Measuring the impact of international corporate volunteering: Lessons learned from the Global Health Fellows Program of Pfizer Corporation

Final Report

Research Team:

Taryn Vian, MSc
Assistant Professor, Team Leader

Frank (Rich) Feeley, JD
Associate Professor

William MacLeod, ScD
Assistant Professor

Sarah C. Richards, PhD
Adjunct Assistant Professor

Kelly McCoy, MPH
Senior Program Manager

September 21, 2007

Acknowledgements
This evaluation was supported by funding from by the US Agency for International Development through the Child and Family Applied Research Project (Contract #GHS-A-00-00020-00), and from Pfizer. Rekha Chalasani and Patricia Bowen, Atiya Ali, Lisa Herren Foster in the Philanthropy Department at Pfizer, provided logistical support in contacting respondents and feedback on questionnaire design, while Mary Jordan at USAID Global AIDS Office provided input to the study design. At Boston University School of Public Health, Ms. Sushma Hyoju assisted in the analysis of quantitative data, and Ms. Mayra Nicola helped with production of the Evaluation Toolkit.
Table of Contents

I. Executive Summary .......................................................................................................... 1
II. Introduction and Background ......................................................................................... 8
III. Study Goal and Objectives ............................................................................................ 9
IV. Methodology ....................................................................................................................10
   A. Review of prior research ............................................................................................... 10
   B. Study design ................................................................................................................ 10
   C. Data collection ............................................................................................................. 11
   D. Analysis ....................................................................................................................... 11
V. Findings ............................................................................................................................ 11
   A. International Corporate Volunteering: Research on Monitoring and Evaluation ....... 12
   B. Evaluation of Round 5-6 Fellowships ....................................................................... 18
      Questionnaires .............................................................................................................. 18
      Fellow and Fellowship Characteristics ....................................................................... 18
      Characteristics of Partner Organizations ...................................................................... 19
      Role Clarity and SOW ................................................................................................. 19
      Fellow and Fellowship Characteristics as Perceived by Partner Organization .......... 21
      Partner Organization Capacity ..................................................................................... 23
      Accomplishments ......................................................................................................... 24
      Evidence of Impact ...................................................................................................... 30
      Sustainability and Attribution ..................................................................................... 33
      Fellowship Types and Impact Scores .......................................................................... 35
   C. Impact Index ................................................................................................................ 36
VI. Discussion ........................................................................................................................ 41
VII. Recommendations .......................................................................................................... 45
   A. Use process model to understand ICV program impact .............................................. 45
   B. Use impact index and rankings to track progress over time ...................................... 45
   C. Adapt tools for each program’s design and priorities ............................................... 45
      1. Adaptation of Tools ............................................................................................... 46
      2. Data collection strategies ....................................................................................... 46
      3. Data analysis .......................................................................................................... 46
VIII. Conclusion ..................................................................................................................... 47
Annex 1: Interview Contacts .................................................................................................. 48
Annex 2: Fellow Accomplishments by Technical Area, according to Partner respondents 49
References .................................................................................................................................. 51
I. Executive Summary

Goals and Objectives
Boston University’s Center for International Health and Development conducted research on tools and methods for evaluating international corporate volunteer (ICV) service programs from October 2006 through May 2007, using empirical data from the Pfizer Global Health Fellows (GHF) Program. The goal of the study was to design a toolkit to measure the impact of ICV on recipient organizations and their ability to deliver efficient, high quality services. We pilot tested the evaluation tools with a small sample of Pfizer corporate volunteers. The research was done with the approval of the Boston University Medical Campus Institutional Review Board, under an extension of the approval for the first study of the GHF program which evaluated 72 earlier Fellowships.

The first BU evaluation of the GHF program, conducted from October 2005 to March 2006, examined program impact from the perspective of Partner Organizations, Fellows, Pfizer work teams, and the company as a whole. The survey methods and qualitative interviews used were similar to methods used to evaluate other ICV programs, and the data obtained appeared accurate, especially regarding the impact of the program on Fellows, work teams, and the company as a whole. However, the measurement of program impact in the field, involving synthesis of perceptions of Fellows and Partner Organizations as well as evidence of accomplishments and outcomes, was complex and context-specific, and the evaluation tools in this domain were not easily transferable. A logical next step, therefore, was to develop more streamlined tools to capture and summarize program impact from the viewpoint of Partner Organizations, in order to make these tools available to the GHF program and other companies engaging in ICV.

This second phase of research looks in more depth at the impact of an ICV program on the Partner Organization and its capacity to deliver public health services. By exploring this evaluation domain in more detail, the BU team hopes to provide improved tools which can be used by many kinds of public-private partnerships to evaluate their ICV programs.

Methods
The research team developed the evaluation model and designed data collection instruments. Using documentation on past Fellowships representing a range of impact levels, we created two questionnaires: one collected data from Fellows, and the other was for Partner Organization respondents. Partner Organization respondents were representatives of the organizations receiving the volunteer assistance. These included local offices of international NGOs, indigenous NGOs, and government offices in developing countries. In many cases, the same question was asked of both the Partner organization and the Fellow in order to provide comparative data.

