

# Making Evaluation Work

---

*Youth Development Fund Learning Series #2*

By Neil F. Carlson

## INTRODUCTION

---

In 2000, the Edna McConnell Clark Foundation (EMCF) embarked upon a new approach to grantmaking in the field of youth development. Rather than making large numbers of grants across several program areas as the foundation had done throughout its 30-year history, the new approach focused on strengthening a select number of high-performing organizations, and helping them improve and expand the services they offered. The foundation settled on the single field of youth development and set about creating a new grantmaking strategy.

Today, the foundation operates very differently. Its primary strategy is to find and select youth organizations that are in a position to—and that desire to—increase the number of young people to whom they deliver high-quality and effective services. With the foundation’s assistance, potential grantees develop business plans, and if there is a good match between the organization’s objectives and the foundation’s goals, the foundation usually makes a substantial, long-term grant against the business plan’s milestones. Thus, as more young people gain access to better services, and as the organizations serving them grow stronger and more effective, those young people will have better lives. This, in dramatically condensed form, is the foundation’s theory of change.

The foundation has transformed itself from top to bottom to support this theory of change. Drawing freely on strategies from the for-profit sector like due diligence, business planning, and performance tracking, the foundation has remade the process through which it finds and selects potential grantees. It has created new internal structures and multidisciplinary teams to execute these new processes. New technology systems and data management techniques support these new structures and processes.

Throughout this transformation, however, something more fundamental has happened. The Edna McConnell Clark Foundation has become an outcomes-based organization. “Our way of ensuring that we, as a foundation, are serving the interests of young people is to invest only in high-performing organizations that are committed to monitoring their performance and demonstrating outcomes,” says Michael Bailin, EMCF president. “Similarly, we have to be rigorous about assessing our own work and showing what outcomes we are achieving.” As a direct result of this focus on performance, the foundation has placed new emphasis on evaluation at all levels of the organization.

Three years ago, evaluations consisted largely of assessments of foundation programs, first to monitor implementation, and then post-mortem, as they wound down. Now, evaluation-based structures, processes, systems, methods, tools, and techniques are fundamental to every stage of the grantmaking and investment process. For the first time in its history, evaluation tools and processes are integrated with every aspect of the foundation's work—strategic planning, program development, and ongoing management. Evaluation is central to the foundation's stated goals and objectives. It shapes the way the foundation locates and assesses potential grantees. It informs how the foundation approaches business planning with grantees and how it manages its ongoing relationships with grantees. Evaluation is also fundamental to the ways in which the foundation hopes ultimately to contribute to the nascent field of youth development.

This paper examines the pivotal role evaluation now plays in the foundation's work. It begins by outlining how the foundation's evolving theory led the organization to reassess the relationship between strategy, outcomes, and evaluation. Next, the paper examines the internal tools, structures, and processes that now support the foundation's work. The paper concludes with a look at how the foundation has helped grantees articulate or develop their own theories of change and institute the evaluation systems necessary to measure their performance, manage program quality, and monitor outcomes the young people participating in their programs achieve.

# Aligning Strategy and Outcomes

---

## EMCF's Theory of Change

Beginning in 1996, the Edna McConnell Clark Foundation began exploring a new approach to grantmaking. Instead of making a large number of grants across the existing five program areas, the foundation's new president, Michael Bailin, urged the foundation to consider strategies that strengthened individual organizations and the fields in which they practiced. "Our past efforts to reform large public systems had not proven effective, so we decided to develop and test a new grantmaking approach," Bailin says. Five years later, these experiments had evolved into a comprehensive new strategy. Today, the foundation is winding down its existing program areas and is now focused on a single field—youth development. Likewise, foundation support is structured to achieve a specific, modest goal: to help a handful of outstanding youth development organizations provide high-quality, effective services to the greatest number of young people possible. But more important, a significant percentage of the foundation's multi-year grants support business planning, administrative architecture, data collection and monitoring, outcomes, and building evaluation systems. "In short, we're investing in outcomes," Bailin says. "The Foundation is not successful unless its grantees are successful in getting better services to more young people. That's our bottom line."

This focus on outcomes is rooted in a fundamentally new theory of change at the foundation. This theory assumes that, given the foundation's relatively modest resources, the best way to improve the lives of young people from socially and economically disadvantaged backgrounds is to identify outstanding youth development organizations that already have evidence of success in serving young people. From this group the Foundation will select those that are performing the best and that are also committed to achieving educational, job readiness and other important skills and accomplishments that will help young people make the transition to independent adulthood. To help these organizations deepen the quality of their program offerings and increase the number of young people to whom they deliver those services, the foundation would provide financial and technical support for a comprehensive business planning process. After the plans were completed, the foundation would structure subsequent support around large, multi-year capacity-building grants that were pegged to clear benchmarks outlined in the business plans.

