

Hoole, E., & Patterson, T.E. (2008). Voices from the field: Evaluation as part of a learning culture. In J. G. Carman & K. A. Fredericks (Eds.), *Nonprofits and evaluation*. *New Directions for Evaluation*, 119, 93–113.

6

Voices From the Field: Evaluation as Part of a Learning Culture

Emily Hoole, Tracy E. Patterson

Abstract

The authors present the argument that evaluation as currently practiced is failing the needs of the social sector, if the goal is to help organizations achieve their mission. Three case studies are presented in which evaluation played a key role in the development of a learning culture within the organization. Each case highlights the positive potential for evaluation to foster continuous organizational learning and improvement, build relationships, and facilitate organizational development. The authors find that although capacity building efforts by funders were important, the commitment of organizational leadership was critical in transforming the role of evaluation from one of basic reporting and accountability to a true

Note: The authors of the chapter would like to thank the members of the three case study organizations for their willingness to share their stories and the valuable contributions they made to this chapter. From BBBS, Rajan Shore and Mary Rodiham; from Big Thought, Jennifer Bransom and Katy Denson, and Dennis Palmer Wolf, senior scholar with the Annenberg Institute for School Reform, Brown University; and from the Healthy Native Communities Fellowship Learning Team, Nina Wallerstein, Kristine Maltrud, Shelley Frazier, Rebecca Rae, Frances Varela, Jeff Stowell, Marita Jones, and Chris Percy.



process of continuous organizational learning. This transformation requires leadership commitment to development of an evaluative learning culture and an infrastructure to support it. © Wiley Periodicals, Inc.

This chapter examines the link between evaluation and an organization's or program's mission, and how the two are connected through the development of a learning culture within groups and organizations. First, we review the state of nonprofits and evaluation and present the argument that evaluation, as currently practiced, is failing the needs of the social sector, if the goal is to help organizations achieve their mission. Second, three case studies are presented, highlighting the positive potential for evaluation to foster continuous organizational learning and improvement, build relationships, and facilitate organizational development. Finally, best practice and lessons learned from all of the cases are presented and discussed.

Nonprofit organizations enjoy a special status in society, both their governmentally recognized status as well as a moral status, based on the good work they are striving to do. Society empowers and invests in these organizations and their missions, through donations, government and foundation funding, and other means. Because the activities of the organization are carried out in service of the mission, the link between evaluation and mission achievement should be clear and strong. Highly effective nonprofit organizations are successful in achieving their mission. To achieve organizational and programmatic effectiveness systematically, evaluation becomes a mission imperative. Recognizing evaluation's critical role in advancing the mission of the organization leads to increased awareness regarding the vital need to develop organizational learning as an organizational capability.

Torres and Preskill (2001) define organizational learning as a process of continuous growth and improvement through use of evaluation, embedded in the organization's culture, systems, structure, and leadership, leading to alignment of values, attitudes, and perceptions. As key stakeholders, institutional funders such as foundations, governmental organizations, and United Ways need to build evaluation capacity to foster organizational learning if their goal is to help nonprofits achieve their mission (Hoole, 2007). Unfortunately, too often funders uphold standards of accountability without enabling nonprofits to develop a learning culture in service of the mission.

A recent article in the *Stanford Social Innovation Review* exposed the dirty laundry of the evaluation field: "Evaluation is failing to help make the social sector more effective" (Snibbe, 2006, p. 40). What led the author to conclude such a state of affairs? Snibbe points to the wasted time and effort spent on creating reports for a multitude of funders, which contribute little to improving program effectiveness. What is worse, Snibbe contends, is these funders do not seem to use the evaluation findings to help organizations improve their programs or make important programmatic changes.

This highlights another failure in the system: funders themselves are not effective at using evaluation data to make important decisions (Bickel, Millett, & Nelson, 2003). So if funding decisions are not based on the data funders demand, evaluation does not contribute to organizational sustainability. Another challenge to the effective use of evaluation in the social sector is the difficulty of demonstrating impact, leading nonprofits to gather useless and sometimes misleading data (Fredericks, Carman, & Birkland, 2002; Katz, 2005; United Way of America, 2000).

If evaluation is not tied to the organization's mission and does not contribute to the fundamental understanding of the people and issues the organization serves but is instead undertaken from accountability pressures or conflicting stakeholder demands, evaluation fails the social sector. This is why it is crucial for evaluation within an organization to support the development of a learning culture. When a learning culture develops and organizational learning occurs, the opportunity for increased effectiveness abounds: "Organizations realizing this . . . goal are in a much better position to grow and thrive in the 'dynamic, unstable, unpredictable environments' that characterize today's world" (Bickel, Millett, & Nelson, 2003, citing Preskill & Torres, 1999, p. xix). This is how evaluation can help the social sector achieve the noble and essential goals of improving the human condition.

Carman's study of nonprofit evaluation practices (2007) offers empirical data regarding the role of evaluation within the nonprofit field. Results indicate only 65% of the organizations in the study engage in formal program evaluation, whereas 95% report activities to their boards and 90% experience site visits from funders or regulatory agencies. It is encouraging that 87% and 85%, respectively, report conducting firsthand observations of program activities and monitoring program implementation, but there is little indication that the myriad activities nonprofits are engaged in, in the name of evaluation, are contributing to the development of a learning culture or leading to program improvement. Out of 20 possible evaluation and accountability activities, the average number undertaken by organizations was 13. Carman (2007) concludes that most of these community-based organizations have yet to truly invest in meaningful evaluation; they are just collecting input and output data.

