EXPECTATIONS AND CAPACITY OF PERFORMANCE MEASUREMENT IN NGOs IN THE DEVELOPMENT CONTEXT

KATHRYN NEWCOMER1*, LAILA EL BARADEI2 AND SANDRA GARCIA3

1George Washington University, USA
2American University in Cairo, Egypt
3Universidad de los Andes, Colombia

SUMMARY

There are many obstacles to promoting learning as an outcome of performance measurement in non-governmental organizations (NGO) social service providers, especially in less developed countries. Building upon a conceptualization of accountability as a multifaceted set of relationships through which funders, or principals, and non-profit providers, or agents, jointly shape organizational learning, and performance, this study expands our understanding of how accountability mechanisms affect learning within service providers. This paper explores the role that funders play in shaping performance measurement, or monitoring, practices within NGOs serving disadvantaged children in developing countries. We examined the experience of service providers in Egypt and Colombia to assess how the barriers to use of performance data and learning may be addressed. We conducted interviews using the same protocols with program managers in six non-profit providers in each country that provide services to children, and we also interviewed major donors in the arena of children’s services in the USA. We probed the NGO managers’ experiences with performance measurement to identify obstacles and potential solutions to improve the use of the data to promote learning. Our findings support previous research about the potential for upward accountability mechanisms to influence internal learning. We suggest that funders should be held accountable for how the incentives and disincentives they provide to grantees affect their internal learning about how to improve their services. This notion of ‘reverse accountability’ means that funders need to be strategic and intentional when they design reporting mechanisms that affect the learning behaviours within their grantees. In line with our call for reverse accountability, we offer a model demonstrating our notion of the two-way flow of accountability and we offer recommendations to help improve the performance reporting environment for NGOs who are addressing complex problems with less than adequate capacity. Copyright © 2012 John Wiley & Sons, Ltd.

KEY WORDS—performance measurement; accountability; international development; evaluation

Over the last four decades, the demand for systematic data about the effectiveness of public and non-profit programs has increased because of the non-coincidental convergence of actions by a number of institutions with stakes in service delivery. In countries across the world, governments at all levels, foundations, funders of international development endeavours such as the World Bank and the US Agency for International Development (AID), the United Way of America in the USA, think tanks, and academics with interests in improving the delivery of services have all contributed to an increasing focus on measurement of programmatic ‘results’ as part of the exercise of accountability for the funds (Carman, 2009; Dubnick and Frederickson, 2011a, 2011b; Imas and Rist, 2009; Moynihan, 2008; Hatry, 2008; and Newcomer, 2008). Over this period, both the level and sophistication of the dialogue about the merit of measuring and reporting about the operations and results of social services have notably increased (Stone and Cutcher-Gershefeld, 2001; Schalcon and Ronham, 2003; Sowa et al., 2004; Baruch and Ramalho, 2006; Martin and Kettner, 2009; Packard, 2010). However, questions remain about whether and how performance data are used internally to improve services and/or to facilitate learning within service providers (Moynihan, 2009; Alexander et al., 2010; Moynihan, and Pandey, 2010; Moynihan et al., 2011; Metzenbaum, 2006; Ebrahim, 2002, 2005, 2010).

*Correspondence to: K. Newcomer, George Washington University, The Trachtenberg School of Public Policy and Public Administration, Washington, DC, USA. E-mail: newcomer@gwu.edu

Copyright © 2012 John Wiley & Sons, Ltd.
Monitoring programmatic performance in agencies that provide social services is a virtually ubiquitous activity in public agencies and non-governmental organizations (NGOs) across the world. However, the decisions and activities involved in measuring performance typically produce tensions and some discomfort for the service providers. Why? Performance is itself a very amorphous concept open to a multitude of operational definitions and may be measured and interpreted in a variety of ways by different stakeholders (Moynihan, 2008; Newcomer, 1997, 2008). Monitoring programmatic performance entails systematic analysis and measurement of program operations and results, and this can present risks for the staff and other stakeholders who have invested themselves in the program. The risks of looking bad, losing public confidence and/or funding, or being forced to change are typically not welcome.

In addition, for many providers, resources to support measurement and reporting on their operations and results may compete with resources to serve their clients. Despite the uncertainties and challenges service providers face, the current demands placed on them to measure performance outweigh the psychic and monetary costs of measurement. Both routine monitoring data and evaluation studies undertaken can be useful in both day-to-day management and the exercise of accountability as long as the performance data and the study results are deemed both credible and useful. However, learning how to ensure that data are credible and useful and learning how to analyse the data to promote learning take time. The US Government Accountability Office in the USA recently highlighted the need to provide training on how to interpret and use performance data to improve programs in the federal government. GAO noted that even after 10 years of experience with widespread performance measurement as required by the Government Performance and Results Act, analysis and interpretation of the data to promote learning is challenging and requires targeted training (GAO, 2008).

The current environment in which social service providers measure their performance in more developed countries is fairly daunting. The demand for credible evidence from public agencies and NGOs continues to exceed capacity to produce it, especially among the less wealthy providers. Despite extremely burdensome performance data collection and reporting requirements placed on providers, there is little documented use of the data by either providers or funders. Scepticism persists that monitoring and evaluation (M&E) actually improves public or non-profit services. And there is typically spotty evidence of the results of performance measurement and evaluation studies affecting budgeting and funding decisions within governments or other funders (Dubnick, 2005; Metzenbaum, 2006; Newcomer, 2007a, 2007b, 2008; Radin, 2009).

Providers in the international development context operate with the same sorts of expectations and challenges of their counterparts in developed countries, but the capacity to analyse and learn from the data is usually even lower than in the more developed countries (for example, see Edwards and Hulme, 1995; Imas and Rist, 2009). The expectations that funders hold about the data that can be produced by the social service providers they support, as well as the specific reporting requirements they impose, are especially critical. Funders have the power and the means to shape the learning value to be derived from monitoring and evaluating their grantees through the exercise of upward accountability (Edwards and Hulme, 1995; Ebrahim, 2002, 2003, 2005, 2010; Benjamin, 2008). As Teresa Behrens and Thomas Kelly wrote in 2008, ‘funders can and should hold themselves accountable for creating positive social change and learning’ through the way they frame evaluation for their grantees (p. 48).

