Taking Stock

Evaluating the Impact of Education Grants

grantmakers for education

grantmakers ^{for} education

Grantmakers for Education is philanthropy's knowledge source for achieving results in education. Our mission is to strengthen philanthropy's capacity to improve educational outcomes for all students, which we achieve by:

- Sharing successful grantmaking strategies, best practices, and lessons learned that exemplify responsive and responsible grantmaking in education.

- Creating venues to collaborate on projects, share knowledge, develop leadership, advocate for change and debate strategies with other education grantmakers.

- Interpreting data, illustrating trends, and conducting research to improve the effectiveness of education grantmaking and to highlight innovative educational approaches.

Our efforts are informed by eight *Principles for Effective Education Grantmaking*, designed both to guide funders in increasing their impact and to ensure GFE's services and programs help funders accomplish their goals for change.

Taking Stock

Evaluating the Impact of Education Grants

MEMBER BRIEFING

CONTENTS

grantmakers for education

Foreword

In today's climate of accountability, many grantmakers struggle to determine the impact of the activities they support. This report seeks to summarize and make available much of the information from a special member briefing Grantmakers for Education organized to help funders—especially those from foundations with modest budgets—confront this challenge.

Held in November 2007, the program provided tools and framework for how funders can make the best use of evaluation to match their objectives and resources. Although the unique learning format of the event and the give-and-take among participants are difficult to capture in a written document, we've endeavored to organize this report to capture key lessons and advice for grantmakers.

The program—and this report as well—was designed to meet the following learning objectives, identified through a survey of GFE members:

- Understand why and when to evaluate—including calibrating evaluation costs with the scope of a project and a funder's ability to pay—and how to frame useful evaluation questions.

- Help participants determine what counts as persuasive evidence and how to gather and use it.

- Explore different approaches for evaluating the results of grants to understand the trade-offs associated with them.

- Offer advice on how to talk with foundation boards to clarify expectations about what an evaluation can tell them and how long it will take to see results.

- Share strategies on how to identify and select a good evaluator, manage costs and determine who is the "client" of the evaluation.

The program was not designed as a technical seminar focused on methodologies, but as a practical seminar to help frame grantmakers' evaluation options and provide a disciplined way to think about whether and how to evaluate initiatives in light of their foundation's resources.

The sessions followed a variety of formats, including formal presentations, panel discussions, small-group analysis of case scenarios and a "fishbowl" conversation in which the group observed an evaluator interviewing a grantmaker about her evaluation needs. What made the briefing unusual, participants agreed, was the high level of candor among the grantmakers and evaluators who attended.

It was a rare opportunity to learn about the practical side of evaluation, its strengths and limits, and its potential to sharpen grantmakers' thinking.

To help readers navigate the report and to probe topics where they want to learn more, each section of this report opens with a set of key questions to be answered and includes sidebars or text sub-heads to clearly flag suggested advice to grantmakers.

WHAT IS A THEORY OF CHANGE?

In framing questions for funders about what and how to evaluate, several speakers pointed to the importance of a "theory of change."

Simply, a theory of change is a statement or flowchart or plan that explains how an organization's intended impact will actually happen—the cause-and-effect logic by which organizational and financial resources will be converted into desired social results. For a grantmaker, a theory of change makes clear what "success" looks like and the assumptions behind how certain grants will lead to certain impacts.

Used often by funders as a way of clarifying their grantmaking strategy and to help drive choices about strategic trade-offs, speakers said a theory of change also was an essential tool for making decisions about evaluations.

RESOURCES

More information about what a theory of change is and how to craft one is described in the section of the report beginning on page 16.

In addition, grantmakers may find the following resources helpful:

ActKnowledge and the Aspen Institute Roundtable on Community Change. In an ongoing partnership, the two organizations have assembled an extensive online library of materials, tools, and background reading on theory of change and strategic planning. See www.theoryofchange.org.

Mapping Change: Using a Theory of Change to Guide Planning and Evaluation. This brief guide from GrantCraft explains how grantmakers can use a theory of change approach in their work, with grantees and inside their foundations. Available for purchase or free download at www.grantcraft.org.

Bringing Strategic Clarity to Your Grantmaking: Crafting and Using a Theory of Change. This report from Grantmakers for Education summarizes discussions during a special program track at the GFE annual conference in 2006 on crafting a theory of change and using it to increase strategic clarity and impact. See www.edfunders.org.

1 Beginning with the end in mind: A grantmaker's guide to evaluation

MICHAEL BANGSER opened the meeting with a primer on key questions and concepts to understand in deciding when and how to evaluate an effort. His observations are summarized in this chapter.

Bangser's presentation identified the questions that every grantmaker should ask up front—before deciding to evaluate—to make sure that evaluation findings are useful and scarce resources are used wisely. What kinds of evaluation approaches are most commonly used, and which are best suited to which purposes? When does it make sense not to evaluate?

Bangser is a consultant who is a former senior vice president of MDRC and former president of the Hartford Foundation for Public Giving.

Simply, evaluation is learning with a purpose.

And, specifically, it is a structured, cost-effective process that produces reliable answers to important questions and then uses those answers to help make better decisions. Ideally, funders should consider each of these components of this definition when they develop an evaluation plan.

Evaluation should be a *structured process*—one that starts at the beginning and not at the end.

All foundations process information—anecdotal or otherwise—before they make decisions. It only makes sense to think about how to collect, process and analyze information in a structured and purposeful way. Timing also is important; funders should think of this issue from the outset, embedding a clear timeframe into their grantmaking decisions and their discussions with grantees.

Evaluation should answer *important questions.*

As American humorist James Thurber once said, "It's better to know some of the questions than all of the answers." Some of the questions that need to be answered include:

- What do we want to evaluate? Individual grants? Clusters of grants? The entire portfolio? Non-grant aspects of the funder's operations, such as its role in advocacy or convening?