This current state of affairs has developed over time owing to accountability pressures on the funders themselves. The federal government instituted performance measurement for all agencies in 1999; the same year, Porter and Kramer (1999) argued radical changes in the field of philanthropy were needed for the field to achieve its mission. These changes include more focused grantmaking; long-term, close grantee relationships; and greater attention to the results of grantmaking efforts (Porter & Kramer, 1999). High-profile cases of misconduct in the social and corporate sectors, starting with United Way of America in 1992, New Era Foundation in 1997, and continuing today with such organizations as EduCap, America's Clean

Water Foundation, Arthur Andersen, and Enron have led to greater congressional interest, resulting in passage of Sarbanes-Oxley and heightened congressional oversight (Snyder, 2007). Fredericks, Carman, and Birkland (2002) point out the evaluation difficulties arising from the complex funding environment nonprofits find themselves in:

- Conflicting stakeholder demands on programs
- Conflicting stakeholder demands for reporting and evaluation
- Ever-present challenge of developing meaningful outcome measures
- Significant organizational challenges in data collection, management, and analysis

How can nonprofit organizations navigate the ever-changing environment in which they must not just survive but thrive, in order to advance their mission? In this chapter, we profile three organizations that have developed their evaluation capacity, created a learning culture, and continuously focus on organizational learning for the purpose of improving the work they do. These organizations fit the description of Type I organizations (see Chapter 5 in this issue) in demonstrating a long-term commitment to evaluation, use of internally driven comprehensive approaches to evaluation focused on the programs and outcomes, dedication of staff to evaluation, and placing clear priority on evaluation for program improvement. These organizations make evaluation an honest and ongoing discussion rather than a bureaucratic process (Maccoby, 2003). The learning cultures within these organizations share similar characteristics. Learning is driven by leadership and management who are willing to share their own mistakes, reward good ideas, encourage staff to hold each other responsible, and lend support to one another in their learning endeavors (Botcheva, White, & Huffman, 2002; Maccoby, 2003; York, 2003).

Voices from the Field

The three organizations we profile represent a range of organization types, sizes, and purposes. Big Brothers Big Sisters of Harrisonburg-Rockingham is a local affiliate mentoring-based youth-serving organization in a small city in Virginia. Dallas ArtsPartners is a major initiative of Big Thought, a partnership-oriented arts education nonprofit in Dallas. The Healthy Communities Native Fellows program is a national community-based leadership program focused on improving the health of Native American communities. These groups share similarities in their approach to evaluation as a means of learning and a facilitator of growth.

Big Brothers Big Sisters of Harrisonburg-Rockingham

The mission of Big Brothers Big Sisters (BBBS) is to “make a positive, measurable difference in the lives of children by promoting self-esteem, expanding

awareness of life's opportunities, and providing education, guidance, and support through professionally supported, one-to-one mentoring relationships" (Big Brothers Big Sisters, 2007). The local affiliate serving Harrisonburg and Rockingham County (population approximately 100,000) in Virginia serves more than 830 children in two types of programs: (1) community-based mentoring and (2) site-based mentoring. Site-based mentoring programs are based in schools and community sites. The School Mentoring Program consists of "Bigs" spending an hour a week with their "Littles" at school. Half the time is spent on academic activities, the other half on relationship-building activities. Other site-based programs in the community include a partnership with the local Boys and Girls Club, a program at a local trailer park, and a partnership with a local retirement community. Thanks to grants and targeted funding, several specific subpopulations of children and volunteers have been identified for recruitment. In 2007, the organization had a goal of serving 300 Hispanic youths, 75 children of incarcerated parents, and recruiting older adults for their intergenerational mentoring program. These targeted subpopulations fit within the two types of mentoring programs. The organization experienced tremendous growth during the nine-year tenure of the current executive director, Rajan Shore. A 300% increase in the number of children served, and more than a 300% increase in the budget, from \$130,000 to more than half a million, resulted in many changes to the organization. Recently a middle-management level was added to the organization to meet the growing demands.

With the growth in services and increase in specific grant-funded programs, the list of stakeholders for the organization has expanded. The list includes children and families as the key stakeholders; staff, volunteers, schools, and school systems; a multitude of funders, including the federal government, state government, the United Way, other foundations, the National BBBS organization, and individual and corporate donors; potential volunteers and donors; and the community as a whole. These stakeholder groups bring a variety of needs and concerns to the evaluation process. Funders are typically interested in the impact and results of the program. Funders and donors are also concerned about the cost effectiveness and stability of the organization's programs. Volunteers and parents are interested in their personal satisfaction with the program. The board of directors is interested in the national rankings of the organization.

From 2004 to 2006, the organization was recognized nationally for double-digit growth in revenue and in service to youth. In 2005, a case manager for the Hispanic Mentoring Program was named the national Program Staff of the Year, and in 2006 the agency's executive director was recognized as CEO of the Year for agencies with revenue of less than \$500,000.

Despite the large number of stakeholder groups, stakeholder involvement in the evaluation process is minimal. The Board Program Committee completes an organizational self-assessment each year as part of the requirements for affiliation. Although performance outcome data are collected from

children, parents, volunteers, and teachers annually, it is completely voluntary. For Spanish-speaking families, the materials are translated into Spanish or collected orally in Spanish.

Stakeholder demands for evaluation require the organization to report data on the children they serve and their programs in a variety of formats. According to the executive director, they “slice and dice the data any number of ways” to meet funders’ requirements. Even with the wide variety of reporting requirements, the organization stays true to its mission and will not collect data unrelated to its core mission. A recent example is a health-related funder of a collaborative site-based mentoring program wanting the body mass index (BMI) of the children participating in the program. The organization determined this fell outside of the mission, and it found a work-around with the local school nurse, who collects BMI data during an annual physical at the school.