International funders are looking for ‘evidence’ of what strategies work and are paying for external evaluators to conduct impact studies to collect data to shed light on solutions for complex social problems (for example, see Das Gupta et al., 2005; Hossain et al., 2005; Vio, 2005; Armecin et al., 2006; Schady and Caridad Araujo, 2006; Garcia-Moreno and Tines, 2007; Paxson and Schady, 2007; Schneider et al., 2008; Soeung et al., 2008; Waiswa et al., 2008; Khagram et al., 2009). The funders are also asking the providers to collect data on a routine basis to track the use of funds and the progress made by the people and communities served (for example, see Edwards and Hulme, 1995; Manning, 2007; Knutsen and Brower, 2010; Mumsen et al., 2010).

Professionals involved in the evaluation of public and non-profit programs advocate that both monitoring and evaluation can and should be employed to promote learning among the providers and the community about what works, as well as to provide a means of holding providers accountable for the expenditure of donors’ funds (for example, see Ospina et al., 2002; Metzenbaum, 2006; Hatry, 2008; Imas and Rist, 2009; Wholey et al., 2010;
Newcomer and Caudle, 2011). The manner in which funders impose reporting requirements can certainly facilitate financial accountability, but it may also promote learning—if framed appropriately (Ebrahim, 2005, 2010).

Following Alnoor Ebrahim’s conceptualization of accountability as a multifaceted set of relationships through which funders, or principals, and non-profit providers, or agents, jointly shape performance, this study expands upon Ebrahim’s conceptual framework relating learning and accountability stemming from the use of evaluation, more narrowly focused upon here as performance assessment within service providers. We adopt Ebrahim’s definition of accountability as ‘the means through which individuals and organizations are held externally to account for their actions and as the means by which they take internal responsibility for continuously shaping and scrutinizing organizational mission, goals and performance’ (Ebrahim, 2003, page 194). We embrace Ebrahim’s view of accountability as relational in nature, both internal and external, and highly complex, especially when the non-profit service providers are involved in reporting upward to multiple donors—both locally and internationally based.

Ebrahim recently offered a framework that is useful for distinguishing among the processes and responses to accountability processes. He offers a typology of five accountability mechanisms: Disclosures/Reports, Evaluation and Performance Assessment, Self Regulation, Participation and Adaptive Learning. For each of the mechanisms, he suggested that inquiry should examine to whom accountability is demonstrated (upward, downward or internal), for what accountability is shown (finance, governance, performance or mission), what sort of inducements are offered (internal or external) and the organization’s responses (compliance or strategic) (Ebrahim, 2010). In this study, we focused on both Evaluation/Performance Assessment and Adaptive Learning and, more importantly, the relationship between the two.

The notion that accountability is exerted both externally to funders and internally to the non-profit staff and boards because of their shared sense of ‘felt responsibility’ for producing desired outcomes for the children and families they serve is applied in this study as well. And in line with Benjamin’s notion of an account space framework, the manner in which funders’ performance reporting requirements shape providers’ practice is deemed important and examined, but the focus here is upon internal learning, not negotiations between the funders and providers, as was the case with Benjamin (2008).

This paper explores the role that funders play in shaping performance measurement, or monitoring, and the practices and use of the data within NGOs serving disadvantaged children in developing countries. We conducted our research with service providers in Egypt and Colombia and with donors focused upon children in Washington, DC. Although the two countries are not deemed representative of the entire developing world, they provide useful test sites for exploring the reality on the ground in order to consider how funders might promote learning among service providers, and thus improve services to the least advantaged in our societies. Important, two of the authors are faculty members with extensive experience studying foreign funding of NGOs offering social services in Cairo and Bogotá, who teach in universities located in the two cities. And in fact, although from two different continents, on two sides of the globe, for maximum variation, the slum areas selected in both countries suffer from quite similar complications related to poverty and human deprivation.

In both Egypt and Colombia, it is difficult to quantify the number of NGOs that provide services for children because the available administrative information relates to the type of services that organizations provide rather than the beneficiary population they serve. In both countries, the main activities of registered NGOs have to do with human development, poverty alleviation and the provision of health and education services. In the case of Colombia, there is also a growing number of NGOs that are devoted to human rights issues and services provision to victims of the armed conflict. However, for the purposes of this study, the main criteria selection was the provision of services to children. In both countries, NGOs are geographically distributed all over the national territories; however, they are more likely to be based in urban rather than rural areas.

We examined the experience of service providers in Egypt and Colombia to assess how the barriers to use of performance data and internal learning may be addressed. Our study addressed four primary questions:

1. What is current practice in performance reporting?
2. What are the primary challenges that providers face in measuring performance?
3. What is the current practice in use of performance data by the service providers?
4. What sorts of changes do service providers recommend to increase the usefulness and internal learning supported by performance measurement?

In this paper, we first briefly describe the scope and methods of our research. Second, we describe the NGOs we visited in Colombia and Egypt. Third, we present our findings. And finally, we offer conclusions and recommendations about the role that funders might play in promoting learning within NGOs that deliver social services.

SCOPE AND METHODS

Our interest was to explore performance measurement, or monitoring, practices within NGOs providing social services in developing countries. Given that we intended to visit NGOs in two countries, we chose to delimit the study to NGOs providing one set of services. We narrowed the type of services to services for disadvantaged children.

We conducted case studies within six NGOs in each country in late 2010. We developed a set interview protocol, which we employed to interview managers in NGOs in Colombia and Egypt. To help us understand the complex flow of funds to the providers for background, we conducted interviews with DC-based staff in two large funders of international NGOs offering services to children that are based in the USA—UNICEF and Save the Children. The reason for selecting these two donors was because of their focus on funding NGOs that serve children and their broad international reach in this social service arena. For the UNICEF and Save the Children interviews, two staff members in each agency’s DC evaluation offices were interviewed in late 2010, and the documentation on monitoring each agency provides was reviewed as background for this study.