- What do we want to learn? Since grantmakers often have multiple goals, they need to clarify at the outset what they truly want to learn and how they will define success. Here are possible—sometimes complementary, but very different—goals for an evaluation: to learn about needs and opportunities; to figure out whether or not advocacy efforts

affected public policy; to gauge the effectiveness of specific program interventions or policy changes; to improve program operations (see sidebar for more information).

- What can be learned, and what can't? Some questions aren't readily measured or quantified. "Not everything that counts can be counted," said physicist Albert Einstein, "and not everything that can be counted counts." Be wary of the tendency to evaluate what's most measurable rather than what's most important. For example, evaluations often measure activity levels or outputs rather than outcomes or impacts.

Evaluation should yield *reliable answers.*

According to the journalist Janet Malcolm, "Almost everything we know, we know incompletely." Yet it's important to distinguish between incomplete information that is nonetheless informative and incomplete information that is misleading and therefore counterproductive. Here again, a few questions can bring things into focus:

- What are our assumptions about causation? To distinguish important information (and gaps) from unimportant ones, grantmakers need to be clear about what they think will cause the change they want to see. And then there's the counterfactual: Might the change have occurred anyway, absent the intervention?

- What is our theory of change? By articulating a theory about how the change is likely to happen, grantmakers can increase the chances that an evaluation will consider the right indicators at the right points of time (see report chapter beginning on page 16 for more information about a theory of change).

- What about data? Are high-quality, reliable data available? How will the foundation (or someone else) gather and store them?

- How should we interpret the results? Grantmakers need to be rigorous about the significance of findings—and also distinguish among types of significance. Results that are significant statistically might not be significant in terms of policy. It's also important to manage one's own expectations, remain aware of uncertainty and consider the importance of looking at subgroups. Avoid the "lure of simplicity," evaluator Patti Patrizi has warned, but don't get bogged down in technicalities.

Evaluation should lead to *better decisions.*

Of course better decisions are good, but it also pays to ask about the importance of the decisions at hand and the consequences of getting them wrong. Is an evaluation a diversion of resources from other work, or is it a means to strengthen grantmaking? Think of an evaluation as the beginning rather than the end of the discussion; use the evaluation to frame questions, choices and areas for improvement rather than (necessarily) providing definitive answers. Questions that lead to better decisions include:

- Who's the audience for the evaluation? Will the funder's board or staff use the evaluation to improve grantmaking decisions? Or is the evaluation really aimed at grantees, policymakers or practitioners?

- What would constitute persuasive evidence? What evidence will be persuasive to the intended audience? How reliable and timely do the findings need to be? Is the project "mature" enough to draw firm conclusions? Will the audience base decisions on research evidence, or does it need other sorts of information as well?

- Are the findings generalizable?
  How relevant are the findings to other locations or conditions?

- How will the findings be communicated?
  If an evaluation is going to influence decisions, people need to know the results. Make a point of developing a communications strategy early in the evaluation process.

Evaluation should be *cost-effective.*

Certain factors affect both the cost of an evaluation and its ultimate value. A lot depends on who carries out the tasks associated with the evaluation. Foundation staff? Evaluators? Grantees? Other questions include:

- Is the expense in proportion to the potential value? Consider the cost of the evaluation in relation to the size and importance of the grant, how it fits within the foundation's overall grantmaking portfolio and the importance of the decisions to be made.

- Who pays? Engaging grantees in the evaluation is good, but it's also good to do so without unreasonably shifting the evaluation costs onto them.

- Who has the capacity? Is it necessary to involve an independent evaluator? Or can the grantee or foundation carry out the necessary work? What about increasing ongoing capacity to "think evaluatively" within the foundation and among grantees?

Applying the right evaluation questions to different situations

To apply general observations about evaluation to the kinds of situations grantmakers confront regularly, participants examined and discussed three typical scenarios. Questions about evidence lie at the heart of every evaluation: How good does the evidence have to be for the kinds of questions you're trying to answer?

In each scenario, a grantmaker is faced with a decision about whether, and how, to support evaluation. Working in small groups, participants at the briefing analyzed each scenario. This section captures their collective advice about how best to navigate each situation and find a path forward.

PAUL GOREN, vice president of The Spencer Foundation, ROBERT GRANGER, president of the William T. Grant Foundation, and CHRIS TEBBEN, deputy director of Grantmakers for Education, designed and led the session.

Summarized on the following pages are the scenarios and a set of suggested opening questions a grantmaker could use to work through the choices at hand.

SCENARIO 1:
THE BOTTOM-LINE BOARD

Several members of your board have been raising the concern that the foundation is not really accountable to anyone and that you need to figure out if what you are doing is "making a difference."

This scenario rang true to many participants. "It reflects reality!" one exclaimed. "Boards want information like this, and they want it now." Indeed, several grantmakers had dealt with similar situations in their organizations. The group suggested that a grantmaker in this position start by working through a handful of initial questions:

- What does the board really want to know? What sort of data will answer their questions?

- Could a tighter logic model help align the foundation's resources with the board's expectations?

- If the foundation ultimately decides to commission an evaluation of its work, who would be the primary client? The whole board? The board and staff?

- What information would be collected in any evaluation? What information would be necessary, and what information would be "nice" to collect?

- How can the foundation create a reasonable match between the direct and opportunity costs of the evaluation and the decisions on the table?

A grantmaker who recently received a similar request from her board explained that she and her colleagues developed five "impact target areas" that board members care about. She and her colleagues are now figuring out where and how to get information that shows impact in those areas.