In many cases, the staff push back when program managers or agency leadership are tempted to consider grants or activities that would take them away from their mission or could compromise the staff’s ability to meet national program standards. In general the organization is much more sophisticated now about grant and funding opportunities, evaluating what the funder is asking for, whether the organization currently has such data, and what type of funding it is (one-time, one-year, or multiyear) in deciding which opportunities to pursue and what it will require them to do. When in grant-funded collaborative partnerships with other organizations, BBBS of Harrisonburg-Rockingham prefers to maintain responsibility for evaluation and reporting. They believe they have the technical systems and capacity to meet funder requirements. They work with the partner organization to get the necessary data, and then furnish a comprehensive report to the funder.

Having received United Way funding for 30 years, the organization has experienced the outcome measurement movement from the beginning and can point to both positive and negative aspects of the United Way’s efforts. They believe outcome measurement helps the United Way meet the needs of its stakeholders, and they value the stamp of approval they gain in the community as a funded program. From BBBS’s perspective, United Way’s focus on outcomes made agencies take a second look at the impact of their services—which is positive for programs and stakeholders. Organizations with strong evaluation capacity had little trouble adapting to the United Way’s outcome measurement efforts. However, they fear the emphasis on outcomes will reduce the willingness of grassroots organizations, which lack the infrastructure needed to support the measurement work, to apply for United Way funding. Grassroots organizations have a better chance of competing for United Way funds when training and support is offered. Further, they report that the year-round, time-intensive nature of the requirements and accompanying process is clearly a negative for agencies. Unless the United Way streamlines the process and considers multiyear funding, they believe agencies will be forced to examine, every year, whether the costs outweigh the benefits of association. A handful of agencies of various sizes and

capacities have left the local United Way recently, finding the costs of association greater than the return. Yet BBBS of Harrisonburg-Rockingham remains dedicated to United Way, its most consistent funder, which currently funds 12% of the organization's budget.

As a local affiliate, BBBS of Harrisonburg-Rockingham benefits from the efforts of the national organization. National clearly values and supports the learning of its affiliates. Through work with Public/Private Ventures, the national office funds rigorous impact evaluations of the community and site-based mentoring programs, something the affiliates would not be able to afford locally. National develops and provides evaluation tools and facilitates learning across the local organizations. As a pilot site of the Hispanic Mentoring Program, the organization worked with an external evaluator and found the experience to be positive. Though they did not receive data as often as they would like (this is a very data-driven organization), they were able to push the evaluator to provide them additional data that they valued. This is an indicator of a true learning culture within the organization, not to be satisfied with the standardized data across the sites but asking for, and receiving, locally relevant data for understanding and program improvement. The program and evaluation have been under way for 3 years now and will run an additional year, but the organization is planning now to take on the data collection and analysis because they value and want the data for their own use. An additional, unexpected benefit of participation as a pilot site was getting in on early implementation of the AIM performance management system. AIM is an on-demand, real-time data collection and reporting system. As a small agency, they would have been unlikely to receive the system so soon, but because of their visibility in the pilot program they were able to advocate for and receive the system.

In the AIM system, staff enter key data aligned with national standards of practice as part of their daily routine. The system is both visual and motivating for staff, allowing goal setting and baseline data collection. Individual-level data are shared between the staff member and his or her supervisor, and aggregate data are shared with the entire staff. To use the system requires an annual fee of roughly \$4,500 from the organization, which they consider a significant but worthy investment. The AIM system allows the organization to reach a higher level of performance management by connecting the process to the outcomes. For the agency, evaluation is not just about results but about how they do their work. As Mary Rodiham, Director of Partnerships and Fund Events, states, with this system they "have learned to care about issues in a whole new way. When you can't see it, you can't value it." As a result, staff output and input have increased. Staff success is celebrated and areas for improvement are targeted. Because of the transparency of the system, they are able to see which factors can be controlled and which can't. This promotes more productive problem-solving sessions, with staff developing action plans they believe will work.

The move to the new system and process was intimidating at first, but it has been transformative for the staff. Management made sure the system

was not viewed as judgmental by framing it as a way for staff to do their work better. Staff are empowered to make an impact on what they can, and to realize what they have no control over. For example, the system allows the staff to monitor reasons for match closure. The staff is able to distinguish between closures that could not be prevented, such as a child moving out of the service area, from those the staff has greater control over, such as volunteer dissatisfaction or “incompatibility.” Whenever they identify these “intra-agency” factors, they can conduct reviews of the agency’s inquiry, enrollment, and interview processes to evaluate whether the correct information is being gathered to allow the best, high-quality match possible. Regular file audits also permit the agency to monitor the quality of services and help benchmark whether service improvement is possible or necessary.

Another key example of the learning culture within the organization and creative problem solving on the part of the staff is the recent discussion among the community-based program match support staff. The community-based match support contact completion numbers are much lower than for site-based programs. National program standards require match support specialists to contact the children and mentors monthly. Site-based programs have an advantage because all the children co-located in a school or community site are assigned to a specific staff member, allowing staff to complete many of their contacts at one time. Community-based matches are individual matches who meet together during the week at different times and locations, so community-based match support specialists have children who are located all across the service area. To leverage the advantage site-based staff enjoy, community-based staff asked for a printout of the schools their children attend and are investigating how they can work as a match support team to complete the contacts rather than working alone. This process improvement would not have been possible without access to the metrics available through the performance management system (AIM) and a learning culture among the staff.