---

1 Other implementation partners work with Pfizer to support the ICV program, including placement agencies such as Health Volunteers Overseas (HVO) and American Jewish World Service (AJWS). These are not included as Partner Organizations for the purposes of this evaluation.
Pfizer organizes and sends Fellows to the field in cohorts, called “Rounds.” Each Fellow has a unique scope of work, and Fellowship dates and duration may vary, but the cohorts receive orientation and follow-up support as a group. We tested our data collection tools on a sample drawn from 44 Fellowships from Round 5 and Round 6 which took place from January 2006 to early 2007. We chose countries where multiple Fellows had been placed or adjacent countries. This strategy gave us data on 20 Fellowships in five countries: India (6), Namibia (1), South Africa (6), Tanzania (2) including Zanzibar, and Uganda (4). Most of these Fellowships (14, or 74%) were under six months long, while 5 (26%) were six months or longer.

Interviews with Partner organization respondents were conducted in person, while Fellows were mostly interviewed by phone. Data were analyzed using SAS version 9.1. We used the qualitative data to provide examples and to further explain or clarify these results. We used principle components analysis to explore data on organizational development status, and to create an overall impact score as well.

Findings

1. Lessons from the IVS literature

Beyond reporting of project results, there has been little evaluation research on international volunteering programs. Challenges for evaluation include the many different outcomes to volunteering, and the difficulty in creating indicators that can be used cross-nationally (Hills and Mahmud 2007; Hutter and McDowell 2007).

One firm which has done consulting work looking particularly at corporate volunteering is the London Benchmarking Group (LBG). LBG’s model examines inputs, outputs, and impacts of volunteering programs and has been applied to a BD (Becton, Dickinson and Company) employee group volunteering program in Zambia. Another company which has focused on evaluation of their ICV program is the transportation multinational TNT. Five years ago, TNT created the TNT-World Food Programme (WFP) Partnership, which sends corporate volunteers to assist WFP field offices. TNT first tried to calculate return on investment in ICV, but found that some impacts could not easily be translated into budget savings. TNT now uses Key Performance Indicators, which capture different kinds of impact, but are harder to summarize and compare across fellowships.

2. Evaluation of Round 5 and Round 6 Pfizer Global Health Fellowships

In our testing of the evaluation toolkit with Pfizer Fellows, we interviewed respondents from 20 Partner Organizations in five countries. Partner Organizations in the sample were mostly large, local NGOs with prior experience hosting volunteers. Fellows were mostly female (74%) and had 17 years’ work experience on average, 8 of which were at Pfizer. Eight Fellows (42%) reported having lived or worked in a resource-constrained country before. Areas of expertise included microbiology, oncology, HIV/AIDS,

---

2 Zanzibar has its own Ministry of Health and functions as a separate nation from Tanzania in organization, financing, and delivery of health services.
3 London Benchmarking Group (http://www.lbg-online.net/)
pharmacy, and medical research, as well as many types of management skills. Study findings include the following:

**Scope of Work and Role**
- Most partner organization respondents (85%) fully understood the role of the Fellow as being related to technical assistance and/or capacity strengthening.
- Among Partner respondents, priority areas for capacity building were management functions (67%) and research/clinical systems (14%); however, many Partner respondents found it hard to choose among the different options, and noted that personnel development was a closely related priority. Among Fellows, 47% mentioned management functions as the top priority, while 36% mentioned personnel development.
- While 90% of Partner respondents stated that the Fellow understood the assignment very well, only 42% of Fellows thought the Partner Organization understood the assignment very well.

**Fellow and Fellowship Characteristics**
- Most Partner Organization respondents said that Fellows were very prepared (57%) and flexible (81%).
- Almost all Partner respondents (20 out of 21) said the Fellow interacted a lot with staff, a characteristic that research indicates may be associated with greater impact in terms of skills transfer and technology transfer.

**Partner Organization Capacity**
- Overall, Partner respondents in this sample felt they were strong in organizational capacity at baseline. Most felt they had clear mission and goals, good strategies, high levels of training and motivation, and financial flexibility at the time the Fellow arrived.
- Some Partner respondents perceived they had weaknesses in management systems, training, and numbers of personnel.
- On average, Fellows rated adequacy of staff time and numbers higher than Partner respondents did, at baseline, while Fellows were more likely to perceive that the baseline level of staff training and financial flexibility are lower.

**Accomplishments**
- Accomplishments were reported in nine areas: curriculum development; training; leadership or strategy; networking & partnerships; management systems; financial stability or revenue generation; marketing; information systems; and research programs or systems. One-third of Partner respondents mentioned accomplishments in the areas of marketing or research, while 90% of respondents said the Fellowship had developed staff skills.
- Fourteen Fellows (74%) said they had conducted some kind of needs assessment.
- When asked to assess organizational capacity at the end of the Fellowship, most Partner respondents said that the level of organizational capacity had improved or stayed the same. (In the case of motivation, two respondents noted a change for the worse, though changes were attributed to issues not related to the Fellowship.)
• Partners had a more optimistic assessment of the changes in organizational capacity by the end of the Fellowship than did Fellows. Fewer Fellows rated change as “much better” and more rated the situation as the same or worse.
• 86% percent of Partner respondents and 84% of Fellows thought GHF performance had exceeded or significantly exceeded the responsibilities outlined in the scope of work. Partner respondents were more likely than Fellows to report that GHF performance had significantly exceeded the SOW.
• Most Partner respondents and Fellows said the organization had sufficient time and financial resources to execute changes, and Partner responses were generally optimistic that changes could be sustained.