SCENARIO 2:
FANTASTIC FINDINGS

You receive a prospectus from a community-based organization that provides college-application counseling to low-income students in two high schools in a Midwestern city. They say that 95 percent of the kids they work with go on to college as a result of their services. They want you to support them to expand their services from two high schools to all eight high schools in the city. If you pay for the start-up to replicate the program to these new schools, the school district says it will absorb ongoing operating costs for all eight programs after two years. Your question is whether or not the program's claims about their success rate are defensible.

In this case, the group zeroed in immediately on a key question: What does 95 percent really mean? Getting a firm answer to that question—a necessary step before making any funding decisions—would probably mean digging deeper in the following areas:

- Which students are served, and how selective is the program? The more selective and voluntary it is, the less likely it is that the outcomes represent the program's true effects.

- What does the intervention entail, and what is its logic model? Do program activities match program goals and the target population? In other words, is there a logical connection between the program's services (inputs) and results (outputs)?

- How can we know the program is making a difference? What is the benchmark for comparison? How strong is that benchmark given concerns about selectivity?

Beyond those questions, participants said they'd want to know more about certain subjective issues regarding the organization and the model. To many, the grant-seeker's extravagant claims and thin evidence suggested a need for caution.

SCENARIO 3:
SIGNIFICANT STAKES

You've been asked to fund an innovative
staff-development program for middle-
school math teachers to improve their math
knowledge and thereby improve students'
math achievement. After the program runs
for three years, you'll want to know if it made
enough difference to warrant being picked
up by other school districts. A lot of resources
are at stake, so you are going to push for a
strong evaluation.

As they began to discuss this scenario, participants
found that they would need much more information
about the intervention itself before they could move
to replication and scale-up issues. They also realized
that good decisions would depend on having answers
to some key background questions: Who are the
teachers? What are the schools like?

Assuming a grantmaker could get answers to those
questions, they imagined planning an evaluation
with a series of further questions that might look
something like this:

- The project's logic model says the intervention
  will increase teachers' knowledge. How will we
  know if that has actually happened?

- How will we measure change in student achieve-
  ment? How much change will be necessary to
  call the program a success?

- What are the classrooms like after teachers go
  through the program? What will have changed, and
  how do those changes raise student achievement?

- How much will it cost? $5,000 per teacher?
  $10,000? How much does it cost to make a significant
  difference? Do high-cost and low-cost models
  yield significantly different results?

A lot could be learned from a formal evaluation of a
program like this, they concluded, and advice from a
researcher would be enormously helpful in clarifying
what data can be gathered and assessed and how any
data should be judged. Also, it's important to ask rigor-
ously how much you would really need to know to justify
replication of the program and ongoing expenditure
by the school district—maybe more than initially
appears to be the case, maybe less.

CLOSING ADVICE:
WHAT DECISIONS ARE ACTUALLY AT STAKE?

Robert Granger summed up the session with thoughts
about clarifying initial questions. "It's helpful," he said,
"to keep in mind what's actually at stake. Figure out
what decision is actually on the table."

To do that, he advised, find out what key people really
want to know—and do it "in a low-stakes environment."
Check in privately with your board, your staff, potential
grantees, educators and school-district officials before
putting them together with an evaluator: "People
don't want to feel stupid publicly," he said, "but in
fact they rarely want to know as much as an evaluator
thinks they do."

Granger closed with an observation about evaluation
as a tool for policy advocacy. "People in Washington say
you need a big research study to communicate with poli-
cymakers," he explained, but it's not true. "In my experi-
ence," Granger said, "people in government have low
standards of evidence. They want to know fairly simple
things about what works, where it works and how much
it's going to cost." Formal evaluation can help answer
those questions, but only if the evaluation fits within a
wider effort to have an effect on policy.

2 Managing practical evaluation challenges: A discussion between a grantmaker, grantee and evaluator

Designing an evaluation is clearly a lot of work. Using a roundtable-discussion format, program participants got to listen in as three people who approach evaluation from different perspectives—a program officer, a grantee and an evaluator—described the nuts and bolts of the process. Their advice covered selecting an evaluator, estimating costs, establishing relationships and handling controversial findings.

We've summarized highlights from this discussion in this chapter. The section headings in boldface flag the key take-aways and the speakers' advice to grantmakers.

Roundtable participants included LANDE AJOSE, director of BTW informing change, HELEN WILLIAMS, program director of The Cleveland Foundation, and JOHN EASTON, executive director of the Consortium on Chicago School Research. MICHAEL BANGSER moderated the discussion.

Choose the right evaluator.

MIKE BANGSER: Helen, you're a grantmaker. What are some of the factors that go into your decision to use an external evaluator?

HELEN WILLIAMS: Using an external evaluator is a huge investment, so you need to be clear about what value an evaluation is going to add to the process and what you hope to accomplish. Evaluators add value in three major ways: first, they bring credibility to the process and the results; second, they have expertise and knowledge that's relevant to the field; and third, they do the work so the grantee doesn't have to. If you're going to capitalize on that value, it helps to be clear from the start that you're going to be willing to accept the results of the evaluation and take it to the next level. There's no point in funding a report that just sits on the shelf.

BANGSER: How do you choose an evaluator?

WILLIAMS: I look at evaluation reports in the relevant field to see how different evaluators are reporting and framing the issues. I also talk with people who are networked with evaluators to see what they think. After that, I choose two or three potential evaluators and talk with them systematically about the scope of work and what the foundation wants. Then I invite proposals. I involve others in reviewing the proposals, including the grantee, some of my colleagues at the foundation and other funders.

BANGSER: Helen, have you ever done an RFP for an evaluation?

WILLIAMS: We did one for a study of charter schools. It worked out all right because there's not a huge universe of people out there who are capable of the sort of work we wanted. We figured out the universe and sent letters of inquiry. Another idea is to do an RFP but don't ask for full proposals. Specify clearly what you

want to learn and ask for a two-to-three-page letter. It doesn't take too much of an evaluator's time to respond, and what they send is useful to you and to them.