Examples abound of organizational learning made possible by developing evaluation capacity and the value placed by BBBS of Harrisonburg-Rockingham on a learning culture. The organization knows, if it is to be successful in getting a match to last 6 months, the match is much more likely to go on to reach the one-year point (the minimum commitment they ask volunteers to make). One strategy used to help matches meet and exceed the 6-month retention level includes setting the expectation of a minimum 1-year match involvement early and often in the volunteer enrollment and training phase. Although BBBS Standards of Practice require only monthly contact with all match parties, more frequent and meaningful contact with agency staff, early in the match, has been surfacing as a best practice across the nation. Customizing match support to the needs of individual mentors, youths, and families is another strategy to sustain long-term match relationships. For example, if the youth a mentor is paired with has ADHD, then offering training specific to the needs of the child and strategies to positively

interact with the child would also be important. If the child has an incarcerated parent, there may be another set of trainings that would be helpful to set the mentor up for success in working with the child.

They also understand the impact of the program as it grows and changes over time, with longer mentoring relationships leading to more positive impact and increased likelihood children will maximize their potential—the ultimate benchmark for the organization. Research by Grossman and Rhodes (2002) indicates length of relationship is critical for the level of impact on youth. Impact cannot occur if the match does not stay together long enough to develop a solid relationship. BBBS of Harrisonburg-Rockingham understands the complex variables in mentoring relationships and is able to effectively intervene in areas they know they can affect. AIM allows managers to monitor 6-month retention rate, average active match length, average match length of closed matches, and reasons for closure. With this information, staff are then able to find correlations between successes and challenges in maintaining strong match relationships and youth impact. For example, if 30% of matches are closing between 4 to 5 months because the volunteer lacks contact with the agency, they can theorize that they need to offer stronger early match orientation, training, and support in an effort to increase match length.

BBBS of Harrisonburg-Rockingham understands the families they serve and their changing risk factors over time, so the staff is better able to evaluate client needs and develop strategies to meet each child's situation. Clearly the organization has kept the mission in the forefront as it has grown and revised evaluation processes over time.

Executive Director Shore reports that the journey to build the organizational capacity necessary to effectively use evaluation for organizational learning was “scary and painful.” She found and hired staff with the technical and database skills necessary to use data management systems. At times, program staff members were pulled away from their work to do evaluation and reporting. But over time they invested in technology to help with this effort, the most recent example being the AIM system to maximize efficiency. The recent creation of the middle-management team will further enhance their ability to maintain growth and reflect on and use data. Funder expectations for evaluation and grant funding for program expansion are acknowledged as playing a critical role in their growth and taking the organization's performance to a higher level. Shore states she would not go back from where they are now; they are working with more knowledge and are providing better-quality services with greater efficiency. She reports seeing a recent shift in thinking at the board level toward more long-term strategic thinking in revenue generation, staffing, and growth. In the past, during lean funding periods, the typical response was to cut nonprogram staff and expect program staff to work harder and longer with fewer resources. The board now realizes if it is to achieve the goal of becoming a \$1 million dollar organization by 2013, this type of thinking will need to change.

BBBS of Harrisonburg-Rockingham clearly values the role of evaluation in developing a learning culture within the organization. Strong management support for creation and adoption of systems and integrated daily use of data signal to staff the importance of learning and continuous improvement. There is also clear evidence that the combination of a data-rich environment and learning culture fosters creative problem solving among staff focused on program improvement. This emphasis on having the best possible system, methods, instruments, and programs is part of an overall ethic within the organization to reach the highest attainable performance, in order to improve the lives of the children, the families, and the community they serve.

Big Thought and the Dallas ArtsPartners (DAP) Program

This second case study focuses on evaluation of Dallas ArtsPartners, a major initiative of a partnership-oriented arts education nonprofit organization called Big Thought. At the center of this case is the story of how one nonprofit organization with many partners, came to balance the accountability functions of evaluation with equal value on possible learning through evaluation.

Big Thought's mission is to "inspire, empower, and unite children and communities through arts and culture" (Big Thought, 2007). The first Big Thought program that underwent a major evaluation was Dallas ArtsPartners, a community-based collaboration with a mission to "build an effective infrastructure that ensures that children and teachers throughout the city have the chance to learn about and participate in the cultural life of the city" (Wolf, Bransom, & Denson, 2007). Since 1998, Dallas ArtsPartners (DAP) has offered a coordinated program of services for all elementary schools in the Dallas Independent School District, the 12th largest district in the country, by bringing local cultural organizations together. DAP serves 156 schools, 101,000 children, and 6,000 elementary educators. Each year, DAP coordinates 15,000 programs provided by cultural organizations, including in-school performances, field trip performances, artist residencies, master classes, workshops, and guided tours. DAP started as a \$1.5 million project that has grown over time, receiving funding from the school district, other public agencies, foundations, and individual and corporate donations as well as in-kind donations from the cultural community. The school district invests approximately \$1 million for DAP services to children (for example, tickets to the symphony or a resident artist coming to a school), while the city as well as public and private funders support the program's infrastructure—staff, technology, research and evaluation, training for teachers, and so on.

The DAP evaluation was a major social and fiscal investment for the organization. To merit this investment, Big Thought was committed to the study producing not only the usual results but also an organization with the ongoing capacity to reflect critically on its programs. In addition, given the high number of DAP stakeholders (Dallas Office of Cultural Affairs,

the Dallas Independent School District, a cultural community of 50 organizations), the process of conducting the evaluation had to build a climate where stakeholders would feel comfortable discussing and addressing gaps, uneven results, and need for improvement. Big Thought and its partners saw that collective learning had to become a hallmark for the evaluation, running throughout the process from building outcomes to collecting data to understanding results. All stakeholders had to be equal partners for the work to be successful.