By doing all these things—identifying, selecting, and supporting high-performing grantees—young people being served by them could be expected to realize improvements in at least one of four areas:

- improved educational skills achievement and attainment;
- a successful transition to employment and self-sufficiency;
- rewarding civic engagement; and
- success in avoiding high-risk behaviors like early pregnancy, substance abuse, and crime.

It was a relatively straightforward value proposition: that if the foundation could find, select, and make capacity-building grants in the right organizations, and provide them with the right kind of support, grantees will help their clients achieve improvements in one of the four outcome areas. But it also signaled a more fundamental change: for the first time in its history, the foundation was explicitly linking its mission and operating strategy to the success of its grantees. This meant that the foundation now had two chief responsibilities. It had to find organizations that constituted the “best bets” in the youth development field, and it had to structure its support for those organizations so they had the best possible chance for success.

### **Internal Alignment: Creating Evaluation-Based Systems, Tools, and Processes**

It is difficult to overstate how profoundly this theory of change and its underlying value proposition altered the foundation’s operations. It’s one thing to say that the foundation will be judged by grantees’ ability to execute their business plans and to demonstrate the effectiveness of their youth services programs; it’s quite another to prove that success. This new focus required the foundation to create new processes and evaluation tools that would help it make “good bets” on potential grantees. It also needed to develop new strategies, methods, and tools to help grantees move from being “good bets” to having programs and services whose effectiveness could be demonstrated with rigorous outcome data.

From the outset, the foundation knew it needed to create a new method of finding and evaluating potential grantees. In 1999, the foundation hired the Bridgespan Group (then known as the Bridge Group), a nonprofit consulting firm spun off from Bain and Company, a leading strategy consulting firm, to help develop a new approach. Drawing on research and analytical methods of equity investors in the for-profit sector, the foundation began developing methods to find potential grantees (what it calls “sourcing”) and then assessing them through a customized due diligence process it created. “Due diligence provides far more information about a grantee than what a funder typically learns from information supplied in response to a request for proposal,” says Nancy Roob, vice president, institution and field building.

### **EMCF Due Diligence Categories**

**Compelling Product or Service:** Does the organization show plausible or demonstrated effectiveness in making a difference in the lives of young people served?

**Leadership and Management:** Does the organization have a track record of achieving its objectives and serving its mission, with a vision for future growth, depth in senior management, and a recognition of the need to fill gaps?

**Operational Viability:** Do the organization’s structure, processes, systems, and relationships have the potential to support growth?

**Financial Health:** Are the organization’s finances in order and does it show a capacity to manage its financial affairs?

**Outcomes Measurement:** Does the organization show concrete efforts at measuring its performance and a commitment to developing systems to evaluate and improve its outcomes?

**EMCF Compatibility:** Are the organization’s culture and management suited to a highly engaged relationship with the Foundation?

*See Appendix A, “Due Diligence Categories and Indicators,” for a complete list.*

The foundation developed six categories for evaluating potential grantees: *Compelling Product or Service (program quality)*; *Leadership and Management*; *Operational Viability*; *Financial Health*; *Outcomes Measurement*; and *Compatibility with EMCF*. Yet while private-sector tools had some analytical strengths, many of the indicators investors used in making for-profit investments — return on investment and market share, in particular — didn't translate easily to the nonprofit sector.

“What we were doing in due diligence a sociologist would call qualitative inquiry,” says David E.K. Hunter, Ph.D., the foundation's director of evaluation and knowledge development. *Did an organization have a compelling product or service? Was it in good financial health? Did it have strong leadership? Did the organization have strong operations? Was the organization committed to outcomes measurement and evaluation?* These were all critical questions, Hunter says, but the foundation needed to systematize due diligence, developing clear, objective indicators for each category. Under Hunter's guidance, the planning team drew on the experience gained in a pilot project to test and refine its methods. The team mapped out grantmaking criteria for each category. Working inductively — from specific evidence to general conclusions — they created specific indicators for each of the six due diligence categories. “It was an enormous amount of work, but in the end we have reliable, fairly rigorous indicators for making ‘best bets’ on grantee organizations.”

### **Internal Restructuring: Creating Evaluation-Based Systems**

As the foundation's due diligence process evolved over the course of the first few grants, it became clear that program evaluation and outcomes measurement had to be integrated from the earliest stages of the grantmaking process. “From our early grants, it became clear — and this is a major lesson learned — that contrary to what we assumed, you could not immediately bring in Bridgespan alone to do business planning right after due diligence,” Hunter says. “That's because something was missing with each of these grantees, and that was a theory of change.” To accommodate this requirement, the Foundation had to restructure its own internal systems and processes to enable evaluation staff to work with grantees on their theory of change as part of business planning.

