Learning in Context: Reflections on the Education Team's Approach to Evaluation

Ashleigh Halverstadt, Senior Evaluation and Learning Officer, S. D. Bechtel, Jr. Foundation

The S. D. Bechtel, Jr. Foundation, a spend-down foundation sunsetting in 2020, invested in four major education initiatives during its final decade of grantmaking. A firm believer in the importance of building and sharing knowledge, the Foundation also made significant, complementary investments in evaluation that were intended to help grantee partners improve their work and to capture lessons learned that funders, nonprofits, policymakers, and other education actors might benefit from. This essay offers a high-level comparison of the evaluation approach taken in each initiative and shares reflections on why we took the paths we did.

The key to a thoughtful evaluation strategy is a solid understanding of purpose, audience, and context. It is essential to ask: given the nature of the issue area and the intervention, who needs information, what information do they need, and why (i.e., how will they use it)? There is no singular best practice for evaluation in philanthropy; the best evaluation approach is the one that delivers the right information to the right people in the right way. The Foundation's approach to evaluation of its Education Program is an example of this philosophy.

The Foundation's commitment to young people is focused on advancing STEM education (science, technology, engineering, and math) and character development. In STEM, three initiatives – Math in Common®, Next Generation Science Standards (NGSS) Early Implementers, and New Generation of Educators – support the professional development of current and future teachers in California school districts and at California State University (CSU) campuses, helping to align instruction to new academic standards in math and science. The Foundation's fourth Education initiative – the National Character Initiative – supports national youth-serving organizations to strengthen the character development practices of the adult staff and volunteers who work with youth.

Evaluation plays a critical yet distinct role in each of these initiatives. For each STEM initiative, the Foundation has commissioned WestEd (and in one case, WestEd in partnership with SRI) to provide formative evaluation, summative evaluation, and/or technical assistance to grantees.[1] And, together with WestEd and Williams Group (our strategic communications firm), the Foundation has invested

---

[1] The difference between formative and summative evaluation can be best understood through their distinction in purpose. Formative evaluation explores a program's strengths and weaknesses to inform program improvement, while summative evaluation assesses the overall value and effectiveness of a program to inform decisions about the future of the program, inform the design of subsequent initiatives, and/or build knowledge for the field.

heavily in the dissemination of evaluation findings to key audiences. The Character Initiative has taken a more decentralized approach, focused on supporting grantees to build their own evaluation capacity. A retrospective, led by American Institutes for Research (AIR), will surface cross-cutting lessons learned from this approach based on the perspective of grantees and Foundation staff.

In all four initiatives, the Foundation is supporting communities of practice, reflecting a hypothesis that one of the most powerful things we can do to enhance learning is create opportunities for knowledge to be co-created and shared.

Considerations of purpose and audience, as well as contextual factors (including the nature of the program intervention and the state of the field) shaped our thinking as we charted these four paths.

Notably, in every case, the Foundation placed grantees and other external stakeholders at the center of the work, prioritizing their information needs above our own. This was an intentional design choice, consistent with a core belief that evaluation should not be an extractive exercise – that it should deliver value back to those who participate in it and ideally support collective learning in the field. When an external evaluator is involved, as in our STEM work, this approach has the added benefit of enabling strong, reciprocal relationships between the grantee(s) and the evaluator; these relationships are essential to the successful implementation of evaluation work.

Initiative	Evaluation Purpose and Audience	Evaluation Approach
Math in Common® Supporting K–8 Common Core State Standards in Mathematics implementation in ten California school districts to develop successful approaches and share promising practices statewide	Inform and support participating school districts in their efforts to implement math standards	Formative
	Capture lessons that can benefit all California school districts, as well as state education policymakers	Summative
	Strengthen the ability of participating school districts to collect and use data to continuously improve instructional practice	Technical assistance
NGSS Early Implementers Supporting K–8 Next Generation Science Standards implementation in eight California school districts to develop successful approaches and share promising practices statewide	Capture lessons that can benefit all California school districts, as well as state education policymakers	Summative
New Generation of Educators Supporting partnerships between California State University (CSU) teacher preparation programs and K–12 districts in their efforts to transform teacher preparation practices to align with new standards and district needs	Inform and support participating CSU campuses in their efforts to transform teacher preparation practices	Formative
	Strengthen the ability of CSU system and campuses to collect and use data to continuously improve teacher preparation programs	Technical assistance
Character Supporting youth-serving organizations in California and nationally to improve and sustain the character development practices of adult staff and volunteers at scale	Strengthen the ability of youth-serving organizations to collect and use data to continuously improve adult practice in service of youth character development	Capacity building
	Capture lessons that can benefit other funders and youth-serving organizations exploring similar approaches	Retrospective

