

Accountability, Evaluation, and Organizational Learning

TODAY, contributors, donor agencies, scholars, and relief and development practitioners are all asking: Do NGOs practice what they preach? How do we know? Are claims of closeness to the grassroots a reality? Have NGOs become too dependent on public money? How effective are their programs and projects? How credible is the evidence? Would donors reach the grassroots better by connecting with local NGOs in developing and transitional countries? If donors must be transparent about evaluation results, why not NGOs?

In short, in this field and many others, accountability is the central issue of our time. We turn to this issue as this book winds toward its close. In fact, however, elements of the accountability challenge are present implicitly in all of the preceding chapters: coping with globalization, transformation, new challenges in relief work, poverty reduction, partnerships, advocacy work—each of these has an accountability component. As that is the case, what remains to be discussed here are the components within the accountability equation that have not yet received close attention: accountability's contingent character and complexity, the role of evaluation systems, and the learning process that should emerge through collective reflection on past performance. Before turning to these three major themes, we will prepare the way with some background discussion and address why this issue emerged as salient in the 1990s.

Introduction: The Challenge

NGOs are at present challenged as never before to demonstrate *results*. Internet users can, within minutes, compare and contrast who is doing what and with what outcomes. Current and prospective donors ask questions about impacts. Nor is there any indication that these pressures will abate. Calls for stronger accountability grew steadily throughout the late 1990s. In Europe, Roger Riddell and Mark Robin-

son detailed the weak evaluation capacity in European-based NGOs.¹ Michael Edwards and David Hulme made accountability the major theme of *Beyond the Magic Bullet*.² Jonathan Fox and David Brown examined at length the conceptual thinking in the NGO demands for World Bank accountability.³ Now NGOs are hearing comparable calls for their own accountability.

In the early 1990s, NGOs pressured the World Bank and the IMF to be more “transparent” and accountable. By the end of the 1990s, the World Bank enacted and implemented new regulations on public access to project evaluations. Sometimes the NGOs were asking the Bank to provide more public access to their project information, including evaluations, than NGOs themselves were prepared to allow. Of course, they were not alone—few official bilateral development agencies were disclosing project documentation in the early 1990s. Yet inexorable pressures in the information revolution, coupled with calls for performance-based management, are now pushing NGOs for more accountability.

Accountability and Public Funding

One reason often given for needing more NGO accountability is growing dependence on public money. Ian Smillie’s reckoning, for example, is that “by the early 1990s, 75 percent of British food aid was being channeled through NGOs, and 40 percent of Swedish spending on emergencies and refugees was going through Swedish NGOs. By 1996, 46 percent of French emergency funding was being spent through NGOs, and half of all the EU’s European Community Humanitarian Office (ECHO) funding was being spent the same way. Between 1992 and 1997 the United States Agency for International Development (USAID)—with the largest emergency budget in the world—spent over 60 percent of it, not counting food aid, through NGOs.”⁴

International organizations are also increasingly working through NGOs: “The World Food Program channeled an estimated \$580 million through NGOs in 1997, UNHCR, about \$300 million. Twelve percent of UNICEF’s country level programming is directed towards NGOs. Further NGO grants, co-financing and contracts are made by UNDP, the United Nations Fund for Women (UNIFEM), the International Fund for Agricultural Development (IFAD) and other UN agencies.”⁵

Yet some of the NGOs with the strongest accountability cultures are the same ones that are taking the lowest amounts of public money. Thus the motivations or “drivers” for accountability within NGOs are more nuanced and complex than the role of public money; financial drivers are not determinative.

Accountability and Advocacy

The previous chapter detailed the increase in advocacy work. As the visibility of that work increases, it underscores and provokes demands for accountability and its correlate, responsibility. If NGOs are to be advocates in public policy, to whom are they accountable for that role? They are not membership organizations; hence they cannot readily claim to represent distinct membership needs.⁶ While their claims to credibility on policy issues derive from their claims to “field”-informed knowledge, they are expected to answer for these claims as much as for their programs and projects.

Policy advocacy also raises legal issues of standing which requires that any claimant for a legal action must have some immediate interest at stake providing that agency, or person, with a right to public recourse. In the case of the NGOs’ meetings with the UN Security Council, the NGOs had standing for seeking recourse, given the risks to their own staff in civil wars. But in their advocacy work on issues around a broadly shared common good—education for all, or debt relief—they are interpreting and advocating on behalf of others. They are taking up a cause on behalf of the poor, but cannot claim to be elected representatives of the poor. Norman Uphoff argues in light of this point that NGOs cannot be referred to as a “third sector,” as that should be reserved for membership organizations and community groups grounded within localized systems of accountability.⁷

Most national governments have regulatory standards in regard to NGOs, usually insisting upon fiduciary responsibility, accounting, and auditing reports. From time to time, various national regulatory bodies for monitoring nonprofit work will look in more detail at the advocacy work of an NGO as well. The Charities Commission in Britain, for example, has from time to time examined Oxfam GB’s advocacy work. In France, the Cours des Comptes selectively undertakes intensive financial audits of NGO accounts. A critical review from the Cours des Comptes can spell real trouble for an NGO—affecting its fundraising appeal if public credibility is tarnished. As the work, size, numbers, and financial scale of NGO activity increases, we can expect that the legal and regulatory environment of NGO work will be impacted by national accountability standards.

Definitional Problems

Accountability, an elusive concept, is more discussed than clarified. For example, Jonathan Fox and David Brown say that “accountability refers to the process of holding actors responsible for actions,” but without clarifying who holds whom accountable, or how. Mark Moore argues it has to do with responsibility. For some, the term means “giving an

account” for decisions, with particular attention to decisions that entail expenditures. For others, it means responsibility to, above all, the people affected directly by projects and also to donors as well as the NGO board and other stakeholders. For the purposes of this discussion, accountability means answering to stakeholders, including beneficiaries, boards, donors, staff, partners, and peers for the results and impacts of performance and the use of resources to achieve that performance. In short, these agencies must respond to demands that emanate from multiple sources. There are inevitably both normative and legal drivers in the accountability equation.

While accountability is used frequently as synonymous with transparency, there are distinctions between the two. Accountability is providing an account for decisions, actions, and their consequences. Transparency is about providing information—not necessarily an explanation. Although they are interrelated, these concepts are nonetheless separable. Quantities of information may be provided in ways that do not directly address or reveal the connections between decision-makers, decisions taken, and their consequences. Moreover, techniques, or recourse for holding decision-makers accountable, vary—from lawsuits to elections. The forum and processes for accountability vary. Information is essential, but not determinative for achieving accountability.⁸

Commitment and Complexity in Accountability Processes

Throughout our sessions with them, our interlocutors frequently discussed their abiding commitment to strengthen accountability. In a discussion on what was involved in doing so, there was broad agreement when Reynold Levy, CEO, International Rescue Committee, said that accountability to beneficiaries had to be the highest priority. James Orbinski, president, Médecins Sans Frontières International, added that it was hard, however, to prioritize as answering to volunteers, beneficiaries, and donors were all equally important—creating a circle of accountability. Charles MacCormack, president of Save the Children US, added that one part of the problem is how broadly or narrowly to set aspirations, as it is easier to measure and be accountable for narrower ones.⁹

For all of them accountability means adhering to organizational core values, their mission, and performing up to their own standards. The ways to “render an account” range from answering to the collective wisdom of shared values across the NGO family; to using public speeches and public education campaigns to exercise responsible leadership; to providing financial and programmatic data via Web sites, annual reports, press releases, and documentation. It also means telling people what their operational work is achieving and answering to boards, contributors, beneficiaries, and other stakeholders for their decisions.

The several tiers of groups and agencies these NGO families operate with requires a multi-directional flow of information that reaches from headquarters to the field and to the people directly affected by programs and projects and back again to the NGO. Robert Chambers asks: *Whose Reality Counts?*¹⁰ One must ask in turn, Who decides? Boards, stakeholders, donors, partners, project participants? Those in NGOs work with these uncertainties every day, while fielding operations or executing advocacy campaigns. Unfortunately, all too often work pressures preclude reflection on critical questions.

Whose reality *does* count? And therefore, to whom, when, and how does one give an account? This chapter looks at these questions and conveys what various NGOs are doing now about accountability. We will also consider how much progress there is with “becoming learning organizations.” As in all chapters, we will then turn again to the four ethnology questions addressed at the end of every chapter. But first we turn to ask why this issue is now being raised everywhere—and with such virulence.

Why Now? What Has Changed?

What has put NGO accountability on the public agenda? After all, for the first decades of their lives, most NGOs enjoyed widespread public approval without many probing questions. Why has that changed?