Over the ensuing months, the foundation undertook a radical restructuring of its operations. The first and most important step was creating multi-disciplinary due diligence teams, which essentially broke down the barriers between the program, finance, and evaluation departments within the foundation. Led by a portfolio manager, due diligence teams include staff from the finance and evaluation units, each of whom are responsible for conducting due diligence in the germane category. Team members conduct document reviews, complete “red-flag” analyses, and go on site visits. Like other foundations, the Edna McConnell Clark Foundation retains distinctions between program and evaluation activities. Nonetheless, Hunter says, the foundation has made it a priority to find areas where both departments can work together.

The effort to share certain responsibilities between the two departments has helped the foundation weave in evaluation and assessment thinking throughout its operations. As an example, evaluation staff have recently assumed responsibility for making initial “pre-screening” assessments of grantees that the foundation's portfolio managers identify. Evaluation staff determine whether these organizations have “compelling” products or

services. Only those organizations that pass this review are referred to the portfolio teams for a full, multi-disciplinary due diligence assessment. Members of the program and evaluation staff also share responsibility for tracking and evaluating grantee data at each step of the grantmaking process. When conducting due diligence, for example, the foundation collects and analyzes a discrete set of documents from each potential grantee—financial records, program documents, internal evaluations, and annual reports, among others. Members of the due diligence team cull these documents, and discuss their findings before settling on a rating for each due diligence indicator.

As foundation staff record their findings from due diligence, as business plans are developed, and as performance milestones are set and tracked, electronic copies of these documents are sent to the evaluation department, which inserts them into the “Document of Record Archive.” Then, using Atlas.ti, a qualitative data-management software program, the evaluation department codes each document, mining data for corresponding relevance to the due diligence indicators, grantee performance against business plan milestones, and the emergence of unanticipated challenges, successes, and other critical issues.

As the number of the foundation’s grants grows, the ability to track data through the stages of due diligence, business planning, and subsequent organizational development will become an increasingly important strategic management tool. “What the system allows us to do is code all the documents of record that come out of due diligence and business planning, and then to match them against subsequent performance tracking,” Hunter explains. “Among other things, we can then analyze backwards from how the organizations are doing to see ultimately which indicators seem to have been powerful in predicting success.”

### **Grantee Relationships: Theory of Change, Setting Goals, Evaluating Success**

Over the next few years, the foundation’s internal databases and analytical systems will allow staff to conduct studies across grantee organizations. For instance, they will be able to look for common issues faced by grantees as they grow, specify the nature of youth outcomes that grantees are achieving, quantify growth in the numbers of young people being served by high-quality programs, and identify unexpected challenges faced by grantees. The long-term goal, Hunter says, is to allow both the foundation and its grantees to use outcome data and evaluation-based systems for program management, and quality improvement. But that level of sophistication is a few years away. For now, Hunter says, the foundation is focused on using the business planning process to help grantees design, implement, and improve their internal evaluation and outcomes measurement systems. “Our objective is to take grantee organizations through business planning, and help them develop a plan to move their youth services from apparent effectiveness to demonstrated effectiveness<sup>1</sup>, and to demonstrate that they’ve done so through the use of rigorous evaluation data they collect routinely as part of their day-to-day operations.”

During the course of conducting due diligence at various nonprofits, the foundation discovered that, even among the highest-performing and most developed youth development organizations, some did not have a clear, verifiable theory of change or an evaluation system to gather and analyze the data to determine what, if any, outcomes, are being achieved. That was understandable, given the funding constraints under which most nonprofits labor

---

<sup>1</sup> See Appendix B, “Continuum of Knowledge About Youth Development Program Effectiveness.”

(many funders are loath to give money to support the development of internal systems or other organizational capacity, preferring instead to fund programs), but it was nevertheless an obstacle to grantees' success. If organizations were going to be successful, they needed robust theories of change that fulfilled four criteria:

- Meaningful—If the grantee accomplished what it set out to do, would it make a significant contribution to the lives of young people?
- Plausible—If the organization does what it hopes to do, is it likely to get the results it expects?
- Doable—Can the organization actually do what it plans?
- Measurable—Can the organization gauge its success?

As grantees entered the business planning process, most had pieces of all four criteria, but few had linked them together systematically. Drawing upon the foundation's own tools, methods, and internal expertise, the Office of Evaluation and Knowledge Development carved out a central role in business planning, helping grantees clarify their theories of change, and developing clear performance indicators and measures to help them assess key outcomes—both for themselves and for the young people they serve.

The role of evaluation has evolved through the foundation's grantmaking experiences, Hunter says. "At first, the evaluation department was deeply intertwined with due diligence. But now we have become integrated with business planning, during which we do theory of change work with each grantee alongside the portfolio manager and in the presence of Bridgespan." Once the evaluation team has helped a grantee establish a clear, rigorous theory of change, the business plan is essentially a way to execute that theory. "Then the Bridgespan people can do what they do best. They go in and start crunching the numbers and analyzing the organization's ability to execute its theory of change, and figuring out what it would have to plan for—organizationally, financially, technologically—to accomplish, with measurable milestones, what it sets out to do."

