the design and implementation of a program or intervention. Researchers are responsible for

conducting evaluation and assessment activities, and highlighting trends and evidence from research that is relevant to the work of the program developer, whether it is a state, district, university, or non-profit organization.

In addition to the relationship between the grantee and the evaluator, advisory boards can also promote feedback loops on project implementation. Advisory board members are often practitioners who provide insight into the context, or they may have content area expertise in the program or intervention components. This form of partnership can be as important as the grantee-evaluator partnership. When evaluators and grantees meet jointly with an advisory board, the evaluators and grantees can share data and have conversations about progress towards goals. The advisors can weigh in and offer alternative perspectives or help make connections to others. If advisory boards include school staff where a program is being implemented, those staff members can act as champions for the program in the school and help to improve program implementation. Having family members and students on advisory boards is also a useful practice to provide ongoing insights about project implementation.

Researchers can provide ongoing formative feedback about what is working and how teachers, students, and families are experiencing the program. Together, the researchers, program staff, and advisors will be able to create a more comprehensive, better informed program with a better chance of success, and generate more useful evidence than if operating alone.

This cross-project summary focuses on how four EIR funded project teams use interim evaluation data to adjust their program implementation in the following four ways:

- Conducting a pilot study;
- Developing and reviewing meaningful performance metrics;
- Collaborating on measuring fidelity of implementation; and
- Engaging in continuous data collection, analysis, and review.

EIR Project Examples

Summer Academy for Integrated Language Learning (SAILL) Program



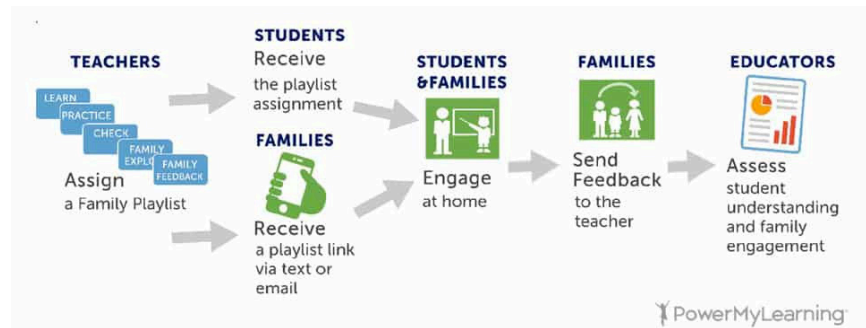
Funded in 2019, the Summer Academy for Integrated Language Learning (SAILL) is a 5-week summer program for newcomer English language learner (EL) students. SAILL offers contextualized English language instruction through computer science coursework. Teachers who offer the instruction receive professional development and both students and teachers have opportunities for peer-assisted learning. Students are able to earn credits toward high school graduation by participating. The SAILL project is a partnership between San Francisco Unified School District (SFUSD) and Rockman.

EIR Project Examples

Family Playlists: Improving High-Need Students' Academic Performance and SEL in Response to COVID-19



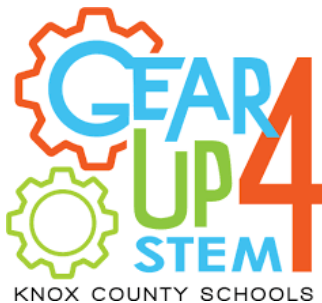
Funded in 2021, the Family Playlists program includes teacher professional development (PD) and a mobile technology tool called Family Playlists to improve students' math performance and social emotional learning. Teachers are trained to use Family Playlists to have students teach academic concepts to their families and have families then provide feedback. Teachers receive a range of family feedback including video recordings of the students' presentations to their families. The Family Playlists project is a partnership between PowerMyLearning (PML), Miami-Dade County Public Schools, and Abt Associates.