3. Impact Measurement
Fellowship impact was measured using nine indicators, covering the areas of organizational impact, attitudinal impact, and leveraging impact. The results for this sample are illustrated in the table below.

<table>
<thead>
<tr>
<th>Impact Measures</th>
<th>% who reported impact by Fellowship</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impact Category</td>
<td>Measure</td>
</tr>
<tr>
<td>Organizational impact</td>
<td>Service expansion</td>
</tr>
<tr>
<td></td>
<td>Efficiency of operation</td>
</tr>
<tr>
<td></td>
<td>Service quality</td>
</tr>
<tr>
<td>Attitudinal impact</td>
<td>Motivation &amp; pride in work</td>
</tr>
<tr>
<td></td>
<td>Self-confidence</td>
</tr>
<tr>
<td></td>
<td>Time management</td>
</tr>
<tr>
<td></td>
<td>Strategic thinking</td>
</tr>
<tr>
<td>Leveraging impact</td>
<td>Financial resources</td>
</tr>
<tr>
<td></td>
<td>Material resources</td>
</tr>
</tbody>
</table>

For organization impact, includes responses “a little” and “a lot”
For attitudinal impact, includes responses "moderate", "high" and "very high"
For leveraging impact, includes "yes" responses.
% Diff = (% Fellows - % Partners) / % Partners

We tested alternative ways to create an index and ranking to combine and report on the Fellowship impacts described above. The results from the preferred method, a Sum of Scores Index, are shown below. The index is used to calculate an overall impact score for each fellowship, and to show where fellowships fall in the range of possible impact rankings (e.g. low, moderate, high, or very high impact). The index was tabulated separately for Fellows and Partner respondents.
### Sum of Scores Index for Measuring ICV Impact

<table>
<thead>
<tr>
<th>Ranking of Fellowship</th>
<th>Impact Score Range</th>
<th>Partners</th>
<th>Fellow</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very High Impact</td>
<td>9-15</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>High Impact</td>
<td>16-21</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Moderate Impact</td>
<td>22-27</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Low Impact</td>
<td>28-33</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

These results show that 60% of Fellowships scored by Partner respondents and 75% of Fellowships scored by Fellows could be considered Very High or High Impact.

#### 4. Sustainability and Attribution

Almost all Partner respondents said that skills or materials imparted were being used at the time the Fellow left. In about half these cases, the interviewer noted that evidence was available to support the assertion. Regarding attribution, Partners readily attributed impact to the presence of the Fellows. For example:

- 100% of Partner respondents stated that the Fellow had helped to speed up change
- About 80% thought the Fellow had contributed new ideas
- Two-thirds thought that the outcomes would not have happened if the Fellow had not come, and close to three-quarters of respondents said they would have had to pay someone else to do the work the Fellow did.

Although we do not know whether the Partner would really have prioritized the tasks or allocated resources to pay someone if the Fellow had not come, these findings do suggest the Partners assigned the Fellows to do work which was of significant value to the organization.

### Discussion

In developing our ICV Evaluation Toolkit, we considered several alternatives for measuring impact. One alternative was to measure performance in comparison with objectives of the Fellowship. While this method can work in some cases, it runs into problems when the scope of work is changed over time in order to be responsive to needs
identified later in the Fellowship. Adaptation of scopes of work results in better technical assistance, but it also complicates the evaluation process.

Similarly, we found it hard to measure impact by **documenting accomplishments with quantifiable indicators**. This method was not able to capture all types of impact. For example, some tasks could be measured by indicators such as increased number of HIV+ patients on antiretroviral drug treatment, a direct impact measure. More often, though, the scope of work involved tasks which were less directly related to impact yet still important for capacity strengthening, such as the development of a plan. This type of task is measured by a binary (yes/no) indicator, which says little about the quality or usefulness of the plan.

The focus on quantifiable indicators appeared to be limiting in a cultural sense as well. One Partner respondent resisted the questions about quantifiable objectives, stating that this approach was “very American…we don’t really think that way.”

Finally, we tested the reliability of asking respondents to categorize Fellowship impact according to the typology developed in the previous Pfizer GHF evaluation. With these Fellowships, as with the first Pfizer evaluation, most Fellowships were rated a Type 2 or Type 3 (Type 2 showing impact on quality or efficiency, and Type 3 showing impact on service expansion); however, we found that Partner Organization respondents and Fellows had different opinions as to the Type which should be assigned to a Fellowship.

**Recommendations**

1. Given the diverse needs of recipient organizations and the many types of assistance provided, we believe that ICV impact should be measured using a **process model with multiple measures and a composite impact index**. First, an evaluation strategy which includes multiple measures to examine different dimensions of performance is likely to be more useful for making decisions regarding program improvement. Multiple measures also allows for the collection of contextual information which is important for interpreting impact measures, and for measuring sustainability and attributing how much of the impact is due to the Fellowship. Using the composite impact index in addition allows ICV programs to summarize the impact using weighting that reflects program priorities. It also provides a way to compare the perceptions of Fellows and Partner respondents, and to monitor the progress of the ICV program over time.

   The Boston University ICV Toolkit, submitted under separate cover[^4] includes questionnaires to collect data on the multiple measures and a spreadsheet for calculation of the composite impact index.