BANGSER: What are some things you ask about an evaluator's qualifications?

WILLIAMS: I ask for references from funders and from grantees. I also want to know specifically who from the firm will do the work. And I check into their communication skills. Are they going to come in and say, "I'm the expert"?

LANDE AJOSE: I think it makes sense to do interviews. Talk to as many people as possible and find out about their people skills.

BANGSER: What about an initiative with other funders involved?

AJOSE: We did one project that involved three funders. Our first job was to put them on the same page. My advice is to look for an evaluator who can occasionally recognize the need to stop and get things straight. Sometimes that means standing up to the funder.

Agree on clear expectations for what is being evaluated.

BANGSER: Lande, imagine you've been chosen by Helen to evaluate a project. What's in your mind? Are you and Helen thinking alike, or is there tension between your points of view?

AJOSE: There might be some tensions or differences at the beginning. I need to get those out in the open so we can discuss them and come to agreement. So, for me, the big "E" is expectations. I try to use our first conversation to frame those. We're both looking for defensible results. We can agree on that. Then there might be divergence in how we're thinking about time and resources. Sometimes a foundation doesn't

want to invest time up front. They're often eager to get results too soon, and we need to address that. The objective is to "right size" the project up front, to ensure it's the right scale and scope for the information the funders needs. If you want to build the field, you need a lot of investment on the front end. Also, you'll want to be mindful that, later on, there might be tension about disseminating the results, especially if we uncover unflattering findings.

BANGSER: What are the key drivers of cost?

AJOSE: The biggest item is data collection—and it almost always takes more money to get defensible results than funders want to pay or can pay. Interviews, for example, are always much more expensive than funders think they're going to be because there are a lot of steps: you need to design the instrument, conduct the interviews, report on them, analyze your findings. It adds up.

Figure out the right measurements.

BANGSER: When we talk about evaluation, there's a tendency to look for numbers. People want to go to the quantitative side, but sometimes qualitative methods go deeper. Lande, how do you talk with clients about what to emphasize?

AJOSE: Quantitative and qualitative methods produce different information. Do both, if you can. We ask our clients, "What do you want?" Very often, the answer is "a good story and data to back it up." Does that mean a data set? Ongoing analysis? Maybe. It depends on so many factors. Recently, everyone has been saying that their board wants a "dashboard." Personally, I have a preference for really well-done case studies. You can get a lot of data into a case study. You can present a wealth of nuanced information from many perspectives.

Establish trust—especially when the results are negative.

BANGSER: John, you're in an interesting spot because you're an evaluator and a grantee. How do the parties involved in an evaluation develop trust and relationships, given power differentials?

JOHN EASTON: Funders want successful grantees, so there's a basis for trust right there. But it can be very hard if there's an us vs. them dynamic between the funder and the grantee. The relationship between funder and grantee builds over time. It's harder to establish if you're a one-time, short-term funder, easier if you do multiple grants over a longer period. It's also helpful if future funding doesn't hinge on the outcome of the evaluation.

BANGSER: Can you talk a little about establishing trust?

AJOSE: The main thing is to be clear from the start about who the client is, whether it's the grantee or the foundation or someone else. We try hard to define the client up front. We ask the funder, "Is it you? Are you the client? Or is it the grantee?" We don't allow dual clients. In one case, a funder gave grantees 10 percent of the amount of their project grants and told them to choose an evaluator from a list of three. One organization chose us, and we had a conversation about what the grantee wanted to know. The arrangement put the grantee firmly in the role of client.

HELEN WILLIAMS: As a funder, you need a "contract" up front with grantees about what you want to know and how it will be used. We've had grantees who have been funded for years without any questions from us about things we'd like to know. When we change that pattern by suggesting an evaluation, they sometimes see it as changing the game. It's not always easy.

BANGSER: How do you handle controversial findings?

EASTON: We try to adhere to a "no surprises" rule. We agree in advance that we won't put out a study until top leadership has seen it. We also get lots of input from constituencies that matter—the school district, parents, the funders—before we release a report. We require a very careful review, but we have a firm policy that the author has the final say. If an author chooses not to make a suggested change, we require a written explanation of why not.

AJOSE: I was involved in the release of one controversial report, and we realized that we needed to get buy-in during the review process. We drafted the report and got feedback from all parties. It seemed like a long vetting process: it took about three months. But people needed a chance to look at our findings closely and say, "We think this is wrong." That gave us a chance to go back to the data and say, "We know you think it's wrong. But it's right, and here's why." The situation can get tricky earlier than that if the evaluator sees a big problem. Sometimes we shift from being the evaluator to being an organizational coach if we see very devastating stuff.

3 Measuring changes in student achievement: Reflections by two evaluators on competing ideas and different approaches

Many education grantmakers ultimately aim to increase achievement gains, but measuring student achievement in an evaluation is anything but straightforward. To help participants understand different ways of measuring achievement and interpreting any results, the program included an exchange between researchers JOHN EASTON, executive director of the Consortium on Chicago School Research, and JOAN TALBERT, co-director of Stanford University's Center for Research on the Context of Teaching. Easton works primarily with quantitative methods, while Talbert works mainly with qualitative approaches.

At the program, these two leading researchers talked about different approaches for measuring achievement, the kinds of information those approaches yield and how to interpret the results. The speakers also looked at the question of persuasive evidence: How much information, and what kind, is sufficient to make a difference in important practical decisions?

This section summarizes their observations and suggestions for grantmakers.

Be clear about how resources link to results.