In Colombia, in-depth structured interviews following the protocol were conducted on site in Spanish at six NGOs that provide services to children and receive funding from international donors. Similarly, in Egypt, in-depth structured interviews following the protocol were conducted on site in Arabic at six NGOs that provide services to children and receive funding from international donors.

The interviews in both countries were conducted with executive managers in each of the 12 organizations on site, and each interview lasted approximately 2 to 3 h. During each site visit, the authors met with several of the managerial staff, but the top manager, that is, CEO, was the official interviewee for each NGO. Each of the interviewees had a lot to say regarding their work and its complexity and shared many examples, incidents and anecdotes regarding how they dealt with the disadvantaged children and their families. The interviewees’ responses were perceived by the researchers to be candid because they were all assured that the purpose of the study was purely academic and that researchers’ objective was not to evaluate or criticize them in any way but to find ways to help them do their job better.

Despite the interesting caveats provided by Weick’s work on organizational ‘sense making’ and that managers’ perceptions of their work may not be always accurate (2005), in the majority of the interviews, the managers were keen on sharing their problems and challenges with performance measurement. They did not seem to be defensive, nor trying to prove their competency in applying the different tools, as much as they were open to learning and receiving guidance. Perhaps, their ‘sense making’ was apparent in their descriptions of how they perceived the efficacy of their services as related to children and the impact of the surrounding environment on their work, but not as much in the descriptions of the performance measurement processes they used.

Additional background materials on the NGOs were obtained during the site visits. For example, the staff showed the authors performance data printouts, reports and other written materials, as well as files and portfolios maintained on the children served. The websites of all of the NGOs were all examined as well.

The selection of the NGOs in both countries started out by purposively contacting NGOs that met our two key criteria, of offering services to marginalized, disadvantaged children in slum areas, and of being known for receiving funding from more than one donor. Additional NGOs were selected through referrals, that is, each NGO visited was asked to identify other NGOs who were also serving children and receiving funding from multiple sources (in either Cairo or Bogotá).
A PROFILE OF THE NGOs INCLUDED IN THIS STUDY

Colombia

All six NGOs selected work in development and are non-profit social services providers, and among their activities are programs targeting vulnerable children. Beneficiary children from these organizations include children from very poor families, children from internally displaced families (victims of Colombian conflict) and children experiencing abuse and neglect. Five of the six NGOs operate in Bogotá and other cities in Colombia, and one of the NGOs operates in Cali.

Egypt

All six Egyptian NGOs included in this study work in development and are non-profit social services providers, and among their activities are programs targeting children at risk. Under the category ‘children at risk’, the Egyptian NGOs serve homeless children, or as more commonly known in Egypt ‘street’ children, child labour, abused children, children with disabilities and children living in deprivation and in extreme poverty conditions. All six Egyptian NGOs studied operate in slum areas in Cairo in the districts of Ein El Sira, Old Cairo and Mokattam Hills. A profile of the NGOs included in this study is presented in Table 1.

FINDINGS ON CURRENT PRACTICE IN DATA COLLECTION AND PERFORMANCE MEASUREMENT

Colombia

Organizations vary in the way they collect data for performance measurement. Some organizations contract out surveys that are conducted at the household/children level, whereas others use their own staff or field workers (teachers, social workers, etc.) to collect the data, and others even use the beneficiaries themselves to collect the data (for example, one organization asks the leaders of the community to fill out the instruments, and another organization sends an instrument to parents and teachers who have to fill it out and then compare).

For most of the organizations, data collection is a cumbersome process, especially when staff members are responsible for data collection. Usually, field workers are very busy with the day-to-day provision of services, and they perceive measurement as an additional task that might distract them from their main purpose.

With regard to data accuracy and verification, only one NGO reported a specific strategy to check on the completeness and accuracy of the data. They have site visits (usually unannounced) where data on infrastructure, quantity and quality of the services are checked (Table 2).

Egypt

The Egyptian NGOs reported using an assortment of methods at various stages of data collection, including surveys, follow-up phone calls, focus groups, individual interviews and evaluation sheets. There was variety in the studied NGOs regarding the responsibility for performance data collection. It was either the field workers who collected the required data in cooperation with the community leaders or the ‘other designated staff’ from within the NGOs, typically referred to as specialists, program coordinators, M&E officers or program managers. In the case of the umbrella organization, responsibility for data collection was delegated to the community-based organizations (CBOs).

The majority of NGOs did not voice any serious complaints regarding data collection processes as they accepted the fact that it had to be carried out, especially when it is a donor requirement. There was the occasional complaint by the NGOs involved in art education that they have a problem with the art animators who do not approve of keeping records and who perceive themselves as artists and do not want to do anything with paper work. Additionally, there were calls for more simplicity in performance measurement systems used and a request for continuous development of NGOs capacity in M&E in general.

With regard to data accuracy and verification, Egyptian NGOs mentioned a variety of strategies employed to check on the completeness and accuracy of data collected. One NGO mentioned that they organize field visits
Table 1. A profile of the NGOs in the study