Talk to Read: A Culturally Responsive Approach to Literacy Recovery

Funded in 2021, the Talk to Read program provides teacher professional development with the goal of supporting second-grade students' literacy learning and recovery. Specifically, this program seeks to improve the following student outcomes: sight word fluency, sight vocabulary, and reading comprehension. Continuous improvement is built into the Talk to Read program through formative data collection, teacher feedback interviews, and monthly project team meetings to review feedback and revise the design. The Talk to Read project is a partnership between the Curators of the University of Missouri Special Trust and the American Institutes of Research.

Gear Up 4 STEM



In 2018 Knox County Schools received funding to develop a four week STEM camp for elementary school students. The camp takes place in the late summer, during fall and spring breaks, and after the school year ends. It introduces students to the Engineering Design Process, exposes students to a variety of STEM careers, provides hands-on learning experiences, and uses local partners to show students how STEM is relevant in the local community. The partnership is between Knox County Schools and AnLar.

Conducting a pilot study

Pilot studies present opportunities for close study of the components of programs and interventions that are hypothesized to be critical for driving positive change. During a pilot study, researchers and grantees limit the scale of investigations in important ways. The scale may be limited in terms of the overall number of students, teachers, schools, districts, or communities involved. It can also be limited in the kind or quantity of measurement instruments used in the study. Researchers and grantees make decisions about how the scale of early stage studies will be limited based on what they hope to learn from and improve as a result of the pilot.

Many EIR projects incorporate a pilot year. Pilot studies can serve several purposes:

- Testing the feasibility of assumptions underlying a theory of action
- Understanding what influences fidelity of implementation
- Testing the validity and/or feasibility of measurement instruments
- Allowing for adjustments to programming that can lead to sustained and systemic change

The Talk to Read, SAILL, and Gear Up 4 STEM projects each partnered with their external evaluators to conduct pilot studies prior to full implementation of their impact studies. These projects used the pilot phase to gather feedback about various aspects of the projects and test data collection instruments.

Guides, such as the [Institute of Science resource titled Learning Before Going to Scale: An Introduction to Conducting Pilot Studies](#) (2021), offer templates and specific advice for all stages of the design and execution of pilot studies.

Part of the **Talk to Read** project involved asking teachers to complete a reflective journal entry at the end of each professional development session. Early feedback during the pilot phase revealed that teachers did not like the activity. However, by the end of the year, the initial negative feedback about the reflective journal entries had turned positive. Teachers valued having written entries that they could look back on to see their growth through the process. Allowing the time to fully pilot this part of the program showed that it might not be well received at first, but that over time teachers see the value in it.

For the **SAILL** project, the pilot study period was used to determine how to best assess students. Because this project works with newcomer English Learners (EL), the grantee and evaluators wanted an assessment that would capture the growth of English proficiency among students. Through the pilot phase, evaluators were able to examine where available or existing English Language Proficiency (ELP) assessments would place students before the program and if the assessments were sensitive enough to detect changes. They wanted to make sure whichever assessment they used would measure what SAILL was trying to improve. The SAILL program was focused more on improving speaking English and less on writing, so assessments that were heavily focused on writing were not appropriate. By presenting decisions about assessments to the SAILL staff in terms of the findings from the feasibility study, the evaluator and SAILL team were able to jointly make informed decisions. Another lesson learned in the pilot period was how to use electronic signatures for consent. The use of electronic signatures from parents and students allowed more flexibility and ease in tracking. This practice ensured consent was obtained in a timely manner so that the program could begin. One

additional lesson learned through the pilot phase was a need to adjust record keeping strategies to systematically track teacher collaboration. Teacher collaboration is an important component of the program and integral to understanding fidelity of implementation.