JOHN EASTON: To measure change in a meaningful way, it's important to ask by what mechanisms we think student achievement will go up. What's our theory of change? The assumptions are often implicit, not stated, and evaluators need to work with stakeholders to clarify those assumptions as part of the research. What's the theory? Is it tight? Or, if an intervention is not successful, where did it break down?

Why, for example, do people think small schools will make a difference? The theory of change for small schools goes something like this:

Small schools have some autonomy and freedom.

↓

Kids get to know teachers; teachers work together.

↓

Instruction and student engagement improve.

↓

Student achievement increases.
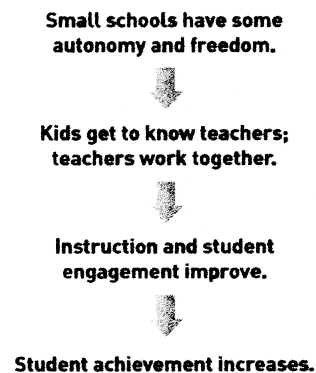
We did a study of small schools and found that they got better attendance but no change in student achievement. A theory of change lets us look for measures at each step, to see what activities are making a difference or not.

JOAN TALBERT: We tend to focus on how to measure that last step. In John's framework, we ask, "What changes in student achievement do we expect to see as a result of better instruction and improved student engagement? And how can we measure those changes?"

One thing we look at is the quality of implementation. Interventions that are implemented in a lot of schools (10 or more) are inevitably implemented unevenly. It's possible to look at variations in implementation and connect those to student outcomes.

Another possibility is to look at measures from sources other than tests. For example, we sometimes identify relevant indicators from national studies and surveys. In one case, we used a survey with items about student effort and attitudes that we could compare with national norms. The data were soft, but people actually care a lot about those outcomes. The funders were also pleased to see comparisons with national norms.

Determine the best—and most realistic—experimental design.

TALBERT: As an evaluator, you need to have a way of knowing what would have happened without the intervention. Randomized-control designs (often called the "gold standard" of research because they test the efficacy of an intervention or treatment by randomly assigning them to some subjects and not others) are very unusual in education. There are serious questions about feasibility, cost effectiveness and ethics. You also need to ask if the "treatment" will be stable enough to study it and make no refinements.

But there are some other, less intensive, less expensive evaluation designs that can add very useful insights into how a program is affecting student achievement:

- If you have good measures of the things you care about, you can often look at matched pairs of schools and compare the results. This works best if you have lots of cases—20 or so—to gauge whether students are learning more.

- You can also look at change over time, as long as you have consistent data over time. In California, for example, the data systems have changed numerous times, which limits our work. As evaluators often say, "If you want to measure change, don't change the measures."

- You can also look repeatedly at a single point in the continuum. One study looked at fifth-graders each year for several years. This can be a good alternative to following one cohort, and it's sometimes easier to get reliable data.

EASTON: Every two years, we do a survey of all Chicago teachers and students in many grades. It helps us see the effects of interventions, even if we don't learn about ultimate outcomes. But early survey results suggest that we can expect future growth in student achievement.

The truth is, there's always a "tortured link" between an intervention and student achievement, but survey information can tell you things with real policy implications. For example, teacher surveys can give important feedback about how much teachers trust each other and their schools. The results help foundation leaders see the value of actions designed to strengthen the professional community.

Decide what's good enough: How much change is enough change?

EASTON: You need to think about the long term. Is change at this rate likely to get us where we want to go? You also need to think about the quality of the measures. In Chicago, test scores have gone up. Is that a victory? Maybe, maybe not. They're lousy tests.

TALBERT: Only a fully rounded portrait can answer that question. You need statistical evidence, and you also need case studies. A successful case lets you look at how forces come together when an intervention works well. If you can describe what it took in that successful case, you can then ask how likely it is that those components can be replicated or spread. What if it takes a strong principal? How many strong principals are out there? Can you grow enough principals to support the intervention?

Use results to improve the program.

TALBERT: If you're a funder and you've invested a lot in an intervention, evaluation can validate the resources you've already spent. But the field has changed. No one wants an evaluation report at the end of the process anymore. More evaluation is being done earlier, when it can affect the intervention.

EASTON: It seems as if a shift has taken place in how we think about evaluation. We used to ask, "Is the intervention working?" Now, we're more likely to ask, "How well is the intervention working?" and "Can it work outside a certain narrow set of conditions?"

4 Evaluation with a small "e": How to gather useful information when you can't afford an evaluation

Recognizing that many meeting participants come from smaller organizations with no or limited evaluation budgets, another session offered ideas for how grantmakers can use their grant-reporting process—a process already in place at most foundations—to answer questions about performance and results. LANDE AJOSE, director of BTW informing change, provided key frameworks and ideas—all summarized in this section.

Key questions covered by Ajose included: How can you learn from your grantmaking if your foundation doesn't have an evaluation budget? How do you capture meaningful information regarding small grants when an independent evaluation would cost more than the grant?

Identify the challenge and get clear about expectations.

The session opened with participants naming some of the evaluation challenges they face as grantmakers, especially in smaller foundations, including:

- As a new foundation, we didn't have specific reporting requirements, so we didn't get back useful results.

- Our trustees are impatient. They expect too much, too quickly.

- It's hard to find a balance so you can get enough information but don't spend too much time on evaluation.

- We have different values about evaluation among our staff and between our staff and board. It seems as if leadership is comfortable making decisions based on gut instinct.

- We'd like to find a way to measure the cumulative impact of small grants made over many years.

- It's hard to find appropriate evaluation techniques when your organization makes small one-year grants.

The group confirmed that expectations are the main issue in most of the challenges they mentioned. It's important, they agreed, to try to clarify expectations and align the scope of your evaluative efforts with the resources available and the stakes involved. It's also important to be sure expectations are clear both among staff members and between the staff and the board.