As we saw in the earlier chapters, *the whole context has changed*. Only in part are the changes in roles being assumed by NGOs responsible for the changed demands for accountability. The fact that many NGOs are receiving greater amounts of public money drives some of the demand, but it, too, is only part of the story.

The second part of the explanation is embedded in the changed ideological climate of the new era. The politically charged and ideological climate of the Cold War era dampened the development debate generally and the role of NGOs as well. Especially in the United States, during the Cold War, there was little tolerance for anything that could be interpreted as left of the political center. As the Berlin Wall came down development theory and practice shook loose as well. Amartya Sen’s book, *Development as Freedom*, exemplifies this new climate; his role as honorary president of Oxfam epitomizes new roles and orientations in this millennium.¹¹

Discussions of poverty today that call immediate attention to vulnerability, to powerlessness, were not well received when they were articulated during the Cold War. Liberation theology, which did make those points, was labeled dangerous and subversive. But now even the World Bank argues that poverty is about powerlessness, and the European Union has a policy on social exclusion. The drafters of the World

Bank's 1980 World Development Report on poverty would have found themselves in serious trouble had they tried to incorporate "vulnerability" and "powerlessness" into their analysis. But NGOs were making that point during the 1980s—and even earlier. Oxfam and Save the Children have long records of working on "empowerment" as central to development, and they have trained trainers to carry that message. NGOs are often concerned to help empower the poor—to give them voice, to facilitate their becoming more powerful.

Coming with this new freer climate, however, is a call for accountability, for demonstrating results, for giving an answer to constituents about what works. Is accountability the current price for the current, more open environment? Maybe. But maybe not. Maybe accountability is the current price, and responding openly may help keep the open environment—open. In general, that is what we implicitly argue in this chapter. But we also want to make the risks of some of the current pressures for demonstrating results more apparent. Demonstrating results is a popular idea. Interpreted too narrowly, it has unintended negative consequences.

Michael Edwards and David Hulme point out that the new emphasis on empowerment adds a political dilemma for those charged with providing accountability: "If the organization's overt or covert goal is empowerment (making those who have little power more powerful), then transparency on this issue will, at best, make it easier for vested interests to identify what is happening and thus more effectively oppose it, or, at worst, lead to the deregistration and closure of the organization for being subversive."¹²

It is not clear that they are right. In the 1970s, many NGOs, including Oxfam GB, were active in Latin America and Central America during the struggles in those countries—doing just what one would have thought not possible: working with opposition groups struggling for more democracy. They were not deregistered, or closed. Moreover, contrary to Michael Edwards's and David Hulme's claim that no NGOs have been successful at democratization, the carefully researched work of Brian Smith in *More Than Altruism* documents that several NGOs were successful in reaching and facilitating alternative paths to influence, and that many of the democratic leaders who emerged later in Central America came out of the NGO movement seeded by this work.¹³

Yet in some ways they are right—and their point should be heeded. Empowerment comes through long-term, incremental steps—it is a fragile process, easily uprooted. Demanding "results" can mean either pushing for quick fixes, or insisting upon digging up the seedling to examine its roots before it can bear fruit. Domestic poverty work often is abandoned when narrowly measured results are demanded—and not found. Empowerment is sometimes more tolerated at a distance than it

is at home; it seems easier to deal with powerlessness in small countries than with the powerlessness of poor people within Northern post-industrial countries. John Gaventa writes about the speed with which the War on Poverty in the United States in the 1960s was dismantled when it began to be effective. It is a powerful case in point.¹⁴

The “indicator movement” irrevocably changed the way that development theorists and practitioners think about “results.” Techniques and skills for the measurement of results have changed in the last three decades. As early as the 1960s, the Yale University Political and Social Data Center was developing social and political indicators to parallel the indicators in use by economists. Knowledge spread that something akin to social accounting (measuring results) was possible. By the 1980s, the multilateral development banks had picked up on performance indicators—and so had managers coming out of business schools. This led in turn to more calls for improved accountability in the public sector. It was only a matter of time until this would be echoed in development work, even though national statistics and census politics in developing countries meant weak databases for measurement.

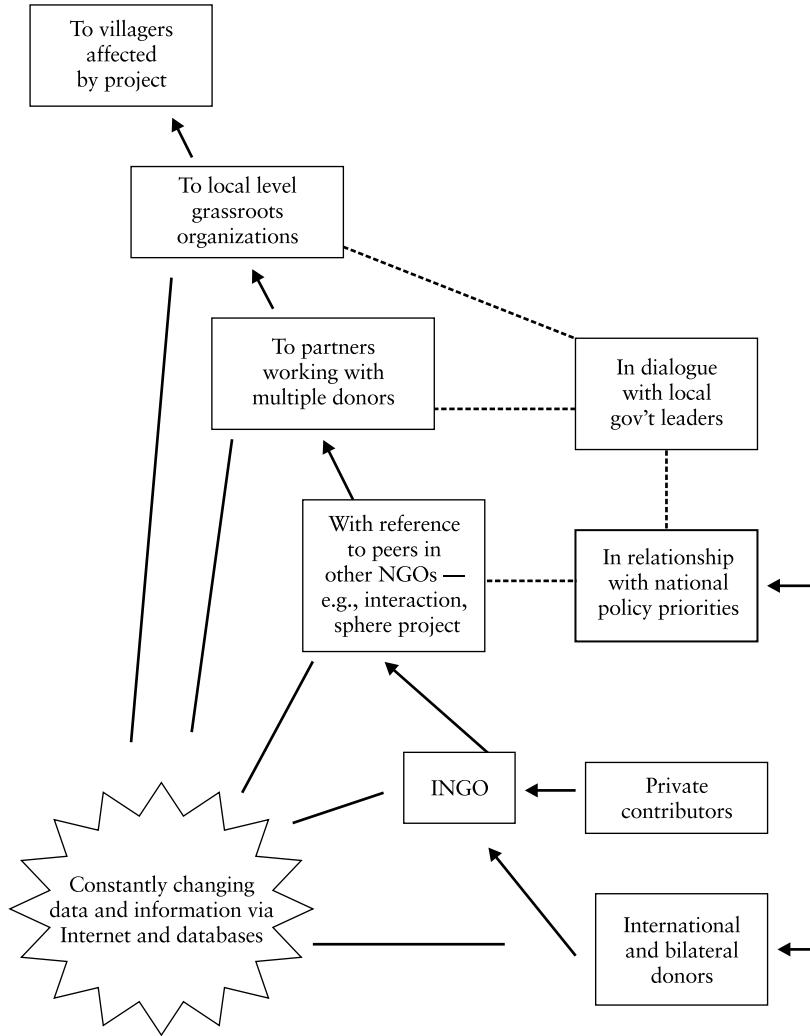
Multi-Directional Accountability

Those who write about domestic nonprofit organizations point out that nonprofit evaluation is complicated by the organization’s multiple goals, multiple constituencies, and market insulation. If this is true of a domestic NGO working where the contributors see the results of their donations, consider how it works for international NGOs. Alan Fowler provides a thoughtful example of the long chain through which resources must flow winding through geographic, sectoral, and institutional spaces to affect objectives, goals, and performance.¹⁵ These interactions determine outcomes. The nature of these layered tiers of agencies affects several dimensions of accountability.

Accountability must be *multi-directional* as a result of these tiers—not only from bottom to top, or from top to bottom. In reality, different parts of any NGO family have to respond to different stakeholders, partners, and people—immediate beneficiaries and others—in order to meet the accountability challenge. Figure 8.1 on the following page depicts the multi-directional nature of the accountability process.

Accountability for emergency relief work is different from that of poverty-reducing work, or longer-term development work. As there are weaker links in emergency situations, and different constraints, it is extremely difficult to measure effectiveness. Both public opinion and the donors funding relief work (for example, UNHCR, or the Office of Disaster Assistance) have reporting and auditing requirements that differ from those of ongoing development assistance with a longer timeline.

Figure 8.1. Multi-Directional Accountability



Given the changing context for humanitarian work, however, changed accountability standards will emerge, especially as the borders between relief and development blur. But to date, differences in approach and techniques have meant different reporting and evaluation guidelines as well. In the case of development work, there is longer-term engagement with people and Southern NGOs over a period of time, making possible more participatory and thorough evaluation.

The very volatility of work with internally displaced people and refugees adds dimensions of difficulty to evaluation processes. That said, the Sphere Project has as its aim to improve the quality of assistance provided to people affected by disasters, and to enhance the accountability of the humanitarian system in disaster response. This project was worked out by the NGOs together with the UN to develop minimum universal standards and codes of conduct for disaster response.