<table>
<thead>
<tr>
<th>NGO name</th>
<th>Number of children served</th>
<th>Cities</th>
<th>Beneficiaries</th>
<th>Main activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colombia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRAN</td>
<td>310</td>
<td>Bogotá</td>
<td>Vulnerable children (abused or neglected)</td>
<td>Temporary child protection</td>
</tr>
<tr>
<td>Fundación Carulla</td>
<td>1400</td>
<td>Bogotá, Barranquill, Santa Martha</td>
<td>Children from families in extreme poverty</td>
<td>Adoption services</td>
</tr>
<tr>
<td>SoyDoy</td>
<td>1200</td>
<td>Bogotá, Ubaé, Bello and La Estrella</td>
<td>Children living in very poor areas</td>
<td>High-quality early education</td>
</tr>
<tr>
<td>Fundación Semana</td>
<td>250 families</td>
<td>Bogotá and El Salado</td>
<td>Families and children victims of forced migration or demobilized from the armed conflict</td>
<td>Community production of nutritional complements based on soy</td>
</tr>
<tr>
<td>Fundación Carvajal</td>
<td>200 000</td>
<td>Cali</td>
<td>Families and children in deprived areas</td>
<td>Management and promotion of inclusive business projects</td>
</tr>
<tr>
<td>Fundalectura</td>
<td>102 000</td>
<td>Bogotá</td>
<td>Children and adolescents</td>
<td>Nutrition and early education</td>
</tr>
<tr>
<td>Egypt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abou El Seoud Social Development Association (ASSDA)</td>
<td>35–37 per year</td>
<td>Cairo; Ein El Sira</td>
<td>Children at risk</td>
<td>Literacy and educational services to help children who are dropouts catch up with the regular schooling system</td>
</tr>
<tr>
<td>New Horizon for Social Development</td>
<td>Approximately 3000 per year</td>
<td>Cairo, Ein El Sira</td>
<td>Children in deprived areas and their families</td>
<td>Training children on safe working methods</td>
</tr>
<tr>
<td>The Coptic Evangelical Organization for Social Services (CEOSS)</td>
<td>Approximately 1000 per year</td>
<td>Cairo, El Nozha, El Gadida</td>
<td>Children at risk (street children, working children and children with special needs)</td>
<td>Direct services to families of children at risk to help improve their quality of life</td>
</tr>
<tr>
<td>The New Fostat Non-Governmental Organization</td>
<td>Not available</td>
<td>Ein El Sira, Cairo</td>
<td>Working children and their families</td>
<td>Advocacy for children’s rights</td>
</tr>
<tr>
<td>Alwan Wa Awtar (A&amp;A) Translation: Colors and Strings</td>
<td>Approximately 4000 per year</td>
<td>Cairo, Mokattam Hills</td>
<td>Poor children in marginalized, underprivileged areas</td>
<td>Umbrella organization that works with community-based organizations providing them with funding and technical assistance to address the needs and dangers facing children at risk</td>
</tr>
<tr>
<td>Alshanek Ya Baladi (AYB) Translation: For You My Country</td>
<td>Approximately 100 children per year</td>
<td>Cairo, Old Cairo</td>
<td>Poor children in slum areas</td>
<td>Helps families of working children through helping them find jobs, start income-generating activities and issue legal papers entitled them to government pensions</td>
</tr>
</tbody>
</table>

67
by the M&E specialist and project officers to check that all data collected by the social specialists are correct. If there are problems with data, they go back to the social specialists to verify the data. A number of NGOs mentioned that they conduct either surveys or focus groups to measure satisfaction among service recipients. They all seem to require accurate documentation of all records, especially financial expenditures, and all transactions require stamped receipts, signatures and IDs.

**CURRENT PRACTICE IN PERFORMANCE REPORTING**

*Colombia*

All of the NGOs in the Colombian sample have local funders (foundations and public agencies), and some of them have international funders such as USAID, International Organization for Migration and International Development Bank. In all but one case, funders require performance reports in different formats and with different levels of detail (Table 3).

A common feature in most of the reporting requirements is a bias towards reporting financial indicators (how the money is being spent) and less on achievement of outcomes. The hierarchy of requirements in terms of frequency of reporting to funders is as follows: inputs (money spent, infrastructure), activities (e.g. how the food is served, if the menu requirements are met), outputs (e.g. number of children served, number of food portions provided) and outcomes (e.g. nutritional status, cognitive and socio-emotional development).

Despite the hierarchy of measures required by funders, NGOs reported that they view outcomes measurement as a priority. In all cases in our Colombian sample, the NGOs are very concerned about being able to measure how well they are achieving their main objectives, and they find that reporting does not necessarily lead towards organizational learning. Also, most of the NGOs in the sample are required by funders to set targets, but this practice varies by funder. Some funders require targets in terms of coverage (usually number of children served), and very few require targets in terms of results. Also, it is important to note that most of the NGOs set their own targets even if they are not required to do so by funders.

*Egypt*

The majority of the Egyptian NGOs interviewees reported that they receive funding from a variety of donors, including international organizations (e.g. European Union, USAID, Drosos, Near East Foundation, Ford Foundation, SIDA, Danida, JICA and Save the Children), private corporations (e.g. Vodafone, ARAMEX, EFG Hermes, Samsung and SADCO), government entities (e.g. the National Council for Childhood and Motherhood NCCM, the Social Fund for Development and the Ministry for Social Solidarity), in addition to individual funders. The average number of donors from which the NGOs receive funding is five.

All six NGOs stated that funders require different formatting for the M&E reports although they may all end up having more or less similar content. The exception to the rule was the some funders do not require reporting at all. For example, the Ford Foundation does not require reporting for some of their funded projects.
Table 3. Current practice in performance reporting