As part of this session, speaker Lande Ajose reviewed the components of a sample theory of change with participants (see figure 1 below), noting that the exact components of a theory can vary; at its simplest level, a theory of change should show how resources a nd activities will yield specific results and outcomes.
She then showed how one foundation had filled in this framework for its arts program (see figure 2 below) and other programs. In addition, she discussed how the foundation finally aggregated its individual program strategies to develop a single theory of change to guide the foundation's entire approach; the goal was to cut across silos to see if the foundation's approach was consistent.

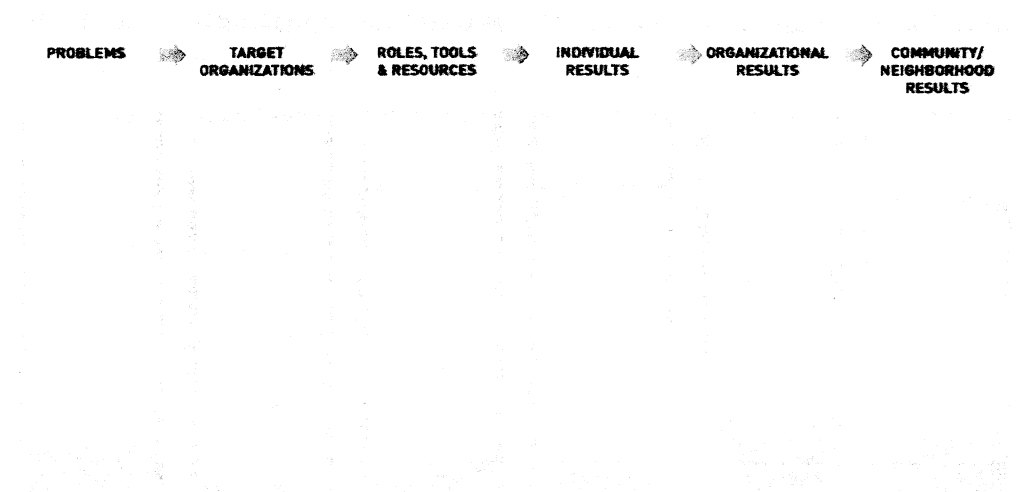FIGURE 1: Arts Program Investment Theory of Change

PROBLEMS	TARGET ORGANIZATIONS	ROLES, TOOLS & RESOURCES	INDIVIDUAL RESULTS	ORGANIZATIONAL RESULTS	COMMUNITY/ NEIGHBORHOOD RESULTS

FIGURE 2: Arts Program Investment Theory of Change

PROBLEMS	TARGET ORGANIZATIONS	ROLES, TOOLS & RESOURCES	INDIVIDUAL RESULTS	ORGANIZATIONAL RESULTS	COMMUNITY/ NEIGHBORHOOD RESULTS
Children and youth from low-income families have limited access to the arts. Low performing public schools are least likely to offer arts education.	Schools and arts organizations that work with low-income students.	Grantmaking: responsive, proactive, initiative, multi-year, core funding	Low-income children have expanded access to quality arts and arts education	Schools and community organizations are capable and positioned to use the arts as a means of building new or stronger connections among people with shared cultural backgrounds.	Arts programs serving low-income children are actively supported by community.
The pressure to assimilate into U.S. culture can create challenges to the expression and preservation of traditional cultures of immigrant communities.	Organizations working with immigrant communities	Network development: collaboration, convening, connecting	Youth are more connected to each other and to community arts institutions	Schools and community organizations are capable and positioned to use the arts as a means of building new or stronger connections across cultural divides.	Artists and arts organizations become part of community and neighborhood improvement efforts.
	Organizations working to preserve traditional arts and culture.	Knowledge development and dissemination	Individuals are more connected to their cultural heritage		Cross-cultural social networks begin to be built.
The arts could be used better to create connections across cultural, generational, religious and age divides.	Organizations that provide cultural commons — bridging cultural, ethnic and religious differences.	Public policy and advocacy	Individuals are more aware and appreciative of the culture of others.	Schools and community organizations are better positioned to sustain arts education.	
			Relationships form that cross familiar lines of age, culture, or socioeconomic status.	Organizational relationships sustained and evolve over time.	

17

Articulate the outcomes you expect to see.

Having a clear theory of change is at the heart of taking an evaluative approach to your work, even if you don't have the funds for an independent evaluation. A theory of change enables a foundation to articulate the outcomes it seeks.

The byproduct is more clarity about what you're doing and what you shouldn't be doing. It becomes a valuable planning tool. A funder can develop a theory of change for an individual grant, an initiative or an entire program area. When developing a theory of change, defining the problems to be addressed, or the purpose statement, is a crucial first step, according to Ajose. Evaluators typically spend more time on the purpose statement than on any other part of a theory of change. The purpose statement usually looks a lot like the desired ultimate result: the neighborhood or community impacts toward which the foundation is aiming.

Also, often board members focus on indicators rather than on broader social impacts. This is understandable: indicators help them define what they're trying to accomplish, and they're hungry for data. But a board's job is not to identify indicators; it's to identify outcomes. Getting the board to name the problems the foundation is trying to address and the ultimate results it wants to advance can help clarify the distinction between outputs and outcomes. It's all right for a foundation's stated desired outcomes to be broad: no single foundation can achieve ultimate results; at best, those will be accomplished through the efforts of many organizations.

Link your theory of change with grant reporting.

For a foundation, linking the theory of change with grant-reporting requirements has several benefits. First, it communicates the foundation's broader goals to its grantees. It also helps the foundation get more information about how a specific grant is contributing to a foundation's desired impact.