While public opinion in general has been sympathetic to the added demands of relief work, and thus has lowered expectations for accountability, in the case of Rwanda, as we saw in Chapter 3, there was in fact more criticism of NGO work. Peter Uvin has suggested that NGOs should be held accountable, as should the international community, for the structural violence in Rwanda.¹⁶ His argument is that NGOs working directly with people at the grassroots had before them all the signals of impending major violence, but that they remained too silent for too long. Enmeshed in their own worlds, they did not blow the whistle. When the violence escalated, and they undertook emergency work, they were often manipulated by terrorists, who outmaneuvered them—for example, by using food programs in refugee camps as foraging places to “refuel” before their next rampage. While some NGOs did decide to withdraw when they became convinced that they were being used by those committed to violence and further bloodshed, others did not decide to withdraw—on equally compelling grounds that the people with whom they were working would be even more immiserated if they abandoned them than if they remained and kept working. The debates about those choices—each of which was very context-specific—will continue for years to come.

NGO emergency relief work is not likely to be the same in the post-Rwanda world. While the Group of Seven industrial powers was largely focused on the wars in Bosnia and Croatia and later on in Kosovo, the numbers killed in those wars, grim as they were, were nowhere near the nearly million lives lost in Rwanda. Nor were NGO staff as deeply at risk.

The sum total of the multi-directionality entailed in these various calls for increased, and differently conceived, accountability and responsibility systems has changed the context for NGO senior managers and staff. Since the accountability process is both multi-faceted and diverse, conceiving of it in terms of a contingency approach may be helpful. Below we turn to this contingency model for accountability—an approach pointed out to us by, among others, David Brown and Julie Fisher, while discussing NGO accountability with them.¹⁷

A Contingency Model for Accountability

In all our discussions with them, NGO leaders confirmed that their accountability is both multi-dimensional and multi-directional. While most perceived their primary responsibility to be toward those directly affected by programs and projects, they added that they must also answer for performance to contributors and donors. One of the implications of multiple and competing audience pulls is that different kinds of information and feedback are needed for different audiences. Therefore accountability is necessarily contingent upon both the demand and the context of the demand. For example, donors often require formal evaluation reports, while other contributors expect some summary financial reports and annual mailings; beneficiaries, on the other hand, need results—though their own interpretation of favorable results can differ from the interpretations imposed by donors. Thus the contingency model of accountability presented below reflects these differing demands.

The way in which NGO family networks differ within themselves is also part of this contingency model. Moreover, often national members work collaboratively within particular countries when working in, for example, Africa or Latin America. Thus, Save the Children UK could be working in Guatemala with Save the Children Sweden in Guatemala—and there is also a Save the Children Guatemala national member. Each of these SAVES has different Guatemalan partners and somewhat different objectives in their work (they are, after all, working collaboratively because of their special expertise on a part of the problem at hand). The modalities for holding them accountable within the United Kingdom, or Sweden, or Guatemala are in each case different, as are the kinds of requests made by contributors, donors, or boards to which they need to respond.

The kinds of processes and “products” that NGOs generate inevitably add to the complexity of their accountability systems. Unlike the private sector with its sharp focus on a single bottom line, NGOs, as Mark Moore has said, have two major bottom lines: their mission’s effectiveness and their financial sustainability. He points out that for nonprofits, “Mission attainment is calculated in terms that are different from revenue assurance. In this important sense, there are two bottom lines: mission effectiveness and financial sustainability.”¹⁸ He adds, rightfully, that this makes for greater managerial complexity than a single bottom line.

The Internet has both eased and complicated the kinds of reports and data that can be presented. Annual reports are often made available via the Internet, easing the mailing costs of getting these to donors. Some contributors can, and do, check Web sites to find out how projects and programs are progressing. For other contributors, mailings that are

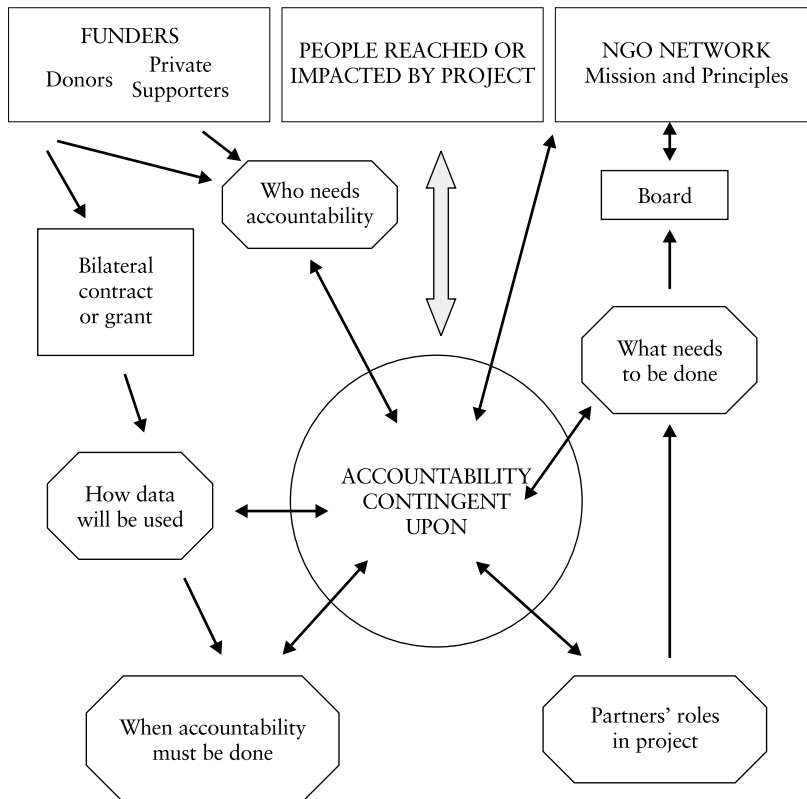
less formal than evaluation reports are needed. Because of changes in technology, the total amount of data produced and distributed has increased. The new modalities of communication are of course additional to all of those which senior decision-makers have traditionally used to meet accountability expectations, including meetings and consultations, discussions with key stakeholders, taskforces, and committees. Having an evaluation system, a staff tasked to perform that function, and budget resources allocated for evaluation are critical components—but they do not add up to the full range of accountability functions, and may even be a relatively small part of the whole.

In short, it is not easy to account to a diverse set of stakeholders who have different and sometimes conflicting demands. While Chambers is right to insist on pride of place in the accountability equation for the views of those on the ground who are directly affected by development projects and programs, they are in practice only *one* of the audiences. Figure 8.2 on the following page is one way to depict this contingency model.

As we have noted before, our key informants consistently cite responsibility to the people their organizations serve as a major priority. The input of those served—preferably obtained through a participatory evaluative process assessing whatever improvements have resulted in their livelihood or well-being—is the critical factor. While putting people first is normatively right, Uphoff reminds us that the fiduciary responsibility to donors cannot be slighted. Donors, on the other hand, tend to focus on individual project results, on their sectoral focus, and on their own policy programming priorities. What can get lost in the fray is the complexity of the needs of the communities in which the NGOs are working.

The contingent model of accountability responds to the fact of widely diverse and equally deserving constituencies and multiple modalities for responding to those constituencies. The managerial questions include the who, what, when, and how of giving an account to these dispersed and sometimes competing constituents. To the people in far-flung villages and communities—or refugee camps or settlements—for which they are working? To the foundations or corporations that contribute to their programs? To peer NGOs? To the bilateral aid agencies whose programs they are administering? To the international donors for whom they perform services? And, in the worse case, if it is all of the above, how is this accounting to be sequenced? Since the NGOs most often work through locally based partner agencies, dividing this challenge with those partners is another part of the equation.

Alan Fowler argues that the linear thinking—as exemplified by the logical framework for project design—imposes a presumed certainty that is not the reality at the village level. Like others, Fowler is concerned

Figure 8.2. Contingent Model of NGO Accountability

that the donor agency may force upon an NGO a narrow window through which to view its work. (Of course, lurking in the background for a donor like USAID is the U.S. Congress, which also will want quantifiable evidence of real results for monies disbursed—and Congress's timetable is governed by a short-term election cycle.) Suffice it to say that the emphasis on accountability to donors can lead NGOs to focus on their immediate projects without examining the broader economic, social, and political realities having an impact on communities. In relief work—where little is understood, and less is appreciated, about the relationships between relief and longer-term development—this problem is exacerbated. What will matter over time is not just how many blankets or food rations were distributed, but how community-level problem solving, for example, began to be reinstated.