2. We believe that the weights implicit in our composite index are appropriate for the Pfizer Global Health Fellows Program. For example, organizational impact (3 measures) and attitudinal impact (4 measures) are weighted more heavily than leveraging impact (2 measures), which accurately reflects the relatively low amount of emphasis which the

[^4]: Also available by contacting the lead author, Taryn Vian ([tvian@bu.edu](mailto:tvian@bu.edu))
Pfizer program puts on the latter dimension of performance. The rankings should be interpreted with caution, however, as they do not take into account the baseline capacity of the organization. Organizations with better capacity to begin with may show a lower impact from the fellowship.

3. The BU ICV Toolkit can be adapted for use by companies running or considering international corporate volunteering programs. Depending on the goals of the program, the factors and weights which are currently used in the BU Toolkit can be changed. These changes will affect comparability of data between companies, and should be explicitly noted.

Some implementation issues which should be considered by both Pfizer and other corporations include adaptation of questions for each corporation’s specific program design; data collection strategies and the choice of on-site versus telephone interviews; and data analysis plans which incorporate statistical analyses and use of qualitative data.

Conclusion
This study designed and pilot tested an evaluation toolkit to assess the impact of Pfizer’s international corporate volunteering program on partner organizations. With adaptations to adjust for different program designs, this toolkit can be useful to many public-private partnerships interested in documenting this aspect of program impact. The new evaluation toolkit found results that were quite similar to the more in-depth partner interview methodology used in a prior Pfizer evaluation. In both evaluations, we found that most Fellowships achieved many accomplishments leading to high levels of perceived impact on service quality, efficiency, and service expansion. This new toolkit provides a streamlined set of measures which allow for comparison of results between types of respondents and across Fellowship sites. This information is essential for continuous quality improvement of international volunteering programs so they can achieve maximum effectiveness and continue to strengthen the capacity of organizations to respond to the key public health challenges in global communities today.
II. Introduction and Background

This study was designed to develop tools and methods for the evaluation of international corporate volunteering (ICV) programs, based on the experience of Pfizer Corporation and the Global Health Fellows (GHF) Program. The study was conducted from October 2006 to May 2007, and grew out of ideas raised by an earlier evaluation of Pfizer’s ICV program. Between October 2005 and February 2006, the Boston University Center for International Health and Development conducted an evaluation of Pfizer’s GHF Program including surveys of 109 Fellows and Supervisors, field visits to five countries, and over 100 in-depth interviews. The research, supported by Pfizer and USAID, was designed to assess the impact of the Fellowship program on partner organizations, Fellows, and Pfizer itself. A key motivation behind the prior study was to determine whether this type of public-private partnership could accelerate service expansion in the countries and organizations where it was working.

The 2005-2006 evaluation reported on perceptions of impact as described by the Fellows and Partner Organization respondents. These findings are presented in the final report of the evaluation (Feeley, Connelly et al. 2006), and in two published papers (Vian, Richards et al. 2006; Vian, McCoy et al. 2007). The evaluation focused on several types of impact, including impact on the Fellows themselves, the Partner Organizations, the Pfizer work groups from which the Fellows came, and Pfizer’s reputation in the global health and development arena.

To summarize the program’s impact on partner organizations, the evaluation sought evidence that Fellowships had affected health and expanded the availability of health services in the countries where Fellows worked. Yet, several structural considerations and data problems made impact hard to measure. Partner organizations were at different levels of their own development, and their work in global health was sometimes not directly related to service delivery (i.e. the International AIDS Vaccine Initiative, which is working in AIDS vaccine development). In addition, the Fellows’ scopes of work were sometimes not well defined, or had evolved over time to meet needs that emerged after work had started. Finally, there is little published research on measures to evaluate the impact of ICV programs, so there was no clear model for evaluation that could be easily applied to the GHF program.

Despite these difficulties, the BU research team created a typology that identified three possible levels of Fellowship impact. By reviewing data from Fellows, their Pfizer Supervisors, and Partner Organizations, the research team was able to place most Fellowship assignments within this typology. The analysis suggested that 31% of the Fellowships could be categorized as Type 3, presenting evidence of a sustainable impact in increasing the level of service or efficiency of the Partner Organization. An additional 48% of Fellowships were coded as Type 2, demonstrating skills transfer and systems development progress, seen as an essential precondition or stepping stone to service expansion. Thirteen percent of Fellowships fell into Type 1, indicating a solid volunteer effort not based on the special skill and expertise of the Pfizer employee, and not
transferring scarce skills to the Partner. Eight percent of Fellowships lacked sufficient data for classification (Vian, Richards et al. 2006).

USAID expressed interest in adapting the BU typology for use in evaluating public-private partnerships more broadly, and quantifying the impact of ICV programs in other contexts. This follow-on study was designed to build the empirical grounding for such a typology and to assess its usefulness in categorizing the impact of fellowship experiences.

The research looks only at one evaluation domain: the impact on the Partner Organization and its ability to expand services. It specifically does not address evaluation measures for assessing the impact of international volunteering on the Fellows or Corporations who participate. Nonetheless, the study results are likely to be relevant to many kinds of public-private partnerships designed to build capacity of service delivery and research organizations in low-income settings. The research was done with the approval of the Boston University Medical Campus Institutional Review Board, under an extension of the approval for the first evaluation of the GHF program (October 2005-March 2006).

III. Study Goal and Objectives

The goal of the study was to develop a methodology and tools for measuring the impact of international corporate volunteering programs on local partner organizations. This type of evaluation toolkit is important to justify philanthropic investments in this type of development assistance, and to provide information for continuous improvement of programs. In addition, the results of the study will help Pfizer to produce higher impact Fellowships, which will benefit Pfizer stakeholder relationships and the professional development of participating Pfizer employees.