In the recently published guidebook *More Than Measuring*, Jennifer Bransom, Big Thought's director of Program Accountability, along with local evaluator Katy Denson and Dennis Palmer Wolf of the Annenberg Institute for School Reform, share design principles they used to build an evaluation that is an integral part of program design, implementation, and improvement.¹

A number of these principles, such as creating community-wide investment and engaging stakeholders, showed up in the earliest phases of the evaluation. Because DAP is a collection of stakeholders with varying definitions of success, creating communitywide investment in the evaluation required collective negotiation to sift through individual needs and find commonly valued outcomes. When the stakeholders first met, each wanted to measure outcomes that would advance its own work. Evaluators as well as Big Thought's executive director saw value in having everyone struggle to find the middle ground. Dallas Independent School District (DISD), a major partner, could not continue to invest in ArtsPartners without proof that the program supported student academic achievement. However, the arts community felt this was only a partial measure of what can result from arts education. Through collective negotiation, the stakeholders agreed to focus on literacy as an outcome. Literacy is linked to key forms of educational accountability such as No Child Left Behind, but when broadly defined it encompassed how children capture information in rich ways (such as through music and symbols) and how they learn to express their ideas imaginatively and with originality. This led to an evaluation that involved multiple measures of literacy, including state test scores and an indexing of children's investment in learning and children's ability to translate their unique experiences and voices into writing. The evaluation forced the stakeholders to wrestle with common outcomes, offering a new level of collaboration and learning across DAP, and enhancing the capacity of all participants to think about shared goals.

The principles of shared and early decision making resurfaced when it was time to finance and staff the evaluation. Committing to multiple measures required resources ArtsPartners did not have. As a result, Big Thought and its partners had to narrow the focus of the evaluation to the students and not include as much focus on the teachers as originally planned. They agreed on the need for a major longitudinal evaluation of DAP's impact and moved forward on faith that they would find the funding needed. In 2001, they wrote an unsuccessful grant proposal to the U.S. Department of Education (DOE). They cobbled together general operating support monies to pay

for the first two years of the evaluation, while exploring other funding options. Eventually, in 2003, they successfully competed for DOE grant funding based on the rigorous and intensive longitudinal study they had begun. This grant paid for the final three years of the evaluation.

Through this process, all the partners learned that you have to be willing to pay for a good evaluation, especially one that builds the capacity of an organization or a network of stakeholders. This includes honoring people's time and talents. "Sometimes in nonprofits we expect everyone to come to the table out of the goodness of their hearts," said Bransom. Throughout the evaluation, Big Thought has put value on people's time and honored it. For example, they paid teachers a stipend to attend workshops to learn how to participate in the evaluation. They also budgeted funds to pay back the cost of having one of the school district's own evaluators, Katy Denson, work on the evaluation.

This investment speaks to another key feature of this evaluation: its shared design. Denson brought the perspective of the school district to the table. She helped to negotiate district categories and what measures meant, whom to contact, and how to write a report that honored the perspectives and interests of the schools as well as the external partners. This evaluation also included an external evaluator, Dennis Palmer Wolf, who is recognized nationally in the arts education world. As a senior scholar at the Annenberg Institute for School Reform, Wolf and her colleagues at the institute could connect the work at Big Thought to national efforts to broaden the definition of learning and expand thinking about partnerships in education.

Through the process of seeking funds for the evaluation, the organization learned that evaluation should be funded as part of every program request in the partnership. Now, when a program director is creating a new program (guitar lessons in neighborhoods) he or she automatically includes evaluation as a percentage cost (either in a budget or in a grant proposal). The organization views evaluation as a process as much as a product. Says Bransom, "The process of evaluation in our organization involves asking questions that inform our learning to make programs better for kids. That has to be the first reason to do evaluation. That is why it has to be part of a program's cost." This learning focus allows them to be open to surprising and even negative findings because, as Wolf observes, "assessment of any kind should be an episode of learning."

Big Thought staff members have learned that the person who collects and analyzes the information ends up knowing the most about a program. So they ask their partners to collect data so they can immediately start learning from it. Involving the artists in the research has allowed them to learn as they go along. The artists have found that being involved in the evaluation can be more meaningful than what they learn from the professional training they receive early in the process. DAP also involves teachers in the evaluation process. Having both teachers and cultural community partners at the table has helped everyone gain perspective.

For example, when, as part of its longitudinal study, the evaluation team looked at the student writing samples, they found mixed results. It was depressing at first because they weren't seeing the impact they expected. But they looked at what was happening during the semesters as the writing quality declined. They presented it to the stakeholders and asked, "What is this telling us?" The teachers realized that the declining results were in a specific area in which the students struggled: expository writing. It was a good opportunity for the teachers, then, to see that they needed to get better at integrating science into writing. The team members' focus on learning allowed them to report the results in a way that was useful and outlined next steps for the program.

"A learning organization has to prove that it will look at all results and take them as they are, not bury them," says Katy Denson of Big Thought. Big Thought and DAP are committed to seeing all information as an opportunity for learning. The DAP evaluators have learned to be prepared to address whatever unexpected answers they unearth when they ask questions. For example, when they interviewed parents as part of an assessment of how to expand arts programs into communities and neighborhoods, they asked parents to share their hopes, dreams, and concerns for their children and how the arts could support those dreams. They wanted to hear parents' concerns about having their children participate in the arts in a neighborhood institution. The evaluators and program design team talked ahead of time about what they would do with the information and agreed they had to be committed to applying what they heard to make their programs stronger.

Healthy Native Communities Fellowship

The Healthy Native Communities Fellowship (HNCF) is one of several national programs developed as part of a comprehensive plan of the federal government's Indian Health Service (IHS) National Prevention Initiative. Its mission is to "increase the capacity of local teams of Tribal leaders and advocates to promote health and wellness within Native communities" (Healthy Native Communities Fellowship Annual Report, 2006). This government program recently took steps to become a nonprofit organization so it can build collaborative partnerships with other organizations, including bringing in other sources of funding for the program. HNCF has implemented evaluation practices and strategies that are useful examples for nonprofits striving to develop a learning culture.