<table>
<thead>
<tr>
<th>Types of formats required by funders</th>
<th>Colombia</th>
<th>Egypt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Funders require different formats with different levels of detail</td>
<td>Bias towards reporting financial indicators</td>
<td>Funders require different formats although similar in overall content covered</td>
</tr>
<tr>
<td>More reporting on inputs and activities and less on outputs or outcomes</td>
<td>Overemphasis on quantifiable indicators and too little opportunity for narratives</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reporting practices</th>
<th>Vary by size of organization. Large organizations have staff or departments devoted solely to data consolidation and reporting</th>
<th>Vary by size of organization. Umbrella organizations may have a separate department</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smaller organizations have a more decentralized process</td>
<td>In most cases preparation of reports is a team effort</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data completeness of reports</th>
<th>Complete and accurate data are required by funders</th>
<th>Data must be accurate and complete to meet donors’ requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>The data are not always complete and adequate to meet internal needs</td>
<td>No additional data completeness requirements were voiced</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Performance targets requirements</th>
<th>Most of funders require targets in terms of coverage</th>
<th>All NGOs mentioned that funders require setting performance targets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very few require targets in terms of results</td>
<td>As NGOs interviewed all deal with international donors, they are required to report outputs and outcomes</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Requirement to report outputs and outcomes</th>
<th>Most NGOs are required to report outputs but not necessarily outcomes</th>
<th>Occasional micro-management by donors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurement</td>
<td>As NGOs interviewed all deal with international donors, they are required to report outputs and outcomes</td>
<td></td>
</tr>
<tr>
<td>Data processing</td>
<td>Occasional micro-management by donors</td>
<td></td>
</tr>
<tr>
<td>Data analysis</td>
<td>Lack of clarity concerning donors’ reporting requirements</td>
<td></td>
</tr>
<tr>
<td>Sometimes, data quality and verification techniques</td>
<td>Lack of performance measurement and reporting skills within NGOs</td>
<td></td>
</tr>
<tr>
<td>Gap between measurement for reporting and ability to use the information for evaluation and performance improvements</td>
<td>Occasional inadequacy of performance measures</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Challenges</th>
<th>Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data processing</td>
<td></td>
</tr>
<tr>
<td>Data analysis</td>
<td></td>
</tr>
<tr>
<td>Sometimes, data quality and verification techniques</td>
<td></td>
</tr>
<tr>
<td>Gap between measurement for reporting and ability to use the information for evaluation and performance improvements</td>
<td>Multiple reporting formats required</td>
</tr>
</tbody>
</table>
Some NGOs commented on the language requirement for the reports, stating that it differs from one donor to another. Some funders require reporting in Arabic and others in English. Some NGOs commented that there are different preferences among funders for quantitative versus qualitative data. USAID requires more quantitative data, whereas Drosos requires more narrative reporting in the Performance Reports.

The majority of NGOs stated that they have to report on both outputs and outcomes to their funders. An interesting observation by one of the NGO officials was that donors sometimes confuse outcomes with impacts. It seems that the different usage of terminology by different donors, and its subsequent translation into the Arabic language, sometimes leads to confusion.

All NGOs interviewed reported that they have been able to meet the reporting requirements of funders, relate their achievements and set performance targets. Although they are convinced of the usefulness of reporting on their performance, they do not find it to be easy. However, they believe they have no other option. Reporting periods ranged anywhere from monthly, quarterly, annually, bi-annually to only end of project reporting requirements.

As for who prepares the performance reports, there is much variation in practices followed within the Egyptian NGOs. The NGOs reported that the responsible party is typically the project head, or the Program Manager or the M&E officer (when such a position is available), or the Executive Director, when the reports have to be in English, as not all staff know English. It is noteworthy that some Egyptian NGOs stated that they hire external consultants to prepare the performance reports required by the funders. There is a growing market in the development industry in Egypt for M&E skills, and qualified experts provide their services directly to the NGOs either by preparing the performance reports on their behalf or by providing them with training.

CHALLENGES REGARDING REPORTING

Colombia

The biggest challenge that is common to all the NGOs in the sample is to have the data systematized so that they can be aggregated and analysed. Interviewees reported frustration that while they are collecting data that are very valuable and they can use them at an individual level, they are not able to benefit from aggregate data. For example, some follow nutritional status individually and are able to identify a problem on an individual basis (if a child falls below a critical point, for example), but because they do not have a way to aggregate data and make more refined analyses (such as comparing sites, or age groups), they face limitations in using the information to inform decision making.

Another important challenge is the design of adequate measures. There are some outcomes that are very difficult to measure but are important to the NGO, such as entrepreneurship (how to measure that people in the community change their outlook from looking for the daily minimum income (rebusque) to be an entrepreneur) and child development (how to measure social and cognitive development in a way that is feasible for the organization).

Egypt

One of the key challenges for NGOs in Egypt is to cope with excessive micro-managing imposed by some donors. Interviewees reported that some donors come in and photocopy every receipt, thus wasting both time and money for the NGOs. As one interviewee summed it up: ‘Some donors are a pain.’

Another challenge stems from a lack of clarity on the part of donors from the beginning regarding their reporting and verification requirements. For example, in the midst of one of the projects funded by the World Bank, the Bank decided to introduce additional performance targets, which produced problems, as indicated by the NGO’s officials who made the point that ‘it is usually better to plan everything from the start.’

A third challenge mentioned is that the NGOs need to have staff develop the skills required for data collection and reporting. Some NGOs mentioned that when beginning work with donors and having to meet their measuring and reporting requirements, it is usually stressful and confusing, but it becomes easier with time and practice.

A fourth challenge identified is that the NGO staff members do not believe that the performance measures fully capture what the NGO is accomplishing. For example, developing indicators for art education is very difficult. In
addition to meeting the donors’ reporting requirements, one NGO involved in art education maintains a separate portfolio for each child that includes all his or her artwork, in addition to a record of behavioural critical incidents, both positive and negative, but the donors do not request any of these data.

A fifth challenge noted by the NGOs stems from the variety and complexity of reporting formats required by different funders, which results in time-consuming efforts to prepare all of the separate reports.

**USE OF PERFORMANCE DATA**

**Colombia**

All the NGOs in the Colombian sample reported that they often use performance data to support day-to-day management of their programs. In particular, data about costs per child served are used in making decisions about providers or salaries and to find ways to provide the same quality of services with fewer resources. Also, data about dropouts (from the program) are used to examine the quality of the services that are being provided and to investigate other areas that are affecting children that the program is not taking into account. For example, in the case of one NGO, an analysis of data about dropouts from an income generation project showed that there were problems among the beneficiaries regarding drug use and domestic violence. This finding suggested to the NGO that they needed to redesign the services so that they also provided social and psychological support to beneficiaries.

In terms of strategic management, the data are also used to revise strategies, and to close or expand programs. For example, in the case of one NGO, because of the success of a technology program in schools in a deprived neighbourhood, they are using the same model to expand the program into other neighbourhoods. Also, another NGO said that performance data have allowed them to ‘change the strategic map’ of the organization. Finally, all but one of the Colombian NGOs in the sample use performance data for marketing success and fundraising.