The study objectives included:

1. Review research findings on impact evaluation for ICV programs;
2. Develop survey tools for evaluating the Pfizer GHF program, looking particularly at the measurement of impact on the partner organization;
3. Test the evaluation tools by collecting and analyzing data from two groups of Fellows (Round 5 and 6), including visits to partner organizations in South Africa, Namibia, Uganda, Tanzania, Zanzibar, and India;
4. Make recommendations for how the evaluation toolkit can be used by Pfizer and other companies to implement effective ICV programs.
IV. Methodology

A. Review of prior research

Because so little published work exists, we conducted phone and in-person interviews with academic researchers and other corporate, government, and NGO leaders involved in this field. The team leader participated in the December 5, 2006 International Volunteering Leadership Forum organized by The Brookings Institution and spoke with some corporate directors of philanthropy and academic experts. The research team also conducted interviews with key individuals at the Center for Social Development at Washington University in St. Louis, FSG Social Impact Advisors in Boston, TNT Corporation’s Public-Private Partnership with the U.N. World Food Programme, and the Washington, DC based NGO Innovations in Civic Participation. See list of contacts in Annex 1.

B. Study design

The research team developed the evaluation model and designed data collection instruments. Using documentation on past Fellowships representing a range of impact levels, we created two questionnaires: one collected data from Fellows, and the other was for Partner Organization respondents. Partner Organization respondents were representatives of the organizations receiving the volunteer assistance. These included local offices of international NGOs, indigenous NGOs, and government offices in developing countries. In many cases, the same question was asked of both the Partner organization and the Fellow in order to provide comparative data.

Pfizer organizes and sends Fellows to the field in cohorts, called “Rounds.” Each Fellow has a unique scope of work, and Fellowship dates and duration may vary, but the cohorts receive orientation and follow-up support as a group and as individuals, as needed. We tested our data collection tools on a sample drawn from 44 Fellowships from Round 5 and Round 6 which took place from January 2006 to early 2007. We chose countries where multiple Fellows had been placed or adjacent countries. This strategy gave us data on 20 Fellowships in five countries: India (6), Namibia (1), South Africa (6), Tanzania (2) including Zanzibar (1), and Uganda (4). There were 5 men and 14 women Fellows included. Among the 20 Fellowships selected, two Fellows have subsequently left Pfizer. We successfully traced one former employee, but another Fellow who left the company could not be traced. Even in the latter case, we were still able to interview the Partner Organization.

5 Other implementation partners work with Pfizer to support the ICV program, including placement agencies such as Health Volunteers Overseas (HVO) and American Jewish World Service (AJWS). These are not included as Partner Organizations for the purposes of this evaluation.

6 Zanzibar has its own Ministry of Health and functions as a separate nation from Tanzania in organization, financing, and delivery of health services.
C. Data collection

The data collection process included contacting Partner Organizations respondents, interviews with Partner Organizations, and interviews with Fellows.

As the Pfizer GHF program uses organizations such as American Jewish World Service (AJWS) and Health Volunteers Overseas (HVO) as placement agencies, contact information for counterparts within local NGO Partner Organizations was sometimes not available, and took some effort to assemble. We e-mailed or phoned the organizations to request participation. In one case (Bihar, India) the partner organization did not seem eager to participate due to inconvenient timing. This site also posed some security issues, so it was not included in the sample. In another case (CMC Pharmacy, India), the partner organization respondent was on leave, but agreed to participate by phone. All other interviews with partner organization respondents were conducted in person.

Fellows were also contacted by e-mail or phone to request participation. Most Fellows were interviewed by phone. Two interviews were conducted in person. As mentioned earlier, one Fellow who had left the company was lost to follow-up. None of the Fellows refused to participate.

Both the Partner Organization questionnaire and the Fellows questionnaire were administered by research team members. Instructions for the interviewer were printed on the forms.

D. Analysis

Using SAS version 9.1, we produced frequency tables and other descriptive statistics for the quantitative data collected. We used the qualitative data to provide examples and to further explain or clarify these results.

We used principle components analysis to explore data on organizational development status, to determine whether there were separate components that would explain organizational development. We also created a composite index for organizational development, looking at Fellows and Partner respondents separately. We followed similar procedures to create an overall impact index as well.

V. Findings

The findings section is organized into three parts. In the first part, we discuss the findings related to other research on the impact of ICV. In the second part we describe the results of the evaluation of the sub-set of Round 5-6 fellowships. Finally, in part three, we present a composite index and ranking comparing the impact of fellowships.
A. International Corporate Volunteering: Research on Monitoring and Evaluation

While there is growing interest in international volunteer service (IVS), the body of literature on impact of IVS is small, and there is even less information available on the impact of international corporate volunteering (ICV).

Researcher Michael Sherraden⁷ of the Center for Social Development at Washington University in St. Louis has defined voluntary service as “an organized period of substantial engagement and contribution to the local, national, or world community, recognized and valued by society, with minimal monetary contribution to the participant.”(Sherraden 2001). International corporate volunteering has been defined as “engaging employees in service projects in international markets,” including local service as well as cross-border service (Hills and Mahmud 2007). Sherraden distinguishes service from employment, in that service is not “well compensated by monetary reward.” In this key dimension, ICV differs from regular international volunteer service. One could argue, however, that ICV also fits the standard definition if compensation is considered an organization-level characteristic, rather than a characteristic measured at the individual level, because most corporations do not receive monetary compensation for the international voluntary service of their employees.