Math in Common® Initiative

TIMELINE:
2013–2020

PROGRAM INVESTMENT:
$51 million

EVALUATION INVESTMENT:
$7.4 million

The Math in Common® Initiative, launched in 2013, supports ten California school districts to implement the Common Core State Standards in Mathematics (CCSS-M) in grades K–8. All ten districts are building systems of support to help their educators align instruction to the shared standards, but each district has taken a different approach, tailored to its local context.

Evaluation of the initiative is led by WestEd. For the first five years of the initiative, the primary purpose of the evaluation was to serve the learning needs of the districts. Because the districts are experimenting with a range of approaches to CCSS-M implementation, the evaluation was predominantly formative in nature during this period, designed to help the districts assess and improve their strategies in real time and to surface themes across districts that all could learn from. WestEd published frequent reports on the districts' progress and challenges, and although the districts were the primary audience for this work, the reports were widely disseminated in hopes of informing other districts and policymakers in their efforts to advance CCSS-M implementation. Throughout this period, WestEd also leveraged its deep understanding of the districts' work to provide responsive technical assistance on a range of issues, including how to collect and use data to improve instructional practice.[2]

As the initiative nears conclusion in 2020, the evaluation has shifted to focus on other districts and policymakers as the primary audience. The goal is to capture and disseminate lessons about what it takes to implement and sustain CCSS-M, so that others can learn from the successes and challenges of the Math in Common® districts. Summative evaluation is a greater focus in this stage, and technical assistance to the districts is reduced.

Using a mixed methods approach, including analysis of student performance on new standardized tests aligned to CCSS-M (which began in 2015), WestEd has examined shifts in the instructional approaches of teachers, changes in student proficiency, change management processes at the district level, and the development and sustainability of the initiative's community of practice.

The evaluation team has published 20 reports on the initiative, including a six-part series of summative findings.

Learn more about the Math in Common® evaluation approach and findings.

---

[2] In both the Math in Common® and New Generation of Educators initiatives, WestEd served as both an evaluator and a technical assistance provider. Some might perceive this dual role to compromise the evaluator's independence. We found, however, that the deep contextual understanding and strong relationships WestEd developed through this model had enormous benefits for the validity and impact of the work.

NGSS Early Implementers Initiative

TIMELINE:
2014–2020

PROGRAM INVESTMENT:
$21.6 million

EVALUATION INVESTMENT:
$3.2 million

Launched in 2014, the NGSS Early Implementers Initiative is a demonstration project that supports eight California school districts to serve as early adopters of the Next Generation Science Standards (NGSS) in grades K–8. WestEd's K–12 Alliance leads the initiative and implements a structured train-the-trainer model of professional learning. K–12 Alliance provides training to a core team of teacher leaders and administrators in each district who then train and support a second tier of teacher leaders. Together, all of these leaders spread NGSS to other science teachers across their districts.

The NGSS Early Implementers Initiative is evaluated by WestEd, independent of the K–12 Alliance. Since participating districts are in the vanguard of transitioning to the standards, the evaluation – indeed, the initiative itself – is intended to provide early insight to other districts and policymakers. In fact, the California State Board of Education and the Department of Education encouraged the Foundation to launch the initiative to advance statewide understanding of how to implement NGSS effectively. Formative evaluation is not a priority for this external audience, nor is it appropriate for an initiative with such a well-codified program model. Instead, WestEd focuses on a form of rolling summative analysis, regularly pushing out reports on how and to what extent various aspects of the model are working.

The evaluation is exploring district approaches to implementing NGSS and the integrated middle school model of science instruction, efforts to build administrator and teacher leadership, shifts in the instructional approaches of teachers, and changes in student access to and engagement in science. Since standardized testing aligned to NGSS only began in 2019, an analysis of the initiative's impact on student proficiency is beyond the scope of the evaluation. Instead, WestEd is using a blend of survey, interview, and observation data to track the initiative's progress.