The fellowship develops teams of people from tribes located across the country to build community connectedness, cultivate the variety of strengths that regenerate and heal tribal communities, nurture the talents and leadership of community members, and develop effective strategies to tackle the problems and issues that threaten the community. In each of the three cohorts to date, 15 to 17 cross-sector teams participated in four retreats over a 12-month period to increase their knowledge, skills, and capacity to promote health and wellness in their community. As part of their participation

in four weeklong retreats, teams develop learning plans on how they will apply these skills to strengthen wellness efforts back home. The program office, located in Shiprock, New Mexico, on the Navajo Nation, is funded solely by Congress through the Indian Health Service.

The HNCF's evaluation is led by the program's "learning team," which is an interdisciplinary group of HNCF staff, University of New Mexico Master's in Public Health Program faculty, and outside consultants. To determine the effectiveness of the program components and recommended program changes according to feedback from Fellows, Fellowship teams (there were various teams of individual Fellows), and HNCF faculty; and to track changes at the individual, team, and community levels to determine HNCF's impact (Healthy Native Communities Fellowship Annual Report, 2006). HNCF's participatory evaluation approach has involved HNCF Fellows and faculty in creating and reviewing the core evaluation questions, giving feedback on the instruments (which has often led to new or modified instruments), and interpreting the data.

The HNCF has four main stakeholder groups:

1. The Health Promotion/Disease Prevention Policy Advisory Committee, representing tribal leaders and key national tribal organizations such as the National Indian Health Board and the National Congress of American Indians (this group provides input from tribal leadership and community-level perspective).

2. The IHS Prevention Task Force, a 20-member multidisciplinary team including IHS area health promotion and disease prevention coordinators and a representative from the Centers for Disease Control and Prevention (CDC). This task force has a broad focus on all health promotion programs within the IHS. Some of the area health promotion and disease prevention coordinators are also HNCF Fellows.

3. The HNCF Fellows, who are interested in seeing program improvements made on the basis of their input through the evaluation and the evaluation's ability to describe the transformation they are experiencing during the fellowship and its impact on their home communities.

4. The director of the IHS, who oversees the program and has a strong stake in its success. He has an interest in demonstrating that the IHS can prevent health problems, not just treat them, and is working at gaining full support for this program from IHS area directors.

From HNCF's inception in 2004, director Marita Jones and Chris Percy, IHS National Prevention Initiative Task Force chair, considered evaluation as a critical component of the program because they wanted to learn about the program's impact at the community level. Although there are no major formal requirements for the evaluation, the most common interest that all the stakeholders have expressed is to understand the program's impact at the community level. For example, the Policy Advisory Committee wants to know if

communities are successfully submitting more grant proposals as a result of the program. The Prevention Task Force is interested in identifying replicable programs and best practices in mobilizing communities toward improved health. The IHS director wants to demonstrate that this program results in long-term community change. The HNCf learning team is aware of these expectations and the challenges of meeting them through the evaluation. The team has taken a proactive approach with the evaluation, anticipating what will be needed and acting accordingly. They have also responded to critiques of the evaluation as it has progressed. For example, on the basis of feedback from stakeholders, they have increased the relevance of the evaluation reports by shortening the length and giving visual snapshots with information in graphs and bullets.

The evaluation plan was primarily negotiated between the learning team and the HNCf faculty. Because of the participatory nature of the evaluation, members of the learning team participate in curriculum discussions and are active in the HNCf retreats. To get input for the evaluation from the participants, the learning team put together a small advisory committee of HNCf graduates who give feedback on the evaluation. Overall, the stakeholders in this evaluation have not created burdensome demands or asked the program to carry out activities that do not fit within its mission.

How did the organization build its capacity to meet the needs of the evaluation and keep the focus on learning? One of their biggest challenges as a new program was to figure out exactly what the evaluation model would be. Developing the model has been an iterative process that continues to be refined over time. At the beginning, the focus was on Fellows and team transformations, with the belief that community change might be too challenging to assess. Yet policy stakeholder interest in community change led to strategies to better document community actions, footprints, and changes at the community level. Triangulation of multiple sources of data (end-of-week retreat evaluations, personal assessments from Fellows, team assessments, team interviews at program midpoint and end, and each team's documentation of its own actions and success stories on a workstation Internet platform) has led to the ability to monitor trends in community changes.

Community outcomes have included a variety of system and capacity changes, including policy actions and new policies achieved, new services and programs, new resources, new collaborations, new participation of elders and youth, and new tribal leadership recognition among others. The background frame for this model was derived from previous participatory evaluation efforts in Healthy Communities in New Mexico (Maltrud, Polacsek, & Wallerstein, 1997)² and by the Work Group for Community Health and Development at the University of Kansas (<http://communityhealth.ku.edu>). For the past two years, these changes (along with representative team and Fellow transformations) have been pulled together into an annual report card, designed for easy use by policy makers. With this strategy for capturing

system and capacity changes, HNCF expects to see health changes in the future that reflect the mission of the HNCF.

The learning team includes HNCF's director, internal evaluation and program staff, and three external consultants with expertise in evaluation. HNCF leveraged a long-time partnership with Nina Wallerstein of the University of New Mexico to bring her evaluation expertise and an external perspective to the learning team. The budget for the program did not automatically include significant funding for the evaluation, but the program's director reallocated funds within the budget early on to adequately support the evaluation. Currently, approximately 14% of the total program budget is allocated for evaluation activities.