In most other aspects, ICV is similar to other forms of international volunteering. Margaret Sherraden and co-authors have developed a typology of international volunteer service that categorizes initiatives by purpose (promoting international understanding versus development aid and humanitarian relief), duration (short, medium or long term), eligibility (expert or not), direction of flow of volunteers (North-South, South-South, etc.), degree internationality (unilateral, multi-lateral, or trans-national), and other dimensions (Sherraden, Stringham et al. 2006). All of the IVS initiatives reviewed by Sherraden et al. have been sponsored by public or private NGO organizations; yet corporate programs can also be described using this typology. Table 1 applies the typology to Pfizer’s GHF program.

According to Sherraden et al., even short-term and medium-term programs can sometimes have long-term effects if a series of volunteers go to the same organization or community over a period of time. This is the case for the Pfizer GHF program in some countries; for example, five fellows have worked in the Pharmacy at Christian Medical College in India over time. Having a “medium-term” volunteering program which also has long-term presence in an organization may make evaluation of impact more complex.

⁷ There are two authors with last name of Sherraden and first initial of M.--Michael and Margaret. The publication dates are different for the articles cited, and Margaret Sherraden’s article is with co-authors.
Table 1: Characteristics of International Volunteer Service Programs, including Pfizer GHF Program

<table>
<thead>
<tr>
<th>Dimension of International Service</th>
<th>Spectrum for this Measure</th>
<th>Where Pfizer’s GHF program fits in</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purpose</td>
<td>International Understanding or Development Aid/Humanitarian Relief</td>
<td>Main purpose is Development Aid/Humanitarian Relief</td>
</tr>
<tr>
<td>Nature</td>
<td>Group or individual placements</td>
<td>Individual placements</td>
</tr>
<tr>
<td>Duration of service</td>
<td>Short (1-8 weeks), medium (3-6 months), long (&gt;6 months)</td>
<td>Medium (3-6 months)</td>
</tr>
<tr>
<td>Degree of Internality (international exposure)</td>
<td>Unilateral (1-way, sending country doesn’t receive volunteers also); bilateral or multilateral (2-way, countries both send &amp; receive); transnational (volunteers from more than 1 country serve together in multiple countries)</td>
<td>Mainly unilateral, but occasional multilateral (e.g. India)</td>
</tr>
<tr>
<td>Eligibility &amp; participation requirements</td>
<td>Not skills-based (i.e. willingness to learn and serve) versus skills-based</td>
<td>Mainly skills-based</td>
</tr>
<tr>
<td>Flow of volunteers</td>
<td>North to South, South to North, North to North, South to South</td>
<td>Mainly North to South, but some South to South</td>
</tr>
<tr>
<td>Number of countries where volunteers serve</td>
<td>Review of 8 programs found range of 6 to 139, with six programs between 16-72</td>
<td>Approximately 30</td>
</tr>
</tbody>
</table>

*Characteristics of IVS and range of options for the measures are derived from Table 2 and text of Sherraden, Stringham, Sow, and McBride (2006)

Sherraden and co-authors present the results from a few studies of impact of IVS. Evaluations of programs whose main purpose is promoting international understanding and learning have sometimes shown long term effects on cultural awareness, future vocational choices, and extra-curricular activities of former volunteers. But beyond reporting of project results, the authors could find little evaluation research on IVS programs whose goal is to provide development aid and humanitarian relief. The documentation that does exist seems to suggest that skills and technology transfer may happen more through such programs because volunteers are “working alongside local professionals,” as opposed to working on independent projects (building a school, for example). Long term volunteering for development and relief is generally individual, not group. This means there may be some confounding of the effect of these two dimensions of volunteering (i.e. nature of placement and duration of assignment).

Expressing concern that international service programs for development and relief may be reflecting the goals of the donors rather than the recipient organizations, Dr. Sherraden and co-authors suggest that impacts should be assessed more broadly, for “host organizations, communities, and [recipient] countries” not just volunteers and sending countries (Sherraden, Stringham et al. 2006). The Global Service Institute at the Center for Social Development would like to do quasi-experimental research where they measure the impact of a volunteering experience on the volunteers, organizations, and communities involved, and compare this with a control group. They also believe more work needs to be done in documenting best practices for IVS (Sherraden 2007).
Katy Hutter and Charlotte McDowell, of **Innovations in Civic Participation (ICP)**, agree that impact assessment is a major issue facing the field of IVS, and that developing evaluation tools, procedures, and research is very important (Hutter and McDowell 2007). ICP is a Washington, DC-based NGO focused on increasing opportunities for youth to engage in quality service programs in their communities. They work in the U.S. and abroad, but don’t currently focus on projects that involve Americans volunteering overseas.

According to Hutter and McDowell, the challenge is that there are many outcomes to volunteering, and many inputs. Two especially challenging issues are how to structure indicators and analyses so they can be used cross-nationally, and the fact that the programs being helped and the volunteering opportunities within those programs have different goals, or sometimes many goals at the same time.