One foundation created a brief checklist that grantees and program officers use as a tool to gather qualitative data on how grants contribute to the foundation's intended outcomes. Without overburdening grantees, the reporting tool generates data on how the foundation's investments are influencing individual young people, schools, community organizations and the community as a whole.

Speaker Lande Ajose shared this example of one foundation's efforts to link its theory of change to a grantee report. In this report, grantees are asked to reflect on how their work is affecting individual, organizational and community outcomes that the foundation has identified in its theory of change.

Supplemental Grantee Reporting Form

Arts Program Investment

INSTRUCTIONS: This profile is to be completed by each grantee as a supplement to their Grant Reporting form.

Name of Organization and/or Grant: _____

Grant Number: _____ **Date Completed:** ___/___/___
                                                                    mm dd yy

Period of Review: _____

INDIVIDUAL RESULTS

1. Did this grant achieve the following individual-level results? *(check one box for each result)*

	Yes	No	Not Applicable
a. Low income children had expanded access to quality arts and arts education	❏	❏	❏
b. Youth were more connected to each other and to community arts institutions	❏	❏	❏
c. Individuals were more connected to their cultural heritage	❏	❏	❏
d. Individuals were more aware and appreciative of the cultures of others	❏	❏	❏
e. Relationships formed that crossed familiar lines of age, culture or socio-economic status	❏	❏	❏

*If your program did not intend to achieve this result, please check this box.

ORGANIZATIONAL RESULTS

2. Did this grant achieve the following organizational-level results? *(check one box for each result)*

	Yes	No	Not Applicable
a. Schools and community organizations grew capable of and positioned to use the arts as a means of building new or stronger connections among people with shared cultural backgrounds	❏	❏	❏
b. Schools and community organizations were capable of and positioned to use the arts as a means of building new or stronger connections across cultural divides	❏	❏	❏

*If your program did not intend to achieve this result, please check this box.

	Yes	No	Not Applicable
c. Schools and community organizations grew better positioned to sustain arts education	❏	❏	❏
d. Organizational relationships were sustained and evolve over time	❏	❏	❏

*If your program did not intend to achieve this result, please check this box.

COMMUNITY RESULTS

3. Did this grant achieve the following community-level results? *(check one box for each result)*

	Yes	No	Not Applicable
a. Arts programs serving low-income children were actively supported by the community	❏	❏	❏
b. Artists and arts organizations became part of community and neighborhood improvement efforts	❏	❏	❏
c. Cross-cultural social networks began to be built	❏	❏	❏

*If your program did not intend to achieve this result, please check this box.

5 Making the most of your evaluation resources: A case study of one foundation's options

To tease out and model the sorts of question a funder might consider when deciding whether or not to embark on an evaluation of a particular project, the program created a "fishbowl" conversation between seasoned evaluator JOAN TALBERT, co-director of the Center for Research on the Context of Teaching at Stanford University, and trustee ROBIN STEANS of the Steans Family Foundation in Chicago.

Looking at the Steans Family Foundation's education portfolio, they considered opportunities for meaningful evaluation and examined the trade-offs involved in terms of cost, effort and possible impact, using a matrix Talbert distributed to session participants. The conversation was wide-ranging, approximating the sort of initial interview Talbert might have with a prospective client.

In this section, we've captured highlights from their conversation that offer especially practical advice for how to make decisions about evaluation.

Profile: The Steans Family Foundation and North Lawndale

The Steans Family Foundation is a small family foundation that has concentrated for the past dozen years on improving the quality of life in North Lawndale, a low-income Chicago neighborhood. The foundation has worked across sectors, but much of its support has focused on education.

The North Lawndale neighborhood includes 27 schools, most of which struggle with the problems that typically beset urban schools: high student mobility (60 percent in some high schools), high teacher turnover and a concentration of the least experienced teachers in the school system. The neighborhood has high rates of violence, unemployment and other challenging factors that influence children's lives. Eighty percent of men have some contact with the criminal justice system, and only four percent of adults have been to college. The foundation sees the value of stabilizing the school environment to benefit children who are subject to many negative, destabilizing forces.

The foundation is supporting three education initiatives that might warrant evaluation:

- A program based in a charter high school that provides students with more consistent, sustained attention from school guidance counselors.

- The Umoja Program, located in a large public high school that is designed to increase student engagement, improve school culture and help students prepare for college.

- The Building Bridges project (which the foundation sponsors with two other funders in collaboration with Chicago's Juvenile Protective Association) that places trainee mental-health counselors in participating schools.

Evaluation Purpose	Evaluation Questions	Evaluation Design	Considerations/ Issues*
Measure impact of the initiative; summative evaluation.	To what extent did the target group(s) benefit from the initiative? Was the design effective on the whole?	Use one of three general approaches: - experimental (randomized control group) - quasi-experimental (comparison group; trend analysis) - non-experimental (statistical estimation)	How well are the outcomes measured (beyond state tests)? Is it feasible to obtain a control or comparison group? Is selection bias controlled?
Monitor grantee performance.	Are the grantees doing what they said they would do? Is the work high-quality?	Measure and report all facets of the program design (quantity and quality indicators).	Do grantees have reliable mechanisms for tracking their work? How well is quality being assessed?
Test assumptions of the initiative; build knowledge to inform future initiatives.	Did the assumed cause-effect relationships hold up? Under what context conditions; through what processes? What were unintended outcomes?	Develop logic model linking design to outcomes. Measure each fact of the model and evaluate expected relationships. Document processes within and across diverse cases.	Do stakeholders agree on the model? Are all facets of the model measured well, over time? Is there sufficient breadth and depth of longitudinal data?
Ensure ongoing feedback to improve the project.	Are participants experiencing the project in the intended way? Does this vary across participants/ sites?	Monitor participant responses to facets of the program. Develop and use "warning indicators."	How reliable and predictive is short-term information? What structures and norms support ongoing learning?