All of these questions and criticisms have led to greater insistence on accountability without much clarity about what it is that should be assessed and which of the stakeholders need what information. The need

for accountability has been accompanied by calls for demonstrated evidence of effectiveness, or results. But ironically, an over-emphasis on quickly apparent “results” can and does undermine long-term sustainable development. In fact, it puts a premium on quick technical fixes, not on well-grounded, lasting, long-term consequences. To further unpack all the issues around impact and results means more focus on evaluation, and on organizational learning as processes to strengthen long-term effectiveness as well as accountability.

NGO leaders have always had to respond—account to—different constituencies. Donors, however, were always more than just another constituency—they were increasingly implementing programs through NGOs. Each donor has some kind of specific evaluation requirements and generally specified in the contract with the NGO how those requirements were to be met. Recall also that each of these NGO families include numerous national members and each member has in turn hundreds of operational partners in widely dispersed country settings. (Table 2.1 in Chapter 2 above reflects this pattern.) Moreover, projects are often multidonor-funded, and donors generally have different evaluation and accountability demands.

While much has been written about the possible threats to the long-term mission of NGOs when or if they become more dependent on donor funding, in general the assumption has been that increased donor funding would require more attention to evaluation than had been the case to date in many NGOs. One of the preliminary surprises our interviews surfaced, however, was that the NGOs that are the least dependent on donor funding are in fact the NGOs doing the most about evaluation. It is not yet clear how to explain this counter-intuitive finding. A part of the answer, however, may be that donors specify a final evaluation, sometimes hiring their own outside team to conduct it. That evaluation then is to meet donor needs, not the learning needs for the NGO. Often, for example, the NGO simply ensures that the evaluation is done, and does little more than treat evaluation as a part of contract compliance, rather than as part of their own ongoing reflection. While it is true that increased donor funding has called into question the role, function, and cost of evaluation, it has not yet become as robust in the process as might have been expected.

Evaluation Systems

While accountability is broader than evaluation, accountability drives NGOs to focus more attention on performance, and hence on strengthening their evaluation systems. Monitoring and evaluation processes are, after all, the ways that data are generated on program and project performance. Improving them depends on staffing and operational bud-

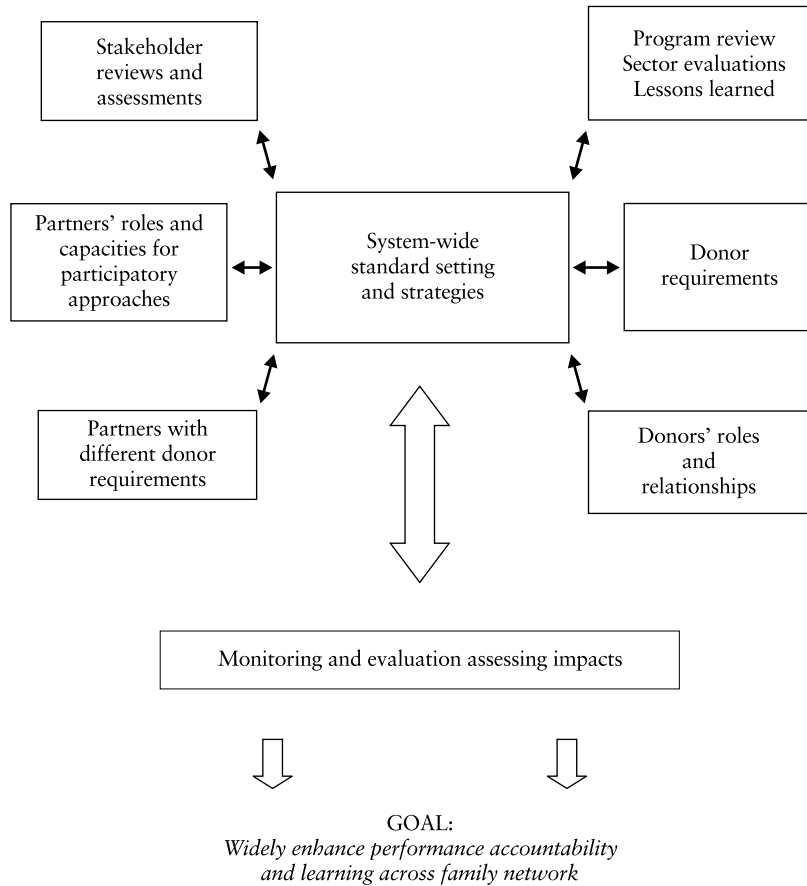
gets, so that strengthening either of these drives up overhead costs. Ironies abound: Donors want to work with NGOs because NGOs' voluntary character means they are less costly than consulting firms. Yet the earlier NGO tradition of voluntarism can be steadily eroded by the same set of actors and factors that led to its initial growth. Now NGOs are often implored to become more "professional." Most donors, however, show little willingness to pay for increased professionalism as it means staff and equipment costs, hence overhead. Support for overhead costs is not available. Professionalism in accountability is not as readily financed as it is demanded.

Monitoring and evaluation systems for large organizations take several different forms, and within the NGO families with which we are concerned, they vary within and among national members. Some are quasi-independent of line operations and report directly to boards. They may be so constituted in order to act as a check, or countervailing power to operations, or because a strong executive board wants the evaluation office to have a "watchdog" function.¹⁹ Other organizations—especially smaller ones on lean budgets—integrate monitoring and evaluation so that the two processes can be iterative, with mutual learning of "best practice" through workshops, seminars, conferences, or a wide variety of other means. The variables within monitoring and evaluation systems are depicted in Figure 8.3.

Evaluation, like any other function, requires organizational commitment of budget and staff to make it happen. Its advocates are invariably quick to note that making use of lessons learned will save costly error, and that investing in evaluation is therefore a sound investment. That may be, but the question remains of how to undertake evaluation of programs over time most efficiently as well as effectively. Portfolio reviews by regions, as CARE has done, make sense. But there are no easy answers to the questions of how much to invest in undertaking these works, at what intervals, or how often. Coupled with that, when or whether to share the findings from evaluations, and how to do so, is a troubled terrain. Predictably, there will be increased pressure from stakeholders for greater transparency. Transparency taken as a rule can put in place pressures to avoid sensitive areas, and some candor can be lost. Privacy rules and practices are culturally contextual, so that transparency practices in one place put burdens on partners in a different context. There are no right answers to these difficult trade-offs.

The evaluation offices in most Northern NGOs are small, spare in resources, and usually focused on setting guidelines and large parameters for the work that is to be done either by partners, or consultants, or both. Oxfam America and Save the Children US have one or two people fully committed to evaluation in their headquarters offices. These staff members are charged with evaluation responsibility where this means

Figure 8.3. Variables in Monitoring and Evaluation



setting guidelines, establishing policy frameworks, assisting with some training for regional offices, and supporting partners. While others may be tasked with aspects of this work, often there is no internal staff to do the work itself. Moreover, since most of these NGOs work with partners, they may ask a Southern partner to manage evaluation research, yet the partner may have even fewer trained staff to do so than the Northern partner. The large bulk of evaluation work that is done for NGOs is done by consultants contracted to undertake various assessments, studies, and impact evaluations. The costs—real and opportunity costs—for undertaking evaluations make them unwelcome demands on the organizational budget. Table 8.1 on the following page provides in the matrix an overview of the current state of these NGOs’ evaluation systems.

Table 8.1. NGO Evaluation Systems

<i>Organization</i>	<i>Past Practice</i>	<i>Current Practice</i>
CARE	Largely qualitative and descriptive evaluations done in field for individual projects.	Last five years standard setting for best practice as well as performance developed—now is system-wide.
<i>Médecins Sans Frontières</i>	Largely qualitative and descriptive except for quantitative health and medical information.	Done on an as-needed basis to drive internal program needs.
<i>Oxfam (UK)</i>	Combinations of approaches using participatory data collection. Careful attention to impact assessments.	Learning around impact assessments. Linked to Strategic Change Objectives at global level. Individual projects continue to respond to donor evaluation needs.
<i>PLAN International</i>	System-wide standards reflected in benchmark indicators.	Strong system-wide standards with agreed-upon benchmarks for PLAN priorities.
<i>Save the Children (UK)</i>	Variety of approaches. Ad-hoc, often driven by donor requirements. Reliance on initiative of project managers.	Serious attention to new indicators for each program area. Systematic, global approach linked to strategic planning.
<i>Save the Children (US)</i>	Donor-required evaluations routinely performed.	Donor-required evaluations done in field with different kinds of teams.
<i>World Vision</i>	Some evaluations done on an as-needed basis.	Increased interest in stronger, more systematic evaluations.