Turning specifically to corporate volunteering programs, **FSG Social Impact Advisors** have identified best practice examples and models of strategic corporate volunteerism. This research, undertaken in cooperation with the Brookings Institution’s International Volunteering Initiative and funded by Pfizer, took place from January 2006 through July 2006. FSG created case studies based on 10 organizations, examining common dimensions. The dimensions which FSG has chosen to explore in detail include:

- Social Sector Focus Areas (not all programs focus on health);
- Implementation Partners (types of beneficiaries, other intermediary organizations partnering with the company to help with logistics, management, etc.);
- Operational Model (local service or cross-border service);
- Corporate Resources Deployed (staff, volunteer time, travel, other funding)
- Business and Social Impact (how ICV fits in the context of business strategy, the leveraging of corporate assets and expertise)

The results of the FSG study confirm that until now most evaluations of the impact of ICV have used anecdotal evidence as well as surveys of employees. Many of the companies included in the study identified impact measurement as an important area of focus for the near future.

Another firm which has done consulting work looking particularly at corporate volunteering is the **London Benchmarking Group (LBG)**\(^8\). LBG’s approach was used in developing a case study of the BD (Becton, Dickinson and Company) Employee Volunteer Partnership Program in Zambia (Dow 2006). Through this program, two groups of 10 BD associates from multiple countries traveled to Zambia to work at five sites in collaboration with the Catholic Medical Mission Board (CMMB). LBG’s model was used to examine inputs, outputs, and impacts of volunteering programs. Table 2 shows how the dimensions of the model were applied to BD’s Zambia program.

---

\(^8\) London Benchmarking Group (http://www.lbg-online.net/) is not to be confused with LBG Associates (http://www.lbg-associates.com/), another consulting firm which also has developed some publications on standards of practice for corporate volunteering programs. LBG Associates reports are available for sale on their web site. We did not access these reports as part of this study.
Table 2: LBG Model for Quantifying Contributions and Impacts of Volunteer Programs

<table>
<thead>
<tr>
<th>Category</th>
<th>Dimension</th>
<th>Application to the BD Zambia Employee Volunteer Partnership</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Inputs</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cash</td>
<td></td>
<td>• BD contribution $151,700 (transport, materials, etc.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• BD employer “match” to employee fundraising ($13,564)</td>
</tr>
<tr>
<td>Time</td>
<td></td>
<td>• Estimated value of volunteer time $23,000 (BD skilled employees were paid during 2 week volunteering experience)</td>
</tr>
<tr>
<td>In-kind</td>
<td></td>
<td>• BD donated $50,000 in equipment</td>
</tr>
<tr>
<td>Management</td>
<td></td>
<td>• 1,000 hours, valued at $48,000</td>
</tr>
<tr>
<td><strong>Outputs</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leverage</td>
<td></td>
<td>• $63,000 in additional resources attracted to this activity, including $20,000 in employee fundraising, in-kind donations from Dell and DHL, and management costs of recipient organization</td>
</tr>
<tr>
<td>Community benefits, local (5 hospital sites)</td>
<td></td>
<td>• Sites received donated basic medical supplies of BD products, and received training in their use</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Staff at 5 labs trained in routine lab work, operations improved (log books, etc)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Awareness raising on healthcare worker safety</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Two new incinerators built; Waste management systems, processes improved</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Other renovations (kitchens, laboratory renovation)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Counterparts received guidance in how to “source” additional supports</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Hands on help from BD volunteers</td>
</tr>
</tbody>
</table>
|            |           | • Local sites closer to qualifying for PEPFAR funding (“leveraging”)
| Community benefits, global (CMMB) | | • Advanced CMMB’s program objectives |
|            |           | • Strengthened CMMB-BD donor relations (benefiting projects in other countries) |
|            |           | • Helped CMMB create new processes for volunteering programs with other donors, expanding use of volunteers |
|            |           | • Improved credibility of CMMB with other local partners, thereby expanding donor base |
|            |           | • Increased CMMB visibility in press |
| Benefits to BD | | • Enhanced skills and experience of volunteers (leadership, training skills, communication, influence, diversity, innovation, and new perspectives) |
|            |           | • BD seen as living out its values and mission |
|            |           | • Changed personal outlooks on life, improved morale |
| **Impacts** |           |                                                             |
| Community, local level (hospitals) | | • Better patient services at hospital (one hospital increased VCT uptake) |
|            |           | • Better safety, less accidents, less risk of infection (unconfirmed) |
|            |           | • Hospitals attracting more donor funding (progress made) |
| Community, global (CMMB) | | • CMMB improved efficiency (no data given) |
|            |           | • New corporate partnerships |
|            |           | • More services to offer partners |
| Business impacts | | • Better awareness of company purpose among employees |
|            |           | • Impacts of employee personal skill development |
|            |           | • Reduced absenteeism due to improved commitment (no data given) |
|            |           | • First hand information about needs for healthcare products in LDCs |
|            |           | • Improved reputation as responsible corporate citizen |


The ultimate goal of the LBG model is to quantify a return on investment (ROI) by comparing the total inputs to outputs. In the Zambia case, however, it was not possible to calculate ROI since the outputs could not be quantified.
Finally, another group which has tried to calculate the return on investment in ICV is transportation multinational TNT. TNT has created the TNT-World Food Programme (WFP) Partnership which has been in operation for five years. This public-private partnership includes three initiatives:

1) Specialist Program. Through the specialist program, TNT consultants work with local WFP offices, providing technical assistance to lower the cost of logistics for bringing in food aid. The goal of the assistance is to improve efficiency so that more of WFP money goes to the actual food rather than logistics. In Sierra Leone, for example, a consultant worked on an improved system for fleet management.