In the early stages of business planning, grantees typically shore up the first three aspects in their theory of change, reaching clarity on what is meaningful, plausible, and doable given their opportunities and constraints. But as long as grantees lacked even basic evaluation systems for measuring and reporting on their outcomes—not to mention having the capacity to use data and evaluation as strategic management tools—the effectiveness of both grantees' and the foundation's work would be undermined. "The idea that we are working to achieve youth outcomes in a delimited universe of outcome areas was a critical part of our work," Hunter explains. "But then came the question 'how the heck would we know and how would our grantees know' if they are achieving the targeted youth outcomes in those areas?"

Nearly every grantee has made strengthening program quality and building evaluation systems a priority in their business plans, and technical assistance provided by the evaluation staff has proven to be very useful, notes Hunter. "Our grantees understand that measuring and tracking outcomes is necessary for quality management and organizational effectiveness. It's yet another thing to demonstrate that your organization is engendering those outcomes." Over time, the foundation hopes its grantees will be able to demonstrate how their programs have moved from apparent to demonstrated effectiveness.

To guide this work, the evaluation unit has developed a continuum of knowledge about program effectiveness (and has also delineated the differences between different types of evaluation design and the claims each allows organizations to make about program outcomes). On the lowest end of the design continuum, grantees collect little or no demographic data about clients, patterns of participation, or client outcomes. At the high end, organizations collect participation and outcome data using rigorous methods, and they use experimental designs, and an outside evaluator conducts evaluations. In such cases, the difference in outcomes between control groups and program participants can be confirmed with a 95 percent degree of confidence. (The chart in Appendix B outlines the evaluation criteria an organization must meet to demonstrate that its programs have moved from “apparent effectiveness” to “proven effectiveness.”)

Hunter says that it is “essential to recognize that this continuum represents a sequence of evaluation system implementation. At its most elementary, an evaluation system tracks demographic information for each of its program participants and in the aggregate — socioeconomic status, age, gender, and ethnicity. The next implementation stage consists of tracking program participation patterns. In most cases, the foundation provides grantees with basic technical assistance. In addition to the technical assistance that the foundation provides for developing their hardware and software systems, it also helps them with the more challenging task of building a culture of evaluation within their organizations. As the culture changes and evaluation systems develop, organizations begin using more sophisticated methods to test and manage program quality, and, in the last implementation phase, client outcomes — using “pre-testing” and “post-testing” measurements, matching client data against comparison groups, and so forth. Yet, even in the earliest stages, performance and rudimentary outcome data have helped grantees be more effective in quality improvement, business planning, and resource allocation.

## Conclusion

---

### Outcomes-Based Organizations — The Foundation and Grantees Together

For both the Edna McConnell Clark Foundation and its grantees, the new focus on outcomes has been exceptionally difficult and profoundly rewarding. In many respects, the foundation's approach is still evolving, as it should be. And it is precisely the fact that the foundation has integrated evaluation into its work at every stage and level that fosters this evolution. "It's just stunning, the degree to which we've been able to weave evaluation and assessment thinking into our grantmaking in a way that we weren't sure we could. And it has been enormously beneficial to both us and our grantees," Hunter says. Strong evaluation systems allow the foundation to tinker and experiment with new approaches because it is able to evaluate and learn from its successes and its missteps.

Externally, most grantees are well on their way to both improving their programs and going to scale, which is precisely where the foundation hoped they would be. The Foundation expects it will take grantees one to two years to implement each phase of the evaluation systems—the hardware and software systems, the internal processes, the data harvesting techniques. Yet, as challenging as it is to create these "hard" systems, it is often much more difficult to change the accompanying "soft" systems—the underlying institutional culture and human resources. And in this respect, the foundation is asking grantees to undertake changes similar to those the foundation itself undertook when it developed its new strategy. These changes, Hunter notes, "inevitably lead organizations to reexamine themselves and push them toward being more effective and reflective. I know it has in the foundation's case, and, most important, our grantees see it that way as well."

Despite the successes shared by the foundation and its grantees, the experience gleaned in the past few years has led the foundation to be much more modest in its ambitions. "When we started this work, we aspired to make a significant impact on the field of youth development," says Michael Bailin. "We've changed our thinking a bit, realizing that, for the foundation, our 'field' for the moment is really defined by our grantees and their efforts. If we can support them and get them to expand proven youth services at significant scale—and, equally important, to demonstrate how they did it—then we will have made a solid, if modest, contribution to youth development in general."