When NGOs work on government contracts, evaluation is generally pre-specified in the contract negotiations—often for a mid-course as well as a final evaluation. Consultants are contracted—either by the NGO, or by the donor to meet the contract compliance needs. Most often in contract work, the donor agency’s field mission will supervise the evaluation work and be the recipient of the final report (along with the partner field organization, or the field office of the NGO). The evaluation staff in the Northern NGO’s central office often do not see these

Table 8.1 (continued). NGO Evaluation Systems

<i>Drivers for Change</i>	<i>Evaluation Unit</i>	<i>Learning Process</i>
Largely internal; donor requirements had been met in previous system.	Yes. Director of monitoring and evaluation position created 1995.	Working towards having a learning system but not in place as an organizational process and system yet.
Largely internal; some external (e.g. UNHCR or ECHO requirements).	No. Staff across family network do evaluation based on knowledge and past experience.	Learning for those who most need the operationally relevant findings.
Internal and external (partners, stakeholders and donors).	Yes. Under Policy Department. Serves advisory and support role to regions where evaluations are done.	Evaluation results and indicators worked into strategic planning process.
Largely internal—strong corporate culture of accountability.	Yes. Evaluation unit with emphasis on system wide indicators. Also strong Auditing unit.	No. Process underway to get all members in PLAN into corporate-wide standards and measuring performance first.
Internal drive to improve on previous ad hoc system to facilitate organizational learning.	Yes. Under Policy and Research Department. Responsible for designing core indicators and supporting project managers in evaluation.	Indicators and change objectives are integral part of programming around core areas. Goal is to use evaluation for learning.
External-donor requirements.	No. A Director of Evaluation sets policy guidance on evaluation.	No. While interested in moving in this direction currently workloads preclude much time available for reflection.
Want donor pressure for more systematic data collection.	No. Evaluations done as driven by regional needs.	No. Other internal organizational restructuring kept this from being possible.

evaluations, precisely because they go to those most directly involved with having implemented the project at hand.

The constant search for improved ways of doing business, gathering and reflecting upon lessons learned, and measuring impact and consequences requires having people tasked to do that work. That in turn costs money and comes out of overhead. The challenge for NGOs, who often have to argue that as large a percentage of the funds contributed to them reach the grassroots as possible, have real difficulty financing

evaluation. Those who want greatest accountability—meaning narrowly that results are quantified and measurable—are not necessarily those private contributors writing their Christmas checks. Pointing to a stronger evaluation system is, however, not a widely favored way to increase an NGO's popular appeal. It increases the overhead, and annual appeals have to point to low, not rising, overheads.

Interestingly, in spite of all these hurdles, there is a great deal of work underway within most of these NGOs on strengthening their evaluation systems. Let us turn to some of the examples of the changes under way.

Program and Project Evaluation

While much is written about the shortcomings and the critical reviews of NGO projects, there is equally compelling evidence of many successes. Roger Riddell and Mark Robinson report on, among other things, a major review by the Overseas Development Institute in London of sixteen poverty-alleviating projects in Bangladesh, India, Uganda, and Zimbabwe, several of which were projects supported by Oxfam, Save the Children, and CARE. The aim of the study was “to formulate an approach for assessing not all projects and programs but, more narrowly, those whose purpose was to alleviate poverty and/or improve the living conditions of the beneficiaries, principally people living in rural areas.”²⁰ Almost all of the projects and programs reviewed were found to have improved living conditions, or to have raised the incomes of those living in poverty. Riddell and Robinson report on these projects in detail, including their immediate context, country context, and what was learned.

The study is especially noteworthy because it appears to be far less known in the United States than in Great Britain, while in the United States there has been more criticism of NGOs without in-depth empirical research on project and program impacts. But it is also true that evaluation techniques and interest in them spread more rapidly in the United States than in Europe. As Roger Riddell and Mark Robinson pointed out in 1995,

For most European NGOs working in the development field, evaluation is still very new and if used at all tends to be more of a one-off affair, most often embarked upon either because things have gone very wrong—the fire brigade approach—or when a particular project is completed but there is a request for future funding, or when a second or third phase of a particular project is to be launched. Indeed, the vast majority of projects and programs funded by British NGOs in developing countries are not subject to any sort of formal evaluation nor bound to specified cycles of

expenditure for committed support, as is common with official aid projects.²¹

Yet even as they wrote this in 1995, Oxfam GB, the largest NGO in Britain, had a separate unit for research and evaluation, though Riddell and Robinson added “but even here, no common framework, guidelines, or procedures have yet been adopted. The same is broadly true of current practice among even the largest NGOs in the Netherlands and Germany and among other Northern European NGOs, such as those in Finland and Sweden.”²²

While the absence of common guidelines might have been true in 1995, it was no longer true in 2000—either for Oxfam GB, or for Save the Children UK. Both of these NGOs, and others, have been rapidly building evaluation capacity and systems appropriate to their needs in the past several years.

Save the Children UK improved what it was getting from field-level evaluation by producing its own guide on how monitoring and evaluation might be done—a guide that was publicly available and entitled *Toolkits—A Practical Guide to Assessment, Monitoring, Review, and Evaluation*. Their point in this publication was to lay out to field staff and partners the tools for improving how they went about doing monitoring and evaluation.

Subsequently, in 1996–97, Save the Children UK collected 245 reports of SAVE reviews and evaluations, and, using a sample drawn from that database, summarized the recommendations looking at how institutional and conceptual approaches could be improved. Among the conclusions was that evaluations too often emphasized the achievement of outputs—numbers of children immunized, or wells put in place—but did not address the larger concern: how lives were improved and whether children were healthier. The report pointed out that: “Evaluation should be seen as part of a wider framework of lesson learning aimed at selecting, targeting, and implementing aid activities that will be most effective. This issue of going beyond counting outputs to address impacts, on how lives are improved is not a narrow issue but one that goes to the heart of institutional change.”²³ The author goes on to say, “Save the Children is already part of this process, developing methodologies for impact monitoring in the field. However this work is currently disparate and great attention needs to be given to integrating evaluation, and its component parts (assessment, planning, monitoring, review, formal evaluation) into program management, providing a strategy for continuous checking, learning, and adjusting within a wider institutional culture of reflective self-criticism.”

CARE USA has also made great strides in strengthening its monitoring and evaluation system since 1995. They too are fully seized of the

problem of assessing impacts—not just fulfillment of project activities. In their case, as there had also been a change internally to a different programming model—the Household Livelihood Security (HLS) system—they could then devise a monitoring and evaluation process keyed to HLS. By the end of 1999, they had developed CARE Impact Guidelines, with a menu of impact indicators for use in light of their goal of strengthening Household Livelihood Security. They are in the midst now of securing across all the CARE national members greater commitment to monitoring and evaluation work in light of these new standards.

Oxfam GB's Chris Roche has recently produced a new book, *Impact Assessment for Development Agencies*, that details with clarity how impact assessment might be done.²⁴ Oxfam has also published a guide, *Monitoring and Assessing Impacts*, that reflects Oxfam's internal change process around these same themes. Oxfam GB has housed its evaluation work in different units over time, but it consistently has made evaluation a core function. Currently it is putting into place a new department on policy planning and evaluation. The earlier, three-volume handbook, *Oxfam Handbook on Development and Relief*, includes several sections on evaluation. Noteworthy throughout these handbooks is their emphasis on participatory evaluation—ways in which those directly reached by a program are engaged in evaluating the program's effectiveness.

Participatory Evaluation

Where a Northern NGO has been working through and with a Southern NGO partner, or is working directly through their own local field office, it is likely that participants in a project can be identified to engage in a participatory evaluation process. This depends on the local partnering NGO's perspectives on evaluation, its evaluation capacity, and its willingness to comply readily with an externally hired evaluation consultant. The kinds of data that can emerge from careful participatory evaluations are particularly valuable. If the line of inquiry to be used is planned ahead of time, the kind of data gathered can help both the partner and the funding Northern NGO to learn in ways that may improve future effectiveness. Learning about what happened, how, why, and with what results provides information and insight that can help inform the next generation of projects or programs.

Participatory evaluation is logical and appropriate. The people who experience the impact of relief programs, or of development projects and programs, by definition have firsthand experience with the impact of those projects, and they should have much to say about effectiveness or impact. And, in general, development professionals undertaking evaluation research endeavor to reach and listen to those directly reached at the grassroots. There is a large and growing literature on how and why evaluation research must include this kind of qualitative empirical

work.²⁵ But there are severe constraints on ensuring that such work is carried out. It is costly, labor-intensive, and requires skill; it takes time—in several different locations. If it is not well done, the findings are not useful.