Marita Jones (HNCF director) is an active member of the learning team. A strong evaluation helps her adequately express the benefits of the program to stakeholders, including Native communities throughout the United States, including Alaska. She also wanted to be on the learning team (instead of just providing oversight) to learn more about data analysis and reporting so she could convey the benefits of the program to future funders and other stakeholders. She has found that being part of the team helps her effectively communicate the value of the program. When people ask, "How do you know it works?" she wants to be able to fully answer the question. The team is striving to understand and document the wisdom of local tribes and native people and translate this wisdom into promising practices and strategies. This aspect is part of Jones's passion in learning more about evaluation.

Jones's administrative perspective has also been helpful to the evaluators. For example, when they shared the draft of an interim report with her, she saw that some teams of Fellows were rated more highly than others on key indicators. She helped the evaluators identify coaching calls with teams of Fellows as an additional data source for the evaluation to better understand this difference. Also, as the HNCF director, Jones has the most interaction with all of the stakeholders (with whom the evaluators don't get that much contact) and has helped them run interference with some of the stakeholders. It has been wonderful for the evaluators to see the skill and interest of their program director in the evaluation process. Having an administrator on the evaluation team helped the evaluators see some of the holes in the evaluation and help connect the dots in the evaluation process. The learning team also includes Shelley Frazier, the director of Just Move It, another national IHS program based at the Shiprock area office. Shelley brings quantitative computer skills to the team as well as skills and ideas about how to make evaluation more valuable and user-friendly. She saw that a lot of the communities were doing great things, but the challenge was how to measure and share these accomplishments. She has gained insights and experiences working on the HNCF learning team that she has been able to apply to her own program. The involvement of a broad group of stakeholders including those involved in other health promotion initiatives at Shiprock has helped strengthen the effectiveness of the team as well as its commitment to evaluation as a learning process.

Evaluation can be a scary word in Indian Country, given past history in which predominantly white researchers misused information collected from Indians for program planning and research. HNCF evaluators are committed to being candid and direct with Fellows about the evaluation purpose and process. During the first retreat of each HNCF cohort, the learning team talks to the Fellows about the evaluation, including the overall evaluation model, what they will measure, how they will measure it, and what they will look at and report back at the beginning of the next meeting. They have also extended an invitation to each cohort for participants to join the learning team, and members of the HNCF alumni advisory group have offered input to the evaluation efforts. The learning team agreed to make sure that any data a Fellow is asked to provide for the evaluation (complete an assessment, evaluation form, interview) will be given back to the Fellows either as confidential copies of individual contributions or as an aggregate report posted on the Website. The team members have learned to be very transparent with participants. Kristine Maltrud, one of the HNCF consultant evaluators, says she has developed extra sensitivity to participants' perceptions of evaluation since working with this project. She communicates as clearly as she can and does not make assumptions about what has been understood. To make sure they gather relevant evaluation data for making ongoing program improvements, members of the HNCF Learning Team present a "participatory evaluation feedback" event at the beginning of each Fellowship retreat during the year. They share up-to-date evaluation results with the Fellows, including changes they have made in the program and changes they plan to make for upcoming groups of Fellows. Throughout the fellowship year, they ask the Fellows and program staff and faculty how the program (curriculum content, materials, logistics, and so forth) can be improved. Many changes and adjustments to the program have been made this way, including changes to the evaluation itself (both instruments and design). This is a clear example of how the evaluation is an integrated component of the program and is focused on helping HNCF achieve its mission.

Conclusion

We hope this chapter offers both a compelling picture of the challenges facing the field of evaluation and hope for the future. The three organizations and programs profiled in this chapter are but a few of the many groups and organizations that are finding a way to make evaluation a meaningful part of their journey toward the realization of their mission. Key examples from the case studies highlight the power of evaluation as part of a learning culture: resourceful problem solving among the staff at Big Brothers Big Sisters, learning-focused recognition of the need to concentrate on expository writing at Dallas ArtsPartners, and creative ways to understand community impact at Healthy Native Communities Fellowship. But many nonprofit organizations will not make a successful transition to a learning-focused

Table 6.1. Best Practices in Evaluation for Organizational Learning

<i>Big Brothers Big Sisters of Harrisonburg-Rockingham</i>	<i>Big Thought and Dallas ArtsPartners</i>	<i>Healthy Native Communities Fellowship</i>
<p>Mission-driven data collection</p> <p>Investing in and actively using data collection and reporting systems</p> <p>Taking advantage of evaluation resources provided by a national organization</p> <p>Making sure external evaluators provide locally relevant and useful information</p> <p>Developing staff capacity to regularly use data to improve organizational performance</p>	<p>Investment early on in negotiating and agreeing on outcomes among multiple partners</p> <p>Shared evaluation design and implementation</p> <p>Persistence in gaining grant funding to augment evaluation resources</p> <p>Involving stakeholders in data collection to accelerate learning from the evaluation</p> <p>Presenting data as part of a context for collective learning and ongoing program improvement</p>	<p>Using participatory evaluation approach</p> <p>Proactively anticipating stakeholder needs and designing evaluation accordingly</p> <p>Taking on challenge of understanding community impact</p> <p>Program funds reallocated early on to cover costs of evaluation</p> <p>Evaluation managed by “learning team” that includes program director as active member</p>

organization without help from evaluators and evaluation capacity building efforts. More important, funders play a key role in shaping much of the evaluation work and the reporting nonprofits currently undertake as well as in funding evaluation capacity building efforts. If funders make sure the evaluation work they require aligns with and supports efforts aimed at program improvement, is always focused on the mission, and develops and funds more effective ways to increase evaluation capacity, then funders can do their part to make sure evaluation does not fail the social sector.