All of the NGOs find performance measurement very useful. Some described it as a ‘fundamental tool’, ‘indispensable’ or ‘a must’. In contrast, most of the NGOs are not satisfied with the potential benefits from reporting. Some declared that there is a disconnect between the reporting requirements and what the organization needs to know in order to learn and improve. For example, sometimes there are requirements to provide details on inputs and activities and not so much on results.

Interviewees were virtually unanimous in saying that they find the performance measures that are most useful to them are the ones that document program participation, and those that document intended outcomes, such as improvements in nutritional status and child development.

**Egypt**

The majority of Egyptian NGOs interviewed mentioned that they use performance data in making decisions about day-to-day management. For example, on the basis of the performance data collected, one NGO made adjustments to how it implements its education and awareness program. They decided to split the soft skills training session from the leadership training sessions and merge the art and literature workshops. Another example was given by an NGO implementing a ‘Theatre for the Oppressed Program’ —theatrical play where children are given the opportunity to simulate the challenges they face in their communities. On the basis of the data collected, when repeating the program a second time, they decided to not to ask for funding to cover costumes, as this was something the children did not enjoy. They also decided to accept more children as they knew from the records that there was an expected high drop-out rate, and to include an orientation period before the theatre program, so that children would have sufficient time to adapt to the experience.

On reporting on the use of performance data to make strategic decisions, some interesting examples were provided. One NGO explained that after they worked for some time with street children and working children, they realized that there was a need for a specialized program for ‘Children at Risk’ and decided to create such a program. Another NGO mentioned that after working some time with young children, they realized they did not have sufficient psychological expertise and decided to focus more on teens rather than on children. Still a third organization
recounted their experience in implementing a project targeting the training of 400 children to move them from hazardous to safe jobs. Data collected from the project revealed that 48 per cent of the children working in hazardous jobs were between the ages of 6 and 12 years. It became apparent that they would not be able to train them because it was illegal to hire them in the first place. Accordingly, they adjusted the strategic goal of the project to include getting the younger children back to school, rather than moving all children to safer jobs.

As for using performance data for marketing purposes, especially marketing the NGOs success to secure funding from existing or new donors, all Egyptian NGO interviewees reported that this happens, but there are variations in practices followed. Some just maintain a summary of their activities and send it out to donors, others wait for donors to approach them and then they show them their previous reporting documents, whereas still others proactively use the data to update their websites and make their organizations more visible on the web.

When the NGOs were asked if the performance data they collect are used in other ways besides those already mentioned, all six Egyptian NGOs mentioned a variety of different uses including, but not limited to, the following:

- advocating for new initiatives such as the ‘Health Insurance Initiative for Street Children’ or the ‘New Child Law’;
- improving staff performance by analysing their points of strengths and weaknesses;
- developing a record ‘documented memory’ for the organization whereby employees can access historical data, or look for ideas for new projects, or rely on when preparing for presentations in conferences;
- orienting and informing new incoming staff about the work of the organization;
- meeting the needs of the reviewers sent out by government authorities; and
- verifying the efficiency, quality, sustainability and impact of programs.

Despite the multiple uses cited for performance data, several NGOs were not completely satisfied with the overall performance measurement and reporting processes and expressed various concerns. One of these concerns was that international donors may insist on the collection of specific, sometimes not totally relevant data, to satisfy their political agendas, for example, to collect information about levels of political participation and engagement of political parties in projects relating to children. Another concern voiced was the overemphasis on quantitative data, which do not always capture the whole picture.

When asked about which performance measures they find most useful, one NGO reported that all the measures are useful without specifying examples. Four of the six NGOs reported that they find outcome measures to be more useful than other sorts of measures. Examples of especially useful measures included, in the case of education service providers, some quantitative data, such as the number of kids who become literate and who are granted certificates of verification from the government to prove it. And in the case of awareness programs, or arts and crafts initiatives, the measures deemed most useful were the success stories reported in progress reports and the comments by animators about the children’s performance in class—both examples of qualitative outcome data.

One NGO mentioned that they would like to have impact measures, for example, a measure of the change in children’s attitudes and behaviours, but they realize that the data required would be very difficult to collect. In summary for the Egyptian NGOs, it seems that the managers perceive outcome measures to be more useful to them in improving their work than other sorts of measures, and they believe that measures do not have to be quantitative to be useful.

AN ANALYSIS OF THE CRITICAL TRANSACTIONS BETWEEN FUNDERS AND SERVICE PROVIDERS

On the basis of the practices within the NGOs in Egypt and Columbia reported by our interviewees, the encouraging finding is that performance data are found to be useful in informing day-to-day management of the NGOs. However, there are four main tensions between funders and service providers affecting performance measurement and organizational learning. Although these tensions have been discussed previously in the literature (Chambers, 2007; Carman, 2009; Ebrahim, 2010), the findings here document the need, not only for their further study, but also for devising strategies to address them:
1. Collection and reporting demands are overly burdensome. Too many data elements with too much detail are required; there is too much micro-managing by donors, and there are too many formats unique to individual donors.

2. There is disproportionate attention devoted solely to financial reporting. Too little discretion is given to NGOs to also discuss achievement of outcomes in ways appropriate to demonstrate achievement of mission-driven goals.

3. There is inadequate capacity within the NGOs to meet funders’ expectations and to spend time learning from the data. As one would expect, the smaller the NGO, the less likely they were to have capacity to analyse and make sense of the data in the NGOs studied here. The fact that some NGOs in Egypt resort to hiring external consultants to prepare the donor demanded performance reports is a case in point. More capacity is needed to address multiple funders’ reporting requirements, involve stakeholders in a meaningful and valuable manner, develop capabilities of staff within the providers to collect data, assess the quality of data, analyse data and use the data internally to improve service and enhance learning.

4. A disconnect exists between reporting and use of data for internal learning and program improvement. Many of the data elements required by funders do not inform providers about how to make improvements in service delivery. There are no incentives given to providers to use the data to make improvement internally. And, the compliance burden in reporting to funders leaves little energy for further attention to be given to learning internally and strategic thinking (e.g. trend analysis).