The first year and a half of the **Gear Up 4 STEM** project was spent developing and then piloting the program. The program was designed as a 4-week camp that takes place during July (before the school year begins), October (during fall break), March (during spring break), and June (just after the school year ends). Each week was five days from 8:30 a.m. until 3:30 p.m. Throughout the week, camp facilitators lead students in engaging in project-based learning modules, completing fun STEM tasks, and using a variety of technology tools. During the pilot year, feedback was collected from camp facilitators, families, and students. The pilot year allowed the camp facilitators to test out and refine the newly developed STEM camp modules. The Gear Up 4 STEM team learned several lessons from the pilot:

- It is important for the camp facilitators to meet and have in-depth planning with each other.
- Getting through 5 modules per week is a challenge. Often, the modules take longer than anticipated. By assigning only 4 modules per week, the facilitators are better able to make adjustments to the schedule to meet the needs of students without feeling rushed.
- Clear communication with the camp facilitators is essential to ensure they feel empowered to edit the camp's daily schedule to meet the needs of the students.
- Camp facilitators need time to test out the technology before using it with students.

The pilot year also allowed time to test and refine data collection instruments. The student survey instruments were tested for both student comprehension and to see if there was variation in responses.

As a result of the formative feedback and discussions with their evaluator during the pilot year, the Gear Up 4 STEM team updated program implementation in several ways. First, they allowed teachers to apply as either part-time or full-time facilitators to proactively address feelings of burnout or fatigue. Second, based on feedback from parents, the program shifted to make sure students produced something to take home each day to increase parent knowledge of what was happening in the camp. Some sites also created daily flyers to share student accomplishments with families. Third, the pilot year showed that recruitment and retention of students would be a challenge. To increase the number of students who registered, the Gear Up 4 STEM team performed more outreach to inform schools about the program. And to increase participation of registered students, facilitators called each family prior to the first week. Lastly, due to the drop in student attendance on Fridays and feelings of burnout from teachers, the program shifted to have a virtual day on Fridays during Fall and Spring Breaks. By making Fridays virtual, it allowed students and facilitators to come together and collaborate, but also have a break before school was back in session. These changes were made and reflected in the implementation.

Additional Example of Using a Pilot Year

The 2020 Early-phase project, “Soft Skills in Computer Science Pathways,” used a pilot year to train teachers in soft skills integration, build infrastructure of an online hub, and to select measurement instruments. Through the pilot they discovered a need to create a collaborative workspace and to have each school track their activities using a consistent tool.

Pilot studies allow Early-phase grantees to be flexible in their program designs and evaluation design. They use information gathered during the pilot to make implementation changes, or evaluation design changes.

Developing and Reviewing Performance Metrics

While impact evaluations can answer questions about the effect of a program on particular student, teacher, or school outcomes after a program has been implemented, practitioners have other questions about project effectiveness throughout the implementation period. Understanding effectiveness starts with the clear articulation of project goals and objectives. As part of the EIR grant funding, project teams set their goals and identify annual performance measures based on what the project is trying to achieve and what success looks like.

Because of the connection between project goals, performance measures, project implementation, and intended outcomes, project teams need to closely communicate and collaborate with external evaluators. Project teams can work with their evaluators to help identify ways to measure progress towards goals. Evaluators may also collaborate with grantee project teams on instrument selection, identifying ways of measuring goals, or finding existing studies with similar approaches that can help inform annual targets.

An example of setting performance metrics comes from the **Talk to Read** program. One of their goals is to increase teacher efficacy and effectiveness in implementing Talk to Read strategies in their core curriculum. The performance measure they identified was the percentage of teachers reporting an increase in the use of strategies learned in professional development. Their target is to grow from 60 percent in year 2 to 65 percent in year 3. Another goal is increasing student sight word vocabulary, sight word fluency, and reading comprehension. Their performance measure is the percentage of students increasing sight word fluency by 50 percent or greater. They ideally want to see all students increase by 50 percent or better, but based on conversations with the project team and evaluator, felt that having 50 percent of students increase their sight word fluency by 50 percent or greater would be reasonable. After their second grant year, they will revisit their targets and adjust if needed.

For the **Family Playlists** project, the project team collaborated with Abt Associates, their evaluator, on measurement selection and data collection. The evaluation team helps the program collect data through teacher, family, and principal surveys and interviews to inform changes.