## **Appendix A: Due Diligence Categories with Indicators**

During due diligence, the Edna McConnell Clark Foundation uses a set of indicators to describe and analyze a potential grantee's capacity in each of six categories:

- Compelling Product
- Leadership/Management
- Financial Health
- Operational Viability
- Performance Tracking And Outcomes Evaluation
- Compatibility With The foundation

### **I. COMPELLING PRODUCT**

*The organization shows apparent or demonstrated effectiveness in making a difference in the lives of young people.*

1. The organization serves young people ages 9-24 from low-income circumstances. Its staff understands the organization's market niche, has a strong rationale for the youth programming it offers, and is highly committed to working effectively to achieve its youth outcomes objectives.
2. The organization has youth outcome objectives that are clearly articulated and can be measured. These include (but may not be limited to): improved educational skills, achievement, and attainment; preparation for the world of work, and a successful transition to employment and self-sufficiency; meaningful citizen participation; and success in avoiding high-risk behaviors.
  - 2.a. Key stakeholders embrace the organization's youth outcomes objectives.
3. The organization's program models are clear, consistent, and well-designed (if possible, research-validated) to promote the achievement of its youth outcomes objectives.
  - 3.a. Programming is delivered consistently for a well-defined target population.
  - 3.b. There is a notable body of anecdotal evidence and at least some outcome data that indicate service recipients are achieving key outcome objectives.

### **II. LEADERSHIP/MANAGEMENT**

*The organization has a track record of achieving its objectives and serving its mission with a vision for future growth, depth in senior management, and, if there are key personnel gaps, a recognition of this fact and a commitment to filling them.*

1. The leadership team demonstrates a high level of commitment to the organization's mission, goals, and objectives.
2. The executive director's and leadership team's professional education, experience, capacity to manage, and tenure qualify them to function well in the positions they occupy.
3. The people who report to the executive director experience him/her to be an effective manager.
4. There is evidence of the leadership team's ability and capacity to manage growth.
5. The organization enjoys a good reputation among local stakeholders, other funders, etc.

6. The executive director demonstrates a substantial understanding of the market within which his/her organization is working.
7. The executive director shows a solid understanding of what it takes to improve the quality and capacity of core operations.
8. The executive director is able to forge critical external alliances.
9. The executive director is able to retain management staff at levels sufficient for the organization to meet its goals and performance objectives.
10. The executive director provides staff with key information, and allows staff input into decision-making.
11. The executive director has a vision for how the organization should grow to provide high-quality services to more young people.
12. The executive director has compelling ideas for how s/he would use EMCF support.
13. Board oversight is appropriate to enabling the organization to meet its goals and objectives.
14. The range of experience and expertise on the board is appropriate to the requirement of assuring that the organization can and does meet its goals and objectives.

### **III. FINANCIAL HEALTH**

*The organization's finances are in order, and it shows a capacity to manage its financial affairs.*

1. In the last five years, overall realized revenue trends are sufficiently robust to indicate that the organization will have the ability to continue to meet its revenue projectors – particularly where these call for increased resources.
2. Overall expense trends are in line with revenue trends and do not present the organization with unmanageable exposure.
3. The mix of revenues is balanced among public and private sources; or, where public sources provide a very large portion of revenues, there are multiple entities (federal, state, local) providing them – so that the organization is likely to be in a position to compensate for unexpected funding reversals by a given agency and to continue to operate without compromising its service mission.
4. Foreseeable shifts in revenue streams do not put the organization at financial risk.
5. The mix of expenses is appropriate to reasonable operational standards.
6. Operating margins (and therefore trends in Net Assets) are adequate to support the organization's financial viability.
7. Asset details and trends are appropriate to support current and planned operations.
8. Liability details and trends indicate adequate debt management.
9. 501(c)(3) IRS-approved nonprofit organization.

10. The results and timing of audits support assessment that the organization is capable of achieving its goals and objectives.
11. Internal controls are sufficient to support the organization in meeting its goals and objectives.
12. Knowledge, skills, and number of financial staff and the organization's financial system are adequate to support the scope of work.
13. The quality of relationships with banks and other financial institutions assures the timely availability of loans and other financial supports necessary to enable the organization to achieve its goals and objectives.
14. Financial reporting to Board and management staff is sufficient in scope, provided on a regular basis, and used to inform decisions.
15. The Board's structure is set up adequately to monitor the organization's finances.
16. Major funding relationships are reliable and adequate to support the organization in meeting its goals and objectives.
17. Other sources of funding (beyond a potential EMCF capacity-building grant) are reliable and adequate to help the organization meet its goals and objectives.
18. Finance & Administration is aware of, and ready for, the organization's future goals and objectives.

#### **IV. OPERATIONAL VIABILITY**

*The organization's structure, processes, systems, and relationships have the potential to support growth.*

1. Staff turnover is at a level that does not significantly impair operational effectiveness or viability.
2. The delegation of responsibilities assigns a rational and reasonable span of control to key managers.
3. The organization aligns staff training and professional development with organizational goals and objectives.
4. Frontline staff skill sets are appropriate to the tasks required of them.
5. Human Resource practices are adequate to support the organization in meeting its goals and objectives.
6. Marketing and public relations are adequate to support agency goals, including recruitment, fundraising, and community visibility.
7. The organization understands the necessity of developing IT systems that support operations in the following domains: collection, processing, and reporting of financial and service delivery data; internal operations; and communications.
8. The organization's structure, staffing, and priorities for resource allocation are designed and aligned to achieve its goals and objectives (especially the youth outcomes desired for service recipients).