How are these tensions viewed by the funders? As noted earlier, we investigated performance measurement reporting requirements within Save the Children and UNICEF. We found that these two large funders of NGOs offering services to children across the world have dissimilar reporting requirements and systems for collecting and using the data they collect from their grantees. These two large funders provide funding to NGOs within countries that then provide grants to the service providers, such as those studied here. Thus, funders such as these two can establish reporting requirements that affect (but not totally determine) the requirements that the service providers receive on the ground.

UNICEF reporting requirements depend upon the length of the relationship and the amount of money involved in the grant. All grants, regardless of the amount of money involved, must submit separate financial and program progress reports. UNICEF provides a general template for the program process report as well as for the financial report. In practice, program process templates tend to be modified and customized by UNICEF Staff to adapt them to particular grantee situations. Financial reports are usually not customized and are more rigid in format.

On the other hand, Save the Children reporting requirement are not as uniform as UNICEF. Most of Save the Children’s funding is based on grants from third parties. Each of the third parties will have its own reporting requirements. Similar to UNICEF, Save the Children requires both financial and program progress reports.

Neither UNICEF nor Save the Children typically requires grantees to report on outcomes. UNICEF strategy with respect to outcome measurement is at the country level, and both UNICEF and Save the Children recognize that individual grantees are unlikely to have the capacity to report on outcomes.

Practice regarding use of the data received from the grantees varies. In UNICEF’s case, program progress reports received from grantees are reviewed by the local UNICEF office where the grantee is located. These data are used to assess how well UNICEF in meeting it targets and to decide whether changes need to be made. During this cycle, grant managers have autonomy to make decisions about continuation and level of funding.

In Save the Children’s case, program progress reports are analysed by field management to assess whether progress and performance criteria are being met. Grants are terminated if the grantee is not performing. Unlike UNICEF, Save the Children does not have an institutionalized formal review process. However, it is highly probable that informal practices are at play within each country office or management unit.

UNICEF is moving away from a one-size-fits-all evaluation approach to a customized approach sensitive of grantees’ capacities and local situations. UNICEF also has a presence in the field where grants are being executed. This presence is supposed to shield UNICEF from bogus feedback and may facilitate provision of support to grantees with incapacity to handle reporting requirement on their own.
Save the Children reports that they try to work closely with grantees to help them build reporting capacity; however, they recognize that they do not work directly with service providers because they give money to intermediary organizations. Because the reporting structure ‘follows’ the money, Save the Children managers oftentimes find themselves in a position where it is impossible for them to monitor the work of the service providers who are not directly receiving money from Save the Children.

Funding that support service providers comes from multiple governmental sources, for example, US AID, and foundations from multiple countries, and local funders, in addition to the large funders in a particular service arena, such as UNICEF and Save the Children. Thus, reporting practices and internal use of data within the funders is likely to vary, thus imposing different requirements and pressures on the providers, as was heard in this study.

External accountability is clearly exerted by all of the funders, both locally and internationally based, to ensure that financial resources are invested appropriately and to minimize corruption, as would be expected. However, other than successfully ‘following the money’, there are institutional barriers, for example, intermediary organizations in the management chain, timing issues and/or capacity deficits that prohibit active encouragement of learning by funders for the service providers, or for themselves.

**FUNDERS’ ROLE IN ENCOURAGING INTERNAL ADAPTIVE LEARNING**

We set out to explore the reality on the ground in Colombia and Egypt in order to consider how funders might promote learning among service providers and thus improve services to the least advantaged in our societies. The lessons we learned lead us to call for explicit and purposeful change in the signalling from funders to grantees to improve learning both on the ground and for funders.

Our findings support propositions that Alnoor Ebrahim set forth about how external, upward accountability requirements affect internal learning from evaluation efforts, albeit the type of evaluation mechanism we examined was limited to performance measurement (2005). In line with Ebrahim’s expectations, organizational learning within an NGO is less likely to flourish when (i) performance measurement is perceived to be a task imposed on managers or outside experts rather than viewed as central to their own work; (ii) there are no strong feedback loops developed between field staff, managers and directors; (iii) time and training are not provided to the NGO staff and managers on how to interpret and use the performance data; and (iv) accountability reporting is driven by the donors’ reporting systems and templates, rather than by internal design to serve internal learning. All of these factors were evident in the majority of the NGOs studied, specifically the weak level of internal levers to support internal organizational learning, reaching an extreme in the case of outsourcing the performance management task to external consultants, delivering the required reports to the concerned donor and then shelving the copy, never to be referred to again. Interviewees reported on specific uses of data to support managerial actions (as reported in Table 4) but, here, was not a more pervasive sense of adaptive learning throughout the organizations in any, but perhaps one, of the NGOs in the entire study.

In order to support internal adaptive learning within the NGOs they support, funders need to have a mental model of the learning processes they want to create on the ground. Providing high-quality and effective services is clearly the overriding goal, but it would help to envision how requirements, reporting processes and capacity shape learning as well. As Sowa et al. noted, understanding how to produce program effectiveness is complex, and ‘To improve outcomes, organizations need to understand how their structures and processes enable or hinder those outcomes’ (2004, p. 715). Incentives need to be designed to support learning and ongoing process improvements. Support to help grantees develop evaluation and monitoring capacity is needed, and recognition that each grantee has multiple donors with differing demands is needed.

The complex adaptive societal systems in which the least advantaged children and families live, whether it be Cairo or Bogotá, present especially complex challenges. There are unlikely to be quick fixes or intervention strategies that work everywhere. ‘One size fits all’ thinking is unlikely to be helpful—in performance reporting system design or in devising effective strategies to help 6-year-old children working on the streets. We suggest that funders play an important role in encouraging learning about what works and where and under what circumstances. We suggest that funders should be held accountable for how the incentives and disincentives
they provide to grantees affect their ability to offer effective services. This notion of ‘reverse accountability’ means that funders need to be strategic and thoughtful when they take actions that affect learning behaviours within their grantees.