The targets for performance measures should be reviewed annually. They can be adjusted based on changes in context, review of data, or external forces such as the COVID-19 pandemic. Working with evaluators to define and capture performance metrics can help project teams identify areas of

success, and areas for improvement. Evaluators can help project teams identify data sources, or can share data they are already collecting. Conversations between project teams and evaluators about the meaning of the data, and whether or not targets are achieved, can help identify ways to improve or change the program. Setting meaningless performance metrics based only on what data is available or what is easily achieved, without linking it to a theory of change, is not aligned with a culture of continuous improvement.

Collaborating on Measuring Fidelity of Implementation

A corollary to developing and reviewing performance metrics is developing and reviewing fidelity of implementation indicators. As part of the evaluation of EIR funded projects, external evaluators develop and conduct fidelity of implementation studies. Evaluators collaborate with grantees to define the measures and understand what is needed to make the program work. Developing the indicators of fidelity is both an art and a science. There can be space for local innovation and ideation, but the actual practices must also be captured and described. Collaboration between the grantee and evaluator is essential. After jointly developing the definition of fidelity, evaluators collect data and report on whether each element meets fidelity. This descriptive information is important, but the thresholds for what constitutes fidelity may need to change. It is possible that the originally developed thresholds are not correct. For example, it might be thought that students need 60 minutes of tutoring per week to improve scores, so teams set 60 minutes as the threshold. But, in the course of the study the evaluator might find that students who received 30 minutes had the same gains in scores as students who received 60 minutes. In that case, the team might revisit the threshold and set it at 30 minutes instead. Alternatively, everything in a program might be implemented with high fidelity, but the expected outcomes are not achieved. In that case, the causal mechanism might need to be revisited.

In the **SAILL** project the evaluator reported on fidelity of implementation at the end of the year. Then the program team reflected on the results and made changes if they felt they were warranted. Understanding the real-time dosage information helps the program team make decisions about where more attention should be focused (e.g., computer science, embedded English instruction).

For **Family Playlists**, the evaluator spent a lot of time understanding the context and outcomes. The Family Playlists team discussed the spirit of the program with the evaluators as they designed measures of fidelity of implementation. Both the project team and the evaluators did their own research on the theory of action and the amount of intervention likely to be needed to have an impact.

The **Gear Up 4 STEM** project looked at fidelity of implementation during the pilot year and determined that changes needed to be made on several indicators. Some of these changes were due to changes in how the program was actually implemented, including the number of modules to be covered during the 4 weeks of camp.

A tension can arise between evaluators and program staff when evaluators design an evaluation of a program as is, but the program ends up changing the way it is implemented based on feedback about what is working. One way to alleviate this tension is through transparent conversations and clear documentation of how the program is being implemented and what is being evaluated.

Engaging in Continuous Data Collection, Analysis, and Review

A common model for continuous improvement involves selecting a practice to test, identifying measures and developing a data collection plan, implementing the practice, collecting data, collectively examining the data to inform improvement, and make changes based on the data. This all involves extensive data collection. External evaluators can help program staff review data they are already collecting, identify other data that the program staff think they should be collecting, and help program staff analyze or interpret data.

Additional Examples of Continuous Data Collection, Analysis, and Review in Practice

The 2018 Early-phase project, “Developing a Student-Driven STEM and Computer Science Curriculum for Rural Students” used a Plan-Do-Study-Act framework for their professional learning sessions. The grantee, Sonoma State University (SSU), planned the agenda with their evaluators, then executed the professional learning sessions. The evaluator observed the sessions. Following the session, the SSU team met with the evaluator to debrief and review end of session assessments. This led to developing an event guide. The cycle of planning the training, implementing training, studying how the training went, and making changes is repeated. The training can be customized based on the needs of the teachers attending.

The 2020 Early-phase project, “Soft Skills in Computer Science Pathways,” used their evaluator to observe the program and provide informal feedback to the Learning Design Team on how training sessions were implemented and how they align with the goals of the program and evaluation.