The larger the project, the more costly it is to sample and reach those directly affected. Project participants are difficult to track down, and baseline data is often missing: refugees often relocate. It is also true that participants directly impacted will not necessarily be able to provide data on aspects of their situation that are needed in order to put into context the data they do have. For example, they know their income level, but know little about the average income level when the project began and even less about the average for the region. The real strength of qualitative data comes in the insights into *why* something did or did not work. Its weakness comes in its not providing generalizable data. Respondents also may not identify long-term consequences (even for themselves) that the program has had. For example, in evaluating a food distribution program in a refugee camp, the refugees themselves will not have access to data on nutritional levels, food availability in different locations to appraise the areas of greatest need, or whether and when the percentage of those being fed are in reality terrorists who are “foraging” until the next battle call is sounded.

Evaluation of Emergency Relief Work

It is no surprise that we are most short of evaluation data on emergency relief work. The UNHCR, one of the main international agencies funding NGO work in relief, most generally requests a final financial audit but not an evaluation—in part because of the inherent difficulty of tracking those who benefited from emergency services. The spate of books criticizing NGOs for their roles in Rwanda grew out of non-sampled interviews, anecdotes, and observation drawn not from interviews with large numbers of refugees, but from interviews and recollections from external observers, agency staff, or journalists hazarding guesses about what happened. The more severe the emergency, the less likely it is that the internally displaced people reached through a relief effort are able to help document their perceptions of NGO effectiveness. Refugees or internally displaced people voice their issues, the media picks up anecdotes and rumor, and these are beamed rapidly to audiences in distant countries. Systematic evaluation from which cumulative learning can take place is not readily available and is expensive—and, by the time it is available, it is much less newsworthy.

The importance of strengthening evaluation systems has gained in salience as a result of the increasing pressure for demonstrated effectiveness. Evaluation in development projects and programming has become increasingly skillful in the past decade, but widespread adoption of the

cutting-edge techniques is still hindered by logistical and financial barriers as well as barriers stemming from the organizational culture and norms within the nonprofit community. There are significant overhead costs and organizational and staffing implications since evaluation research has grown in sophistication and technique in the past decade. Adoption of evaluation systems has also been affected by the fact that many of the techniques for measuring program effectiveness tend to contradict the inherent organizational culture within most NGOs.

Indicators and Benchmarking

Data is central to an evaluation system that allows for comparisons across similar kinds of programs or projects. NGOs, and other organizations, have found that the use of indicators, when these are carefully constructed, can be helpful. But getting those indicators, creating a culture of evaluation and learning, and then bringing staff fully on board across a system of national affiliates is an ongoing process—and not a short-term goal that, once achieved, stays in place. Nonetheless, PLAN International has been working on doing just that, and appears to be one of the leaders among Northern NGOs for its work on evaluation. It is now beginning to implement an evaluation system for their core areas of work: livelihood, habitat, and health. PLAN International went outside to get professional help to come up with indicators for each of the components of these core areas. Wherever it works, it gets baseline data. Of course its operations are long-term and integrated, and it largely operates through programs. Because PLAN works in communities for fifteen to twenty years, it can monitor those programs as well.

Performance indicators can be developed to distinguish between measuring inputs, outputs, results, and impacts. Indicator data collected over time provides information critical to learning where attention needs to go to further improvement. PLAN International's work in this area has attracted attention from other peer organizations; for example, NOVIB (the Dutch Oxfam affiliate) turned to PLAN in order to strengthen its own system. What is especially noteworthy about PLAN's approach is that it is a system-wide approach being used by all of PLAN's operational offices. PLAN's operational work is undertaken by Southern PLAN offices, with much more central coordination and quality control by PLAN International in support of those country offices. PLAN is, after all, more fully multinational than other Northern NGOs. It is globally structured and staffed with strong coordination across national boundaries. The PLAN UK office and PLAN US office exist for the purposes of fundraising only; when people within those offices refer to "operations," they mean not field-level operations themselves, but quality control, standard setting, and policymaking vis-à-vis field-level operations. Actual field-level operations in livelihood, habi-

tat, and health are wholly undertaken by Southern PLAN International offices and generally wholly by locally hired staff.

Save the Children UK is also in the process of developing a more systematic approach to monitoring and evaluation that will be consistent across its projects. This process builds upon their 1999 Strategic review, which prioritized Save's work into six core areas (health, education, social policy, food and nutrition, child labor, HIV/AIDS) and four cross-cutting themes (gender, emergencies, disabilities, advocacy). Preliminary indicators (mainly process-oriented) have been identified but much work remains to be done in obtaining baseline data, developing more impact indicators, and in integrating the new approach within country offices.

At CARE, the experience of introducing benchmarking and portfolio analysis approaches found resistance stemming from the strongly individualistic and independent style of staff and the service culture, which sometimes places a higher value on helping people than on doing things efficiently. Relief situations again accentuate the problem, since the urgent need for action takes precedence over detailed analysis and data collection. Furthermore, the very nature of the work that is being measured is much more ambiguous in relief and development organizations than in the for-profit world. The long-term contribution of any one development or relief effort to the long-term goal of improving people's livelihood is difficult to measure.

John Greensmith, international executive director of PLAN International, adds that the independence of staff impacts on an evaluation system. When he arrived at PLAN, he found the commitment of staff to their work far stronger than in the private sector. He noted, however, that the other side of that great commitment relates to the biggest challenge as well:

These people are very independent—they have to be to do what they are doing—taking more risks, and not your everyday conforming to expectations—and yet that independence means resistance to doing things with consistency across the organization, or being concerned with efficiency. Everyone wants to do their work their own way—and they find it very very hard to work on teams. We are still working on getting the incentives right—and unlike commercial life—pay incentives are not the answer—in fact sometimes people are insulted when that is offered. They demand inclusion, but then including them is not enough either, for that independence factor means that even after the decision is made, they will go off and do what they want to do anyway. The demand for transparency, efficiency, and accountability in our sector is steadily growing, but getting it to happen is challenging.²⁶

An in-depth study of CARE echoes Greensmith's observation on staff commitment and independence, and adds that "in spite of natural barriers, CARE staff were initially receptive to a program impact initiative because there was virtually no CARE-wide data on numbers of projects, beneficiary numbers, costs per participant, and there were few project baselines to compare project performance. CARE staff had a strong personal interest as highly driven professionals in knowing whether their projects were having an impact."²⁷

The strategic management process in short led to several things—attention to their mission, incentives, and attention to benchmarking and looking at impacts. It led initially to training—regional impact evaluation training. In addition the headquarters technical division was asked to create project data baselines by sector (water, agriculture, health, family planning, microenterprise development). The data were even to be collected on a form called the project implementation report. CARE headquarters asked all country offices to approve no new projects without baselines. In an effort to provide different approaches to best practice, three technical approaches were suggested—all of them shared the objective of comparing projects to some performance standard. Staff who believed they had unique projects could do self-comparison, in which staff set performance standards and then monitored for them. For others using a portfolio approach, it was possible to use indicators against national standards.

But CARE's extensive work on benchmarking methodologies had mixed results.

While top-down external rankings, which included good global best practice, were technically strong, they frequently backfired when angry project managers felt they were being ambushed. NGO participative culture made it hard to use external information in a non-threatening way. In a regional management conference in Asia, project staff rejected the top down approach . . . even when participative methods are used, it continues to be a challenge to get field staff to accept data from global empirical studies and broader evaluations. . . . What is more effective is an overall strategy which includes training in program design and evaluation, standards about project design baseline data, benchmarking, and participative evaluation.²⁸

CARE staff went on to develop its learning around the development of the household analysis tool as an example of reconfiguring internal programming approaches in order to improve effectiveness. The household analysis tool also helped reposition CARE's comparative advantage.

World Vision has recognized the need to place greater emphasis on

program quality and impact and has begun this process by sharpening the organization's core focus and priorities, developing clearer impact goals and indicators, and establishing new mechanisms for mutual accountability between national offices in its partners.

Oxfam GB has also undergone a process of internal review that has resulted in the definition of specific challenges and steps related to a range of issues. To address accountability issues, it will establish a quality assurance system involving the establishment of quality standards and a system of auditing, and it will develop a new approach to the ways in which it listens to and assimilates the views of its diverse stakeholders.

Given the difficulty of putting in place a system across all national members of an NGO family integrated evaluation process, there will always be questions whether this is the best investment in light of the relief and development NGOs' scarce resources and high opportunity costs. While the answer is contingent upon the context and situation of any particular NGO, there is a case for considering how staff are to learn over time, and how organizational learning is to be fostered. An old aphorism teaches us that "While the unexamined life is not worth leading, the overexamined life is not worth writing home about either." That is true of evaluation. While an evaluation system that feeds into and encourages organizational learning is eminently worthwhile, over-investing in it given the opportunity costs in a world where more work on poverty reduction and refugee relief cries for attention is not wise.