Drawing upon Ebrahim’s conceptualization of accountability mechanisms, we suggest that performance assessment can be used to institutionalize adaptive learning and strategic thinking and that funders can encourage this by being intentional in developing reporting mechanisms that move providers beyond compliance and support internal and downward (e.g. clients), as well as upward, accountability (Ebrahim, 2010). We extend Ebrahim’s model by suggesting that performance assessment can be used to fruitfully enhance accountability for participation (of stakeholders) and adaptive learning, two of the other accountability processes that Ebrahim delineates in his framework. And we focus on the role that funders can play in shaping these accountability processes.

In line with our call for reverse accountability, we offer here a model demonstrating our notion of the two-way flow of accountability (Figure 1). We also offer recommendations to help improve the performance reporting environment for NGOs who are addressing complex problems with less than adequate capacity. Our recommendations not only reflect the experience of the 12 NGOs studied here but are also based upon the over 25 years of experience of the senior author in building evaluation capacity within NGOs within and outside of the USA.

### RECOMMENDATIONS FOR FUNDERS

**Capacity building**

Funders should provide or fund training in performance data collection, analysis, interpretation and use to support M&E in the field. Training on performance assessment, for example, monitoring, and evaluation for service providers should be offered on site or via webinars in the languages of the providers and should become an integral part of any funder’s program. The training should include guidance on logic modelling, outcomes hierarchies or other tools that help providers clarify desired outcomes of their programs, trend analysis and both quantitative and qualitative data reporting techniques as well. Funders should also encourage and reward providers for performing monitoring activities in-house, as opposed to outsourcing these activities, to enhance a sense of ownership and learning within the organizations. Conducting an assessment of the staff’s capacity in performance measurement and analysis could be one of the criteria for assessing an NGO’s eligibility for new or renewed funding.

### Table 4. Use of performance data

<table>
<thead>
<tr>
<th>Use of performance data</th>
<th>Colombia</th>
<th>Egypt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of performance data in the day management of programs</td>
<td>Quite often</td>
<td>Quite often</td>
</tr>
<tr>
<td>Use of performance data in strategic management</td>
<td>Sometimes (given the limitation of data aggregation and data analysis)</td>
<td>Quite often</td>
</tr>
<tr>
<td>Marketing success to secure funding by donations from existing or new funders</td>
<td>Quite often</td>
<td>Quite often</td>
</tr>
<tr>
<td>Usefulness of performance measurement and reporting</td>
<td>Performance measurement is perceived as very useful, but not reporting</td>
<td>Performance measurement and reporting are perceived as useful especially when not purely quantitative and not overly tedious and time consuming</td>
</tr>
<tr>
<td>Examples for use of data</td>
<td>Cost per child data are used to revise program’s efficiency</td>
<td>Outputs (such as the number of children trained on working in a safer job or the number of children attending an illiteracy class) used to revise activities</td>
</tr>
<tr>
<td></td>
<td>Program participation (and dropout data) are used to revise activities</td>
<td>Data on outcomes (such as number of children trained who move to safer jobs or number of children who pass the national illiteracy exam) are used to revise activities</td>
</tr>
<tr>
<td></td>
<td>Data on outcomes (such as nutritional status or readership) are used to revise activities</td>
<td></td>
</tr>
</tbody>
</table>

### Table 4. Use of performance data

Copyright © 2012 John Wiley & Sons, Ltd.


DOI: 10.1002/pad
**Mission achievement**

Reporting requirements should be leaner and more focused on achievement of the espoused mission-driven goals of the providers. Data requirements should be aligned with the mission and goals of the service providers. And service providers should be given some latitude in developing outcome measures and then asked to clarify how the measures relate to their mission. It should be noted that this is not meant to replace financial reporting, as accountability for the specific uses to which funds are directed is also necessary, but reporting about how mission-oriented objectives are achieved, and if they are not an explanation of why they are not, should also be required.

**Reporting requirements**

Formats, or templates, for reporting should be streamlined and simplified to permit service providers to provide similarly formatted reports to multiple donors. When developing reporting requirements for service providers,
funding agencies should take into account the specific obligations of these same organizations to other funders and try to minimize duplication and harmonize reporting requirements.

Funders should give service providers some discretion in how they characterize outcomes attributable to their work when reporting. Qualitative assessments of outcomes, such as success stories and critical incidents, that supplement quantitative data requirements should be encouraged and guidance should be provided on how to do so in an efficient manner. For example, if an organization is reporting on children’s artistic development, the use of qualitative assessments, through the use of work portfolios, in addition to quantitative assessments should be accepted.

**Stakeholder involvement**

Providers should be encouraged and rewarded for involving their stakeholders (including beneficiaries, for example) in designing M&E processes (for example, see Fowler, 1995). Funders could require providers to list who was involved in deliberations about what to measure, what the data ‘mean’ and strategic thinking about mission-oriented objectives. Guidance on how to effectively engage stakeholders should be provided in the capacity building training described earlier.

**Adaptive learning**

Providers should be consistently encouraged and rewarded for learning from performance assessment processes. For example, funders could require providers to report about changes in service delivery that were made in response to learning from the data, that is, instrumental use of the data, as well as more strategic thinking about problem solving and reconceptualization of the problems their clients face, and redesign of programs and interventions (i.e. conceptual use of the data). Guidance on how to analyse performance data in order to make connections to improve both program management and to planning and redesign should be given high priority in capacity building training, because program managers throughout the world are less skilled in making these connections than one would expect (Moynihan, 2009; Moynihan and Pandey, 2010; Moynihan et al., 2011).

**CONCLUSION**

In the end, we are calling for both funders and NGO managers to make the connection between performance assessment and adaptive learning processes. Service providers must be held accountable for the funding they receive, and they should also be held accountable for mission achievement and for adaptive learning to improve the performance of their programs. Moving beyond a compliance mentality when reporting on performance in both public and non-profit organizations is a challenge, but the rewards benefit all. As Ebrahim recently noted, ‘In the end, accountability is about both being held to account by external actors and standards, as well as about taking internal responsibility for actions’ (2010, p. 25). We recommend that both funders and service providers take responsibility for using performance data to support strategic thinking and adaptive learning to continually improve services delivered to the most needy in societies across the world.

**REFERENCES**