The evaluation team has published 10 reports to date as well as a compendium of the key tools and processes districts have used to support NGSS implementation.

Learn more about the NGSS Early Implementers evaluation approach and findings.

New Generation of Educators Initiative

TIMELINE:
2014–2020

PROGRAM INVESTMENT:
$25 million

EVALUATION INVESTMENT:
$4.9 million

The New Generation of Educators Initiative, launched in 2014, supports the California State University (CSU) system, the largest preparer of teachers in California, to deepen partnerships between teacher preparation programs and the districts who hire their graduates for the purpose of preparing more new educators who are ready to lead instruction aligned with state standards in math and science. The initiative supports system-wide continuous improvement and data collection efforts, as well as eleven individual partnerships between CSU campuses and the local K–12 districts who hire their graduates. These partners work together to design preparation experiences that equip teachers with the knowledge, skills, and beliefs they need to lead effective, equitable instruction from their first days in the classroom.

WestEd and SRI International jointly evaluate the initiative. The evaluation is intended to inform the CSU's efforts to transform teacher preparation practices and to build the capacity of the CSU system and individual campuses to engage in data-driven continuous improvement. Formative evaluation was a priority from the outset, with some resources reserved for technical assistance and summative evaluation. As the work unfolded, evaluators encountered challenges with the availability, quality, and connectivity of several secondary data sources – requiring them to adapt their data collection plan and ramp up technical assistance. The evaluation team ended up playing a critical role in strengthening the CSU's data collection instruments and process and in establishing connectivity between the CSU system and the California Commission on Teacher Credentialing to more efficiently share workforce data. Formative reports continued to be produced along the way, and evaluators will contribute to a summative publication produced by the CSU Chancellor's Office.

In 2018, building on the CSU's increasing commitment to continuous improvement, the Foundation made an additional investment in technical assistance, commissioning WestEd to design and facilitate an improvement research fellowship for CSU faculty and district partners. The goal of this effort is to support CSU-district partnerships to deepen and sustain the use of data to improve their teacher preparation programs. There are early signs that the capacities built through this and other technical assistance from the evaluators will be among the most durable impacts of the initiative.

The evaluation has focused on understanding the components and quality of CSU teacher preparation programs, the pipeline of teacher candidates coming through them, and the capacity of the CSU system to engage in data-driven decision making.

Evaluators have published 10 publicly available reports to date.

Learn more about the New Generation of Educators evaluation approach and findings.

National Character Initiative

TIMELINE:
2014–2020

PROGRAM INVESTMENT:
$130 million

EVALUATION INVESTMENT:
Embedded in program investment and variable by grantee

Launched in 2014, the National Character Initiative partners with the national offices of thirteen youth-serving organizations as well as a network of California technical assistance providers (including the state department of education) to prepare and support the adult staff and volunteers who work with youth. Most national grantees are focused on identifying promising character development practices and then translating that knowledge into improvements to the professional development they provide to practitioners. The California technical assistance providers focus on infusing character development and social-emotional learning practices into the state's quality improvement efforts.

The initiative's approach to evaluation concentrates on strengthening the evaluation capacity of national grantees. Funds for program and evaluation efforts are integrated and are used in a variety of ways, with many grantees investing in internal evaluation systems and others working to build their knowledge base in partnership with external evaluators. This approach centers the grantees as the most important audience for evaluation and seeks to help them build, scale, and sustain systems for continuously improving programming in their individual contexts.

The decision to focus on evaluation capacity building was also shaped by the context for the initiative. The National Character Initiative was the last of the Education initiatives to launch, and most of the major investments were made in 2016 and 2017, with the work of each grantee progressing on its own timeline. Unlike school districts, youth-serving organizations do not have common standards like CCSS-M and NGSS to which they can align their practices or measurement, and program environments vary widely. The initiative comprises three grantee cohorts: large, national organizations; sports and play organizations; and nature-based organizations. An initiative-wide approach to evaluation would have required the grantees to come together around a shared definition and framework for character development. Given the relatively short lifespan of the initiative and the Foundation's intent to put the grantees' learning goals ahead of its own, an initiative-wide evaluation did not seem appropriate.