That said, most of the NGOs we have looked at have a way to go to improve the quality of their evaluation work. They need to do so in order to understand their own effectiveness. Linking that learning into organizational learning more generally would be meaningful for staff—and helpful in maintaining morale and averting "burnout" as well as increasing productivity. While almost all authors commenting on evaluation start from the assumption that with increased official assistance, organizations will perforce move toward strengthening evaluation, we found in our interviews that two organizations with strong evaluation systems, Oxfam and PLAN International, are not accepting significant amounts of official donor assistance. Oxfam America accepts no official development assistance. These agencies therefore moved in this direction in response to internal, rather than external, demands or needs. Further empirical research on the relationships between percentages of budget derived from official development assistance, concern with program impact, and organizational learning would be useful. The implications of organizational learning for staff productivity and morale are significant, and evaluation has a large role to play in a great learning system.

In short, there is a larger case to be made for creating more of an evaluation culture coupled with, and integral to, increased staff learning about effectiveness. An organizational learning process moves

toward re-establishing meaning for staff. Learning—especially when self-directed—is intrinsically invigorating. It is worth considering how this process works.

Organizational Learning

Organizational learning is one of the intriguing concepts currently receiving significant attention as well as scholarship. Peter Senge, director of the Center for Organizational Learning at MIT's Sloan School of Management, works with a large group of professionals on what has become the leading concept in much of the organizational theory and practice field.²⁹ The core concepts are focused on reinventing relationships, being loyal to the truth, developing strategies for personal mastery, building a shared vision, strategies for team learning and systems thinking, designing governing ideas, and treating organizations as communities. In brief, several schools of thought are at work—mixing and drawing upon the classical work of, for example, Chris Argyris, Russell Ackoff, and Jay Forrester, and incorporating material from, for example, strategic planning, quality management, and the emphasis on excellence. Thus in many ways this current model (and its practice) has long, strong roots and is not dismissable as another management “fad.” It is particularly appealing in the context of the development and relief NGOs, since it puts people at the center—thereby building upon the international development management tradition of people-centered development as fostered by David Korten, Louise White, and Robert Chambers.³⁰

The core elements of the organizational learning process are rooted in the field of organizational development. By engaging staff working in groups through queries that evoke reflection and analysis on their work, the participants begin to drive the agenda. Oxfam America has gone furthest with organizational learning. Their president, Ray Offenheiser, says that “organizational learning is driving the strategic management of our transformation process. We had to rethink our organizational model . . . the older traditional organization models are gone in light of global changes. Now our core currency is information and organizational learning is our over-arching principle.”³¹ Oxfam America began by working in groups to develop a strategic plan, and then turned to implementing that plan. To lead off the implementation, they had a week-long workshop at the Goree Institute in Senegal bringing together partners, regional representatives, regional managers, and senior managers (including Ray Offenheiser) to discuss implementation of the theme, “Participation for Equity.” It was one of the first times that a large number of managers, staff, and partners had talked with one another systematically about their work. Those who participated conveyed

their excitement and commitment to building upon and deepening the process. The Goree week produced a series of guidelines that were then used to guide the management of their strategic planning process. These are still being used: speak with authority and substance on key development issues; invest in knowledge for action; link the local with the global; program outcomes lead to social change; and serve partners and work toward their empowerment.

The core elements in this process are that it focuses on being a learning organization in practice—by breaking down the boundaries between departments, and between center and field and partners; puts people at the center of the organization; and flattens organizational structure. It also—and predominantly from the perspective of operations—creates permeability between planning and taking action. No longer is there a separation between those who plan and those who implement; these are seen to be, and they become, interchangeable.

Flattening the organizational structure, however, proved to be one of the difficult parts of the process. All the Oxfams are unionized and Oxfam America is no exception. The union (Service Employees International Union) has detailed rules about structure and these precluded giving staff supervisory responsibility. How then could they get to more movement between those who plan and those who implement—a more horizontal organization? Again, this was worked through bit and piece by managers and staff working in groups, devising ways to proceed and yet be in conformance with the union rules. As many union leaders and members experienced organizational learning as empowering, they basically worked their way through a thicket of problems. Some of the old school union leaders—more accustomed to the fist fight model of adversarial relationships—had trouble making the change; eventually they lost out in union elections.

Oxfam America drew heavily upon training in what has come to be called interspace bargaining—an approach in which small groups work through competing interests rather than posturing for positions in an argument. This training proved invaluable; staff now knew how to do interest-based negotiating. As the time was approaching for a renewal of the union's contract, this proved important. Instead of drawing upon the older adversarial approaches to union contract issues, groups of staff worked through what needed to be done and drafted side letters that reflected the agreements reached. Then during the contract negotiations, these side letter agreements were incorporated into the contract. Now Oxfam America has moved on yet again, moving away from the big bang approach to five-year planning to a more flexible and responsive planning. Organizational learning is firmly driving their planning process, and bringing transformation with it.

There is no doubt that, from the point of view of achieving results,

NGOs need an iterative process of engaging staff from top to bottom in identifying and illuminating what is and is not working. It is very easy to be caught up in processes and meeting deadlines, fundraising, disseminating information, and measuring those “inputs” without getting time to reflect and think about the impact or results of this work. Laura Roper, Oxfam America’s director of program planning and learning—the office where setting guidelines and policy on evaluation takes place—told us that there is a real need “to create space and opportunity for staff to reflect.” Reflection, especially when informed by data on results or consequences of actions taken to date, leads naturally to learning, or at least identifying what remains to be done. Roper has also recently contracted for assistance to get more quantitative data on program results because she knows that Oxfam’s qualitative data from these various processes would be better informed with more quantitative measurement of impact and results.

Ethnology and Beyond

Throughout this book, we have ended each chapter with a step back and a look at the same more probing questions: Can the views of our interlocutors be supported by empirical data? What motivations, as well as nuances and details, may our respondents not have explicitly discussed? And what emerging issues or trends need further exploration?

Empirical Data

The leaders with whom we met are committed to accountability and in many instances also have worked to strengthen monitoring and evaluation systems. Few, however, consider these functions their highest priority. For all, achieving their organization’s mission is the uppermost goal; this is what drives them to work on advocacy, on fundraising, and, above all, on programs. Yet they are fully aware that improving performance requires knowing more about that performance—and hence accountability, and its components, monitoring and evaluation systems are strong intermediate goals.

We must acknowledge that our data on accountability systems is incomplete. The variations in accountability systems among national members of all the six NGO families in our core group of interlocutors militate against our having a complete data set on accountability. What we were able to learn is that these variations are real and pervasive. Our timetable did not permit collecting everything that needed to be known within any given family—let alone account for the differences across the different networks. But what we did learn is that there have been significant and serious efforts to improve evaluation systems in the past

five to eight years, and that new approaches and learning from them are readily shared across NGO family borders.

Motivations Not Discussed

The motivations most likely to be left undisclosed in regard to accountability are those that are rooted in serious differences of perspective within any given NGO family on how to manage the contingencies in accountability. These NGOs' separate country political contexts vary too much for there to be ease with a system-wide approach. For example, Oxfam GB is large enough within the political system of Great Britain that the head of Oxfam GB has ready access to the Prime Minister. Moreover, since a significant percentage of the British electorate uses a payroll deduction system to contribute to Oxfam, most parliamentarians will listen when Oxfam speaks. That is in marked contrast with the situation of Oxfam America—or of any other NGO based in the United States.

As different national members have different constituencies to whom they respond, as well as an array of domestic political and legal forces within the country where they are headquartered, their accountability systems differ more than any other function. While we could, and did, learn about such differences within NGO families, there was no opportunity to explore the myriad implications of those differences for internal management of the accountability process. There are also competitive pressures among all the NGO families. The Internet, for example, makes it far simpler for prospective donors to compare and contrast the information on Web sites when trying to decide about an annual gift. Hence the Internet has inevitably brought some competitive forces to the fore among these NGO groupings. For example, an innovation on one Web site that, for example, highlights significant results achieved puts pressure on others to demonstrate comparable success as well. Leaders were not eager to talk about tensions they are experiencing either within their own NGO families, or across the spectrum with others. Moreover, the very ethos of their missions puts a premium on social service and outreach—which carries over a values-centered premium for being, or at least appearing to be, cooperative rather than competitive like the private sector.

Further Nuances and Details

In researching the NGOs on the topics covered in this chapter—more than on any other—it was striking to see how little is actually known about monitoring and evaluation systems in NGOs. Most of the literature takes a very broad-brush treatment of the accountability factors. One of the results of that is a paucity of hard data on exactly what the evaluation systems are within these NGOs. One aspect of that is that the

NGOs' leaders also are not very likely to go into details. When probing for this information, one is sent on to others—even when it is said that accountability is a senior management goal.