The types of data that can be used for continuous improvement include data on program quality, attendance or participation, student outcomes, student experiences, teacher outcomes, and teacher experiences. One way to help minimize the burden of data collection is by using technology to support data collection which allows for quicker turnaround and analysis of data and provides swift feedback.

One example of this comes from the **Gear Up 4 STEM** project. They used Google Forms and Google Sheets to collect data on student attendance; teacher ratings of students; and feedback from parents, teachers, and mentors. The evaluators were then able to analyze the data and provide the project team with a summary memo a week after each camp session. Getting weekly feedback allowed the project team to make adjustments or follow up on items before the next camp session.

Being judicious about how quickly to respond to program data can be easier when the researcher and program teams work together to make sense of emerging results. Leaders from the projects featured in this cross-project summary all referenced the two-way nature of the learning that happens. In the **SAILL** project, the evaluators present findings at quarterly meetings so that the program team can remain up to date. In these presentations, the researchers offer evidence-supported suggestions and guidance for the program team to consider. The evaluators remain in regular contact with the

SAILL team through calls and one-on-one conversations. The evaluators analyze data and offer recommendations, but they also listen to the expertise of the program staff and allow them to make the changes and put a plan into action.

Looking Forward

Research-practitioner partnerships bring together two teams of intelligent, educated individuals who have experience in different aspects of improving outcomes for students and teachers. Research-practitioner partnerships can be a natural fit to support continuous school improvement. Researchers work with program developers or school personnel to set goals, identify strategies for improvement, and evaluate change. Communication and trust are crucial elements in a partnership. These four grantees have used their partnerships with researchers to improve their program design, and the evaluators have improved their research design by listening to the needs of the education stakeholders. Evaluators can provide formative feedback and real time 'dosage' information to project teams to help them understand outcomes and implement immediate program improvements.

In these partnerships, it is important to know that negative results are not failure but rather an opportunity to learn how to make mid-course corrections. A program might fail to achieve a given performance metric, or might not achieve fidelity of implementation. What the project team does with the information next is important. If the team uses that data and understanding to make "pre-mortem" changes, then the project can be successful. Partnerships and programs can be most successful when they operate in a culture of improvement. Having engaged stakeholders who want to understand what is working and for whom can be effective in improving student outcomes.

Research-practitioner partnerships are intended to be collaborative and supportive relationships. Improving programs and student outcomes are the goals and should be the center of improvement practices. These practices are cyclical. Continuous improvement includes performance assessment, data-driven improvement, support for implementation, and reassessing data. Program developers can engage in continuous improvement on their own, but may find efficiencies by including their external evaluators.

Resources

Adam, T., Gagnier, K. M., and Jones-Manson, S. (2022). Approaches to Build and Maintain High-Quality, Effective Partnerships. Cross-project summary produced by AnLar and the Office of Elementary and Secondary Education; Education, Innovation, and Research Program (EIR). Retrieved from https://oese.ed.gov/files/2022/07/EIR_Approaches-to-Build-and-Maintain-High-Quality-Effective-Partnerships_FINAL.pdf

Chow, K., Nakamura, J., Nunn, S., Wilderson, S.B., Haden, C., and Puma, M. (2021). Learning Before Going to Scale: An Introduction to Conducting Pilot Studies. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Appalachia Retrieved from https://ies.ed.gov/ncee/edlabs/regions/appalachia/resources/pdfs/Pilot-Study-Resource_acc.pdf

National Implementation Research Network resources and publications available online <https://nirn.fpg.unc.edu/publications-resources>

Shakman, K., Wogan, D., Rodriguez, S., Boyce, J. and Shaver, D. (2020). Continuous Improvement in Education: A Toolkit for Schools and Districts. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast and Islands. Retrieved from https://ies.ed.gov/ncee/edlabs/regions/northeast/pdf/REL_2021014.pdf

Supplee, L. (2022). Studying ways to improve the use of research evidence: Distinguishing data from research evidence. Blog post retrieved from <https://wtgrantfoundation.org/studying-ways-to-improve-the-use-of-research-evidence-distinguishing-data-from-research-evidence>