Still, the Foundation is committed to capturing lessons learned from the initiative that can build the knowledge base of the broader youth development field. In late 2019, the Foundation commissioned AIR to conduct a retrospective on the system of supports the Foundation has provided to National Character Initiative grantees. This project will seek to understand how grantees experienced the initiative and to surface insights that would benefit other funders and youth-serving organizations.

Learn more about the National Character Initiative and the efforts of individual grantees to build their evaluation capacity.

The Roads Not Taken

Charting the evaluation paths described above meant leaving other roads untraveled. In STEM, for example, the Foundation considered commissioning a meta evaluation of its three initiatives to surface cross-cutting insights about what it takes to transition education systems to new academic standards. We ultimately abandoned this idea, reasoning that the STEM evaluations are too divergent in approach, focus of inquiry, and methodology to allow a meta-analysis to be productive. In Character, we opted not to align grantees around common outcomes and indicators because we did not want to impose a shared measurement framework and felt that contextual evaluation approaches would have greater relevance, though some might argue that we left a field-building opportunity on the table.

Perhaps more notably, we have chosen not to put our grantmaking strategies under the microscope in the way some foundations do. For example, with Math in Common® we have documented the progress and learning of the ten participating school districts in great detail. However, we have not zoomed up a level to assess the grantmaking strategy itself – i.e., the choice we made to invest in a small cohort of districts and to document and disseminate their learning in hopes of informing and advancing math education statewide.

But in any foundation, choices must be made about what to prioritize. As a spend-down foundation nearing the end of its life in 2020, there is limited utility in evaluating our grantmaking strategies and practices because we have few opportunities to apply the learning. Although strategy evaluation might have generated useful learning for peer funders, with finite time and resources, we decided not to focus on our own theory of change as the subject of study.

Finally, in our concluding years, as we have been deepening our understanding of the role of equity in our work, we have been considering what it means for our evaluation practice. Efforts like the Equitable Evaluation Initiative have stretched our thinking in a healthy way and caused us to reflect on how our evaluation work might have been different had we designed it with the principles of equitable evaluation in mind. In some ways, our approach to evaluation has been consistent with these principles (e.g., we have centered our grantee and field partners as the primary users of evaluation; we have explored the impact of our work on some marginalized groups like English language learners). But, if we could begin again, we would invite our partners to help us think about this more explicitly and look for additional ways to align our evaluation goals, questions, and methods with our strong commitment to equity.

We hope these reflections provide a useful window into the considerations that shaped our Education evaluation approaches. There were many other factors that influenced us, of course, and the road was not always smooth. Some evaluation decisions were made under time and resource constraints, others were made in response to emerging opportunities, and some proved untenable and had to be adjusted mid-course. Yet, throughout the process, we attended to purpose, audience, and context as best we could – making choices that we hope were mostly sensible and will leave useful knowledge in our wake.

Key Takeaways

✓ Begin every evaluation effort with a deep understanding of the purpose, audience, and context for the work. These factors should drive the evaluation approach (not the other way around).

✓ Tradeoffs are inevitable: Not everything can or should be evaluated. Prioritize evaluation in areas where there is a meaningful use for it.

✓ Adaptation is key: Expect to adjust the evaluation approach as the work unfolds, the needs of stakeholders become clearer, and the constraints and opportunities of the context come into focus.

✓ Consider the role of equity in your evaluation practice from day one.

There is no singular best practice for evaluation in philanthropy; the best evaluation approach is the one that delivers the right information to the right people in the right way.

Ashleigh Halverstadt joined the S. D. Bechtel, Jr. Foundation in 2014 and serves as the senior evaluation and learning officer. She works closely with program staff to ensure that the Foundation's evaluation strategies are generating insight about what's working, what's not, and what can be done better in support of the Foundation's goals. Prior to joining the Foundation, Ashleigh was an independent consultant, providing fundraising, program evaluation, and strategic planning support to a portfolio of nonprofit clients. Previously, she served on the development team at New York Road Runners, where she managed public and private grants and spearheaded an initiative to design and implement an outcome management system for the organization's youth programs. She graduated with honors from the University of Washington, earning dual degrees in geography and political science, and holds a Master of Public Administration from the Maxwell School at Syracuse University.