9. Space is not an unreasonable barrier to service delivery, staff retention, and recruitment.
10. The organization has a realistic and responsible approach to risk management. (This includes an assessment of its legal exposure and insurance coverage).

## **V. PERFORMANCE TRACKING AND OUTCOMES EVALUATION**

*The Organization shows concrete efforts at measuring its performance, and a commitment to developing systems to evaluate and improve its outcomes.*

1. The organization collects data on client demographics, or there is written evidence (plans) demonstrating that it recognizes that it is essential to do so.
2. The organization collects data on client participation/attendance, or there is written evidence (plans) demonstrating that it recognizes that it is essential to do so.
3. The organization has concrete plans to use data to manage program quality.
4. The organization currently is committing adequate resources to performance tracking and client outcomes assessment, or there is written evidence (plans) demonstrating that it recognizes the need to do so.
5. The organization's leaders understand what it takes to *demonstrate* that program participants are achieving the specified outcome objectives, and there is written evidence (plans) demonstrating that they are committed to doing so within the next 2-5 years.
6. Leadership and management are able to articulate indicators of success for the organization as a whole and for each of its programs and/or services, and also have written plans for assessing each indicator using clear measures.

## **VI. EMCF COMPATIBILITY**

*The organization's management and culture are suited to a highly engaged relationship with the Foundation.*

1. EMCF and the organization are able to achieve consensus on a vision regarding the nature of the proposed relationship.
2. The organization's culture and leadership seem to be open and its internal operations accessible.
3. The organization worked and communicated with the EMCF team during due diligence in ways that conveyed the likelihood of a sustained stance of cooperation and collaboration over the course of the proposed investment period.
4. The organization has a track record of providing funders with honest and timely information ("they deal straight").
5. Gut check – it "feels right" to be working with this organization.
6. The organization's program(s) is (are) connected to EMCF strategy in the city (out-of-school-time programs).

# Appendix B

## Edna McConnell Clark Foundation Guide to Assessing Youth Development Program Quality and Effectiveness

### Introduction

The Edna McConnell Clark Foundation is working to improve the lives of young people in lower income communities through investing in the core capacities of nonprofit youth serving organizations. We believe that by doing so the Foundation's grants will help these organizations to deliver their programs and services to significantly more young people more reliably and at higher levels of *quality* and *effectiveness*, thereby promoting notably better lives and prospects for large numbers of low income youths. It follows, then, that these variables – combined under the concept of “compelling product” – are the first areas that Foundation staff assess when considering whether to make an investment in a potential grantee organization.

### Indicators of Youth Program Quality

While many different kinds of programs can be effective in promoting desirable youth outcomes, the relevant research documents that certain traits characterize all successful youth development programs – regardless of their type or approach. These indicators of *program quality* include:

1. Physical and psychological safety – including practice that increases safe...and decreases unsafe or confrontational peer interactions;
2. Appropriate structure – clear and consistent rules with predictable and firm enough control;
3. Supportive relationships – warmth, caring, good communication, connectedness, support and guidance, and responsiveness;
4. Ongoing opportunities to belong – meaningful social inclusion regardless of ethnicity, gender, sexual orientation, or disabilities;
5. Positive social norms – rules of behavior, expectations, and obligations for service;
6. Support for efficacy and Mattering – youth-based empowering practices that support autonomy, making a real difference in one's community, and being taken seriously;
7. Opportunities for skill building – predictable occasions and intentional learning experiences to acquire learn physical, intellectual, psychological, emotional and social skills;

8. Integration of family, school, and community efforts – concordance and coordination among activities and values in these diverse venues.<sup>1</sup> Related to these indicators, there is research evidence<sup>2</sup> for the following indicators:
9. Adults interact with participating young people in predictable, stable, and caring ways, and hence care and effort is put into matching adults with youths;
10. Effort is devoted to involving participants’ nuclear or extended family – or other significant adults – in the program;
11. Care is taken in the recruitment, screening, training, and supervising of staff and volunteers, using well-developed guidelines;
12. Where programs offer standardized services, these are tailored to individual participants’ needs, which are assessed systematically; also, one-on-one adult-youth relationships are valued, available, and supported organizationally;
13. Young people participate in the program(s) for a minimum of one year, and often longer; and finally,
14. Young people have a valued voice in, and can make choices about how they participate in the program.<sup>3</sup>

*Therefore, when assessing a potential grantee organization, Foundation staff will look at its programs to assess how well they implement these desirable program traits, and will invest only in those organizations where are present to a significant degree.*

### **A Hierarchy of Attribution Regarding Program Effectiveness**

Once it has been established that a youth program manifests the quality traits listed above, it remains necessary to assess *what is known about the program’s effectiveness* – that is, how likely it is that young people will benefit from participating in. Are young people who are being tutored in literacy likely to improve their reading and writing skills? Are young people being mentored likely to finish high school? Are graduates of an employment program likely to get viable jobs?