* Considerations of cost and feasibility apply across the evaluation types.

After a quick analysis, Talbert and Steans zeroed in on Building Bridges, which seemed most ripe for evaluation and most likely to produce lessons of value to the foundation and others.

Originally based in three elementary schools and recently expanded to four, Building Bridges places trainee mental-health counselors in participating schools for extended residencies supervised by senior counselors. The schools get the benefit of the additional services; the Juvenile Protection Association gets entrée to schools and training opportunities. The association believes that families underuse mental-health services and that the project could increase access.

Throughout the discussion, Talbert and Steans checked their thinking about a possible evaluation against a matrix of evaluation types that Talbert distributed to session participants.

JOAN TALBERT: Before conducting an evaluation, a funder needs to build consensus on three issues: Why evaluate? What would we evaluate? And how would we evaluate, or what would we hope to learn?

ROBIN STEANS: Here's why we might be interested in evaluating Building Bridges. We know there's not enough money in the system to put more counselors into elementary schools. So, does the counselor-and-intern combination make a difference? It stretches dollars, but do we get traction? What accounts for variation among schools? And how do we know if we're having an impact? The schools are so troubled, and the kids so needy, but does the program raise student achievement?

TALBERT: The Juvenile Protection Association probably has questions, too. Staff there might want to know if they get more reach into communities and with families, but those are not necessarily the outcomes that interest you. One might want to know about how the counselors influence classroom environment. What comes first, classroom change or kids' resilience and capacity? You'd need to work out a detailed logic model. What about teachers' capacity to reach kids? Is that part of the model?

STEANS: Teachers seem to be weaving things from the project into their classrooms. That wasn't built into the grant, but it seems to be happening anyway. We're encouraging them to make that aspect of the work more intentional. Now that the Juvenile Protection Association has a presence in the schools, it's doing things like offering workshops for teachers or providing interventions after a child dies. Teachers are going to the counselor's office to talk about their own issues. This sort of development was expected and hoped for, but it wasn't planned for.

TALBERT: This is a perfect example of an important lesson: Leave room for unanticipated outcomes.

It sounds as if some of the preconditions for evaluation are present with this project. The grantee is capable and well-equipped. There's an appetite for learning among the partners. There seems to be a safe learning space. The grantee is probably already collecting some data, although they might not know much about why changes are happening. They might want better information on their training program.

Here are two ideas for what might be possible and valuable with relatively little money:

- First, an evaluator might look at things that are at a distance from the actual intervention—things that would be hard for the grantee to see on its own, such as partnerships that allow the program to

work in different schools. An evaluator could look at student data across schools to get a better picture of preconditions in the range of schools.

An evaluator might also look for "warning indicators" that show when an intervention is in danger of running into trouble. This could be very helpful if you're interested in replicating the project. In one formative evaluation, for example, we identified signs that an implementation was beginning to play out in bad ways. We helped the district develop a tracking system that focused on translation problems that people at the top rarely think about. For example, do teachers see counselors as the solution to all their problems? Do they concede responsibility? Your story about teachers seeking counselors out for their own therapeutic issues sounds good to me; but, on second thought, what does it really mean? You probably also have questions about impact. Are there schools in comparable situations that aren't doing this?

STEANS: There are also variations among the schools in the project. All of them have similar demographics, but there are large variations in environment. How do you explain that?

TALBERT: You might want to look across schools to learn from those variations without trying to assess student achievement outcomes. For example, you might see a connection between school climate and kids' social-emotional data or on-the-ground developmental outcomes. You might find some test-score outcomes that seem relevant, but in fact those would be pretty far down the logic chain.

STEANS: Is there an affordable way for us to sustain a dialogue with an evaluator and also conduct a useful evaluation?

TALBERT: It's important not to lose track of the importance of developing the capacity of grantees to measure their own work—partly because that capacity can improve their work, and partly because of cost. For example, if we wanted to do good case studies of the four Building Bridges schools—even if the evaluator were based in Chicago and had no travel expenses and even if the researchers were simply interviewing principals, counselors, and association staff and doing little or no classroom observation—the project would cost about $90,000 for one year.

STEANS: With a project budget of $180,000, $90,000 seems very expensive. The evaluation would need to be much more valuable than what we get from site visits by the program officer. Is the difference worth $90,000?

TALBERT: Projects tend to be perceived in terms of personalities, but it does sound as if this model could be operationalized. A case study could be helpful at capturing what needs to be part of the principal's job, what needs to be part of the counselor's job, what resources are necessary and other factors that would need to be in place for replication.

6 Parting thoughts

At the close of the program, participants came up with this compelling list of observations and pieces of advice for funders on evaluation:

- Ask the right questions even if you can't get the answers.

- Ask two big questions upfront: What should we measure? And how do we define success?

- Be aware of trade-offs in terms of time, money and buy-in.

- Theories of change and logic models have come of age. Look past the jargon to see what's helpful, probing "What are we trying to do? With what resources? How will we follow what happened over time?" The time dimension is crucial and needs to be connected to the evaluation.

- Integrating evaluation within grantmaking strategy helps frame your decision-making. Or, looked at from the other direction, the more you go down the path of evaluation, the more impact it will have on your grantmaking.

- Implementation matters. It's common to expect results too early, which can compromise quality. There's a tendency to think that the sites with the smoothest implementation will have the best outcomes. That may be the case, but it's not always true—especially if smooth implementation occurs because the program avoids serving challenging students.

- Even if you can't afford to evaluate, learn from other evaluations. Could larger foundations do more to fund learning for the field?

- Costs are variable. Some foundations assign flat percentages of their grantmaking to evaluation, but that doesn't always fit the need.

- Keep plugging away. Integrate evaluation into your work where you can.