For BBBS, this critical support can be seen in the work of the national organization in funding large evaluation studies and providing evaluation resources, as well as the responsiveness of the external evaluator in providing locally relevant data. Both Dallas ArtsPartners and HNCF used evaluation expertise from outside professionals who became full members of the team, with a clear focus on building evaluation capacity. Evaluation funding, either from outside sources such as DOE or from reallocating money from the program budget, was also an important element in these case studies.

The best practices from each case study are summarized in Table 6.1. These practices clearly align with Alaimo's Type I organizations (see Chapter 5 in this issue). Each group visibly demonstrates a long-term commitment to evaluation and evaluation capacity building, gathering committed teams composed of staff, outside evaluation expertise, and in some cases participants, to guide the work of the evaluation and focus on learning. Comprehensive approaches are used and evolve over time to best meet the needs of the organization and program. Prioritization of evaluation use for improvement, rather than just to meet accountability demands, is a hallmark of all these organizations.

The key lesson from the case studies presented in this chapter is the role of leadership at all levels in effectively using evaluation to foster a learning culture, and commitment to using the information gained through evaluation efforts for program improvement to further the mission of the organization. Capacity building efforts by funders are important, but the commitment of organizational leadership is critical in transforming the role of evaluation from one of basic reporting and accountability to a true process of continuous organizational learning. This transformation requires leadership commitment to development of an evaluative learning culture and development of an infrastructure to support it.

Notes

1. The principles are (1) tailor the evaluation to the context, (2) create community-wide investment, (3) engage stakeholders in key decisions early, (4) enhance the capacity of all participants, (5) plan for midcourse corrections, (6) grapple with uneven findings, (7) stay alert to surprises, and (8) share and use the findings.

2. This workbook is available from Nina Wallerstein at the University of New Mexico (e-mail: nwallerstein@salud.unm.edu).

References

- Bickel, W. E., Millet, R., & Nelson, C. (2003). Challenges to the role of evaluation in supporting organizational learning in foundations. Washington, DC: Council on Foundations.
- Big Brothers Big Sisters Website. Retrieved September 25, 2007, from <http://www.bbbs.org/site/c.ieJMIWOyEnH/b.3560037/k.BE4C/Home.htm>
- Big Thought. (2007). Retrieved September 25, 2007, from <http://www.bigthought.org/>
- Botcheva L., White, C. R., & Huffman, L. C. (2002). Learning culture and outcome measurement practices in community agencies. *American Journal of Evaluation*, 23(4), 421–434.
- Carman, J. G. (2007). Evaluation practice among community-based organizations: Research into the reality. *American Journal of Evaluation*, 28(1), 60–75.
- Fredericks, K. A., Carman, J. G., & Birkland, T. A. (2002, Fall). Program evaluation in a challenging authorizing environment: Intergovernmental and interorganizational factors. In R. Mohen, D. J. Bernstein, & M. D. Whitsett (eds.), *Responding to sponsors and stakeholders in complex evaluation environments*. New Directions for Evaluation, no 95, pp. 5–22. San Francisco: Jossey-Bass.
- Grossman, J. B., & Rhodes, J. E. (2002). Test of time: Predictors and effects of duration in youth mentoring relationships. *American Journal of Community Psychology*, 30(2), 199–219.
- Healthy Native Communities Fellowship Annual Report. (2006). Indian Health Service. Retrieved September 25, 2007, from http://www.healthynativecommunities.org/initmaterialist.jsp?initiative_id=7
- Hoole, E. (2007, November). Evaluation capacity building and the role of funders. Paper presented at the Annual Meeting of the American Evaluation Association, Portland, OR.
- Katz, S. N. (2005, June). What does it mean to say that philanthropy is “effective”? The philanthropists’ new clothes. *Proceedings of the American Philosophical Society*, 149(2), 123–131.
- Maccoby, M. (2003). The seventh rule: Creating a learning culture. *Research Technology Management*, 43(3), 59–60.
- Maltrud, K., Polacsek, M., & Wallerstein, N. (1997). *Participatory evaluation workbook for community initiatives*. University of New Mexico School of Medicine, Master’s in Public Health Program and New Mexico Healthier Communities.
- Porter, M. E., & Kramer, M. R. (1999, November/December). Philanthropy’s new agenda: Creating value. *Harvard Business Review*, 121–136.
- Preskill, H., & Torres, R. (1999). *Evaluative inquiry for learning in organizations*. Thousand Oaks, CA: Sage.
- Snibbe, A. C. (2006, Fall). Drowning in data. *Stanford Social Innovation Review*, 39–45.
- Snyder, G. (2007). The nonprofit sector crossroads. Retrieved September 26, 2007, from <http://www.ncrp.org/blog/2007/08/nonprofit-sector-crossroads.html>
- Torres, R. T., & Preskill, H. (2001). Evaluation and organizational learning: Past, present, and future. *American Journal of Evaluation*, 22(3), 387–395.
- United Way of America. (2000). *Agency experiences with outcome measurement*. Retrieved September 26, 2007, from <http://www.unitedway.org/Outcomes/Resources/What/upload/agencyom.pdf>

- Wolf, P. W., Bransom J., & Denson K. (2007). *More than measuring: Program evaluation as an opportunity to build the capacity of communities*. Dallas: Big Thought.
- York, P. J. (2003). *Learning as we go: Making evaluation work for everyone*. Retrieved September 21, 2007, from http://www.tccgrp.com/pdfs/per_brief_lawg.pdf

EMILY HOOLE is the director of the Evaluation Center at the Center for Creative Leadership (CCL), a nonprofit educational institution in Greensboro, North Carolina.

TRACY E. PATTERSON is a manager with the CCL Evaluation Center.