Underexplored Trends and Issues

Of the several trends on the theme of this chapter that warrant further exploration, the one that commands the most attention is the growing interest in understanding the longer-term impact, not just the outcomes and outputs, of projects and programs. In our interviews, staff expressed a keen interest in that, and in having more time for reflection on lessons learned. At the center of this issue is a conundrum: both outsiders and staff would like to know more about when and why NGOs are effective (as many staff believe they are). Replicating that success becomes more possible with such knowledge. Development projects come with hidden surprises and unexpected outcomes, and sometimes beneficiaries most value outcomes that were not planned but just happened. Still, more learning is both possible and wanted. Given the scale of the human needs with which these NGOs are struggling, this search for explanations of effectiveness needs further exploration.

It is also worth examining why the pressure for such learning sometimes comes more from the field or the staff than from headquarters. It may be that those most immediately engaged in the work are most keen to discover, and document, when they have had an impact. Senior managers, however, whose workloads pull them in different directions, are inevitably focused on the big picture.

Notes

1. Roger Riddell and Mark Robinson, *Non-Governmental Organizations and Rural Poverty Alleviation* (London: Clarendon Press, 1995).
2. Michael Edwards and David Hulme, *Beyond the Magic Bullet: NGO Performance and Accountability in the Post-Cold War World* (West Hartford, Conn.: Kumarian Press, 1996).
3. Jonathan Fox and L. David Brown, *The Struggle for Accountability* (Cambridge, Mass.: MIT Press, 1998).
4. Ian Smillie and Henny Helmich in collaboration with Tony German and Judith Randell, *Stakeholders; Government-NGO Partnerships for Development* (OECD and Earthscan Publications, 1999), 9. With worries about these relationships, Edwards and Hulme followed up their first book with one entitled, *NGOs, State and Donors Too Close for Comfort?* (London: St. Martin's Press, 1997).
5. Smillie and Helmich, *Stakeholders*, 16.
6. Norman Uphoff has cogently argued that NGOs should not be considered a "Third Sector" because they are neither people's associations nor membership groups. He goes on to discuss the implications of this for accountability. See

Norman Uphoff, "Why NGOs Are Not a Third Sector," in Edwards and Hulme, *Beyond the Magic Bullet* (West Hartford, Conn.: Kumarian Press), 23–39.

7. Uphoff, "Why NGOs Are Not a Third Sector," in Edwards and Hulme, *Beyond the Magic Bullet*.

8. One of the first major official development sources on these issues was the World Bank's policy paper on governance, *Governance and Development* (Washington, D.C.: World Bank, 1992). This paper argued for three concepts as central to good governance: accountability, transparency, and predictability (rule of law). These were interrelated, but separable, as it is possible for an organization or government to perform one without the other. For example, they could share information, provide data, and release documents without there being a process by which that organization's leaders are held to account for their decisions.

9. These participants held this discussion in the Global Leaders' Meeting in Seattle, Wash., Session 1, May 7, 2000, meeting session notes.

10. Robert Chambers, *Whose Reality Counts?* (Brighton, England: University of Sussex, 1997).

11. Amartya K. Sen, *Development as Freedom* (New York: Alfred Knopf, 2000).

12. Edwards and Hulme, *Beyond the Magic Bullet*, 11.

13. Brian Smith, *More Than Altruism* (Princeton, N.J.: Princeton University Press, 1990).

14. John Gaventa, "Poverty, Participation and Social Exclusion in North and South," in *Poverty and Social Exclusion in North and South*, ed. Arjan de Haan and Simon Maxwell (*IDS Bulletin* 29, no. 1, January 1998), 50–57.

15. Alan Fowler, "Assessing NGO Performance, Difficulties, Dilemmas, and a Way Ahead," in Michael Edwards and David Hulme, *Beyond the Magic Bullet*, 169–86.

16. Peter Uvin, *Aiding Violence: The Development Enterprise in Rwanda* (West Hartford, Conn.: Kumarian Press, 1998). It is worth noting, however, that the large four-volume study by John Eriksson with contributions from Howard Adelman, John Borton, Hanne Christensen, Krishna Kumar, Astri Suhrke, David Tardif-Douglin, Stein Villumstad, and Lennart Wohlgemuth was one of the first places to discuss the problem of structural violence. See Chapter 2, *Early Warning and Conflict Management*, in *The International Response to Conflict and Genocide: Lessons from the Rwanda Experience*, ed. John Eriksson (New York: Steering Committee of the Joint Evaluation on Emergency Assistance to Rwanda, 1996), vol. 1, 17–24.

17. At a meeting at the Brookings Institution, October 29, 1999, with L. David Brown, Julie Fisher, Peter Hall, Virginia Hogkinson, Ray Horten, Steve Smith, Christina Kappaz, and Crispin Gregoire, we discussed, among other issues, the accountability dilemmas, and this model was suggested. It should also be noted that Alan Fowler uses the same term in his outstanding article, "Assessing NGO Performance, Difficulties, Dilemmas, and a Way Ahead," in *Beyond the Magic Bullet*, 169–86.

18. Mark Moore, "Managing for Value: Organizational Strategy in For-

Profit, Nonprofit, and Governmental Organizations,” in *Nonprofit and Voluntary Sector Quarterly* 29, no. 1 (Supplement 2000): 194.

19. The World Bank, for example, has the evaluation function housed in the Operations Evaluation Department (OED) for this reason. OED reports directly to the Board of the Bank and as such has complete independence from control by the senior managers within Bank operational departments. Task managers of projects, however, may also build project evaluation into their project budget and then design the manner in which that is to happen. That would not, however, preempt any independent evaluation by OED. The United States Agency for International Development (USAID), on the other hand, generally houses monitoring and evaluation within a center or separate department, sometimes adding the task of learning as part of that function. USAID generally builds linkages between operational work and the monitoring and evaluation work. There are a wide variety of ways that this function can be organized, each with advantages and disadvantages.

20. Riddell and Robinson, *Non-Governmental Organizations and Rural Poverty Alleviation*, 45.

21. Riddell and Robinson, *Non-Governmental Organizations and Rural Poverty Alleviation*, 44.

22. Riddell and Robinson, *Non-Governmental Organizations and Rural Poverty Alleviation*, 45

23. David Mansfield, “Evaluation: Tried and Tested? A Review of Save the Children Evaluation Reports,” Working Paper No. 17 (London: Save the Children UK, 1997), 9.

24. Chris Roche, *Impact Assessment for Development Agencies* (Oxford: Oxfam and NOVIB Publications, 1999).

25. Participatory evaluation was first promoted in Asia and Africa where it was also often linked with giving people a voice in influencing future policy choices. There was also a Society for Participatory Research in Asia that worked in conjunction with the African Adult Education Association. See Yusuf Kassam and Kemal Mustafa, *Participatory Research: An Emerging Alternative Methodology in Social Science Research* (Nairobi: African Adult Education Society Publication, 1982). That approach spread rapidly to Northern industrial countries, and is reflected in much current material on evaluation. See, for example, Edward T. Jackson and Yusuf Kassam, *Knowledge Shared: Participatory Evaluation in Development Cooperation* (West Hartford, Conn.: Kumarian Press, 1998).

26. John Greensmith, international executive director of PLAN International, telephone interview by Coralie Bryant, June 24, 1999.

27. Marc Lindenberg, “Are We at the Cutting Edge, the Blunt Edge, or the Wrong Edge?” *Nonprofit Leadership and Management* 11, no. 3 (March 2001). See also Christine Letts, William P. Ryan, and Allen Grossman, *High Performance Nonprofit Organizations* (New York: Wiley, 1999), see chapters 1 and 5.

28. Lindenberg, “Are We at the Cutting Edge?” 16.

29. Peter M. Senge, *The Fifth Discipline: The Art and Practice of the Learning Organization* (New York: Doubleday, 1990). This was followed up by Peter

Senge et al., *The Fifth Discipline Fieldbook: Strategies and Tools for Building a Learning Organization* (New York: Doubleday, 1994).

30. See for example David Korten, "Community Organization and Rural Development: A Learning Process Approach," *Public Administration Review*, September/October 1980: 480–511. *People Centered Development*, ed. David Korten and Rudi Klaus (West Hartford, Conn.: Kumarian Press, 1984); and Coralie Bryant and Louise G. White, *Managing Development in the Third World* (Boulder, Colo.: Lynne Rienner Publishers, 1982), and Louise G. White, *Creating Opportunities for Change* (Boulder, Colo.: Lynne Rienner Publishers, 1987).

31. Raymond Offenheiser, president, Oxfam America, telephone interview by Coralie Bryant, January 25, 2001.