There are very few youth development programs whose effectiveness has been proven through rigorous, experimental evaluation. Therefore, the Foundation must rely on other sources of information to assess the likely effectiveness of youth programs, so that it can invest in organizations that probably do achieve the results that they and the young people whom they serve desire. Toward this end we have adopted a hierarchy (Tables I. – III. pp. 3-5) as a guide in order to increase the prospect of finding and investing in effective youth development service organizations.

Of course, what is known about the effectiveness of youth is a continuum – varying from programs for which there is no evidence at all that they are helping young people, to programs whose effectiveness has been proven. In order to be confident that the Foundation’s youth development grants really are helping more young people do better in life – and to communicate clearly with potential and current grantees about these matters – we divided this continuum into discrete levels, each of which, of course, represents its own range.

---

<sup>1</sup> Taken directly from Table ES-1: Features of Positive Developmental Settings, in J. Eccles & J. A. Gootman (eds.), *Community Programs to Promote youth Development*, Washington, DC: National Academy Press, 2002 (pp. 9-10); these indicator are research-based.

<sup>2</sup> E. C. Hair, K. A. Moore, D. Hunter, & J. W. Kaye (eds.), *The Edna McConnell Clark Youth Outcomes Compendium*, Washington, DC: Child Trends, 2001

<sup>3</sup> Although not research-based, indicator 14. is held to be important by many youth development practitioners.

*At the first level are programs where so little useful information about their effectiveness is available that the organizations operating them cannot currently be considered as potential EMCF youth development grantees.. Beyond this line we identify three levels of attributed of program effectiveness:*

1. Apparent effectiveness, the first level, for which it is justifiable to *assume* that young people are benefiting as intended from participating;
2. Demonstrated effectiveness, the intermediate level, for which one can make a *substantiated judgement* young people are benefiting as intended from participating; and
3. Proven effectiveness, the highest level, where the *program's impact* on participants has been *confirmed scientifically through experimental research*.

*For the Foundation to invest in a youth development organization, its program(s) must, at the very minimum, fit toward the upper end of the level of “apparent effectiveness.”*

On page 4, Table I. defines the category of “apparent effectiveness,” specifies the kinds of information that must be available to identify such programs, and indicates the minimal kinds of evaluation activities that have to be in place to provide such information. But even within this category we find a range marked by differences in how and what kinds of information organizations collect about the young people who participate in their program(s), and thus what they can claim about their program(s) effectiveness. At the lower end are those that have not yet been systematic in their collection of basic information about program participants, patterns of participation, and outcomes for service recipients. At the higher end are organizations that have implemented full data collection systems in all three areas, and are using information collected in this way to plan and manage program quality and service delivery.

In our experience, most youth development service providers offer programs that fall somewhere in the range of “apparent effectiveness.” *Thus, a key consideration for deciding to invest in an organization is the degree to which its leadership is committed to improving data collection about its programs' effectiveness.* For some programs, at the lower end of the range in this category, this will mean, raising them within this category to the highest level over the first 1-2 years of the grant (see Table I.). *However, over the course of the whole 3-6 year grant, we expect an organization to achieve the next level of our hierarchy for its programs, that of “demonstrated effectiveness”* (see Table II. page 5).

However, there are some organizations that already have demonstrated one or more of their programs' effectiveness before any relationship with the Foundation exists. In that case, if the organization offers more than one youth program, the Foundation will encourage it to demonstrate the other programs' effectiveness as well, preferably using a quasi-experimental evaluation design. And finally, for programs whose effectiveness has been demonstrated at the top of this intermediate range (see Table II.), under certain conditions the time may be ripe for proving their effectiveness through the use of an experimental evaluation design (see Table III. page 6).

<b>Table I. Lowest Level of Attribution for Program Effectiveness</b>	<b>Definition</b>	<b>Quality of Knowledge</b>	<b>Level of Evaluation System in Place</b>
<b>“Apparent Effectiveness”</b>	Through the more or less systematic collection of various kinds of information about program participants and youth outcomes -- that range from the purely impressionistic and anecdotal to the highly systematic -- the assumption that one can reach an initial assessment claiming a program’s likely effectiveness is deemed justifiable.	<u>Justified assumption.</u> Thoughtfully collected information is tested against personal and professional experience, as well as available knowledge about similar programs, to substantiate the assessment that a program’s effectiveness is “apparent.” This requires an assessment that the program in question is being implemented with a high degree of fidelity to its design and intentions, and that its attributes are indeed comparable to those similar programs to which it is being compared.	<p><i>a. Low End.</i></p> <ul style="list-style-type: none"> <li>▪ Little or no systematic data collected on program participant demographics and patterns of participation.</li> <li>▪ Data are not collected systematically on participant outcomes.</li> <li>▪ Data are collected internally, without the involvement of an external evaluator.</li> </ul> <p><i>b. High End.</i></p> <ul style="list-style-type: none"> <li>▪ Systematic data collection on program participant demographics and participation patters.</li> <li>▪ “Pre-test” and “post-test” participant outcomes data are collected – but without the use of any comparison groups.</li> <li>▪ Data are collected internally, but an outside evaluator may be used for the analysis of client participation and outcome data.</li> </ul>

<b>Table II. Intermediate Level of Claim for Youth Development Program Effectiveness</b>	<b>Definition</b>	<b>Quality of Knowledge</b>	<b>Level of Evaluation System in Place</b>
<p><b>“Demonstrated Effectiveness” (or “Demonstrated Results”)</b></p>	<p>Through the tracking of outcomes for service recipients and the use of comparison methods, the program’s likely effects on service recipients have been identified – but without any way of calculating the degree to which the program has contributed to these outcomes.</p>	<p><u>Substantiated judgement.</u> Rigorously collected data are tested against comparison data as well as analytical methods of, and constraints on, generalization.</p>	<p><i>a. Low End.</i></p> <ul style="list-style-type: none"> <li>▪ Youth outcomes data are collected “pre-test” and “post-test” where this is applicable.</li> <li>▪ An evaluation model is implemented that uses some comparison group(s); or, comparison is made to outcomes benchmarked against the general population.<sup>4</sup></li> <li>▪ Comparative outcome evaluations generally are conducted by an external evaluator, but may be conducted by an internal evaluator if the expertise and level of resources are available internally.<sup>5</sup></li> </ul> <p><i>b. High End.</i></p> <ul style="list-style-type: none"> <li>▪ Youth outcomes data are collected “pre-test” and “post-test” where this is applicable.</li> <li>▪ An evaluation with a rigorously matched comparison group.</li> <li>▪ Comparative outcome evaluations generally are conducted by an external evaluator, but may be conducted by an internal evaluator if the expertise and level of resources are available</li> </ul>

<sup>4</sup> Evaluators call this kind of approach a “quasi-experimental” evaluation design.

<sup>5</sup> We believe such internal capacities are very rare and for practical purposes this means that at the level of “demonstrated” quality or results, grantees will need to implement external or third party evaluations.

<b>Table III. Highest Level of Claim for Youth Development Program Quality</b>	<b>Definition</b>	<b>Quality of Knowledge</b>	<b>Level of Evaluation System in Place</b>
<b>“Proven Effectiveness” (or “Impact”)</b>	Through the use of a random assignment experimental research design, the impact of the program on service recipients has been verified in a statistically significant manner.	<u>Scientific knowledge.</u> Rigorously collected data are tested against the full body of scientific knowledge (including statistical and analytical methods of, and constraints on, generalization).	<p>internally.</p> <p><i>a. Low End.</i></p> <ul style="list-style-type: none"> <li>▪ Youth outcomes are collected using a rigorous experimental evaluation method – which at a minimum means use of “pre-test” and “post-test” measurement where applicable, some form of random assignment (to control for selection bias), and the use of a rigorously designed control group (to permit reliable estimates of attribution).</li> <li>▪ An external evaluator conducts the evaluations.</li> <li>▪ The difference in outcomes between program and control group is confirmed statistically with a <u>90%</u> degree of certainty.</li> </ul> <p><i>b. High End.</i></p> <ul style="list-style-type: none"> <li>▪ Youth outcomes are collected using a rigorous experimental evaluation method (see above).</li> <li>▪ An external evaluator conducts the evaluations.</li> <li>▪ The difference in outcomes between program and control group is confirmed statistically with a <u>95%</u> degree of certainty.</li> </ul>

These considerations apply to organizations that have reached a mid-level stage of development, defined by operational stability and reasonable financial security. Such organizations can operate their programs with sufficient reliability that one can be reasonably confident that in assessing the quality and effectiveness of their programs what one finds is likely to be sustained over time. Thus, the Foundation can have a relatively good idea of the kinds of benefits young people are likely to derive from participating in these programs.

However, the Foundation may also find certain organizations that are at an earlier stage of development – but are offering a program or service that is so creative or novel, or seems to offer such great a potential to benefit participants, that it seems like a “good bet” to invest in the organization in order to help it develop its core capacities and achieve the next stage of its development. *For such organizations the Foundation will set a higher bar for assessing “compelling product” – i.e., program quality and effectiveness.* In practice this is likely to mean that the program or service is relatively narrow and highly focused on a limited number of outcomes it is seeking to promote in its participants. In addition, the leadership of the organization would need to be driven to succeed, and have aligned their limited resources in such a way as to maximize the likelihood of success.