2) Volunteer program and “Hands-on Help”. Through the volunteer program, about 40 TNT employee volunteers per year spend 3 months working with WFP Feeding Programs in 5 countries, while through the “Hands-on Help” program, TNT uses its own capacity to assist in emergencies. For example, TNT airplanes might fly in to provide food aid in the initial days of a disaster (this is an important contribution, since WFP, as a U.N. agency, has to use a tendering process and can’t respond in the first days of a disaster). Also, TNT employees provided assistance during the South Asian Tsunami as part of this program.

3) Fundraising and awareness program. Employees are encouraged to organize fundraising events, big or small, to benefit WFP. In addition, TNT has developed a module for teaching school children in The Netherlands (where TNT is headquartered) about how children live in poor countries. TNT employees go to local schools and provide training using the module.

TNT has been concerned with evaluating their ICV program, and has tried two approaches so far. According to TNT manager Henriette van Eaghen, neither approach is seen as ideal (van Eaghen 2007).

The first approach was “Multiplier Analysis” and is a variation on the Return on Investment methodology of LBG. TNT asked each recipient of TNT specialist or volunteer services to estimate the amount of value received. It was thought that the value would generally be cost savings due to greater efficiency. This was then divided by the cost of the program. For example, if a program received $20K in cost savings, and the TNT consultant cost $10K to the company, this resulted in a 2.0 multiplier. The advantage of the multiplier was that it gave the company one summary measure they could view each year, since they could sum all the multipliers from all the programs and calculate the overall multiplier at the end of the year.

The disadvantage of this method of measuring performance was that the value of some programs wasn’t easily captured in budget savings. For example, before a school can participate in a WFP feeding program, it needs to meet certain standards (have a clean water source, have a latrine). TNT helped schools to achieve the required standards by digging wells and building latrines. This expanded feeding programs to more children in poorer schools, i.e. to children who really needed the help. This kind of benefit isn’t
captured in “cost savings” but still the WFP is pleased to be meeting more needs and expanding aid. In another example, when TNT sends a plane to enable WFP to respond faster with food aid, the benefit is measured more in terms of time savings than in terms of cost savings.

TNT then switched to their current system of Key Performance Indicators (KPI). In this system, for each program intervention the TNT staff member and recipient office will define the aim, or key result they want to achieve. This could be cost savings, but could also be something like “To be able to respond to an emergency 4 days earlier by having a plane available.” The advantage is that different kinds of impacts can be measured. The disadvantage is that it is complicated. It is harder to summarize the KPIs and difficult to communicate them. Plus, there’s no single number to determine overall success of the partnership.

Another issue with KPI’s is that they aren’t helpful in choosing projects. In the past, the “multiplier” was used to help screen projects. People tried to estimate in advance what the multiplier would be, and were inclined to favor projects which would have a high multiplier. But the KPI’s don’t allow this type of screening, so they can’t help TNT in choosing what projects to support.

TNT’s Memorandum of Understanding with the World Food Programme is up in December 2007. At this time, they are anticipating trying to come up with a new way to measure outcomes and impact. TNT hopes to be able to measure achievement of the overall goal of the partnership, which is to reduce the number of hungry people (children) in the world.

In addition to TNT’s program, other Dutch employers have promoted employee volunteering and are working to develop systems for evaluating them. For example, ABN-AMRO Bank has evaluated their volunteering program in terms of participants’ attitudes and behavior toward their employer, including attendance, performance, and potential turnover (intention to leave the company) (Gilder, Schuyt et al. 2005). Fortis, another Dutch Bank, provides a lot of volunteering options for employees, and has worked on measurement as well. In addition to measuring number of hours spent volunteering, Fortis also conducts surveys of the employee volunteer, manager of the employee, and the receiving organization. Data captured include items such as internal motivation and outputs. These data are put into an automated monitoring system, called MBO Monitor. The Bank has been trying to get other companies with volunteering programs to use the database also, which would provide some comparative benchmarking data. However according to Henriette van Eaghen, the database is currently only available in Dutch, which is less interesting to a multinational corporation such as TNT (van Eaghen 2007).
B. Evaluation of Round 5-6 Fellowships

Questionnaires

As mentioned earlier, we interviewed respondents from 20 Partner Organizations in five countries. As one Fellow split time between two departments within an organization, we actually administered 21 questionnaires to respondents. Most respondents (81%) self-identified as having been the supervisor of the Fellow.

We interviewed 19 Fellows, including one Fellow who had left Pfizer following the Fellowship. This Fellow had gone on to work for the Partner Organization and was paid through a donor grant. One other Fellow had also left the company and was lost to follow-up.

Fellow and Fellowship Characteristics

As mentioned in the Methods section, most of the Fellowships were under six months (14, or 74%). Fellows had 17 years’ work experience on average; more than half the Fellows had over 15 years of professional experience prior to the Fellowship, and only 21% had less than 10 years. All had worked for Pfizer for at least five, and about 30% had 10 years experience or more working at Pfizer. Eight Fellows (42%) reported having previously lived or worked in a resource-constrained country.

We asked Fellows to list five areas of professional expertise or skill, a question which is also asked on the Fellows’ application to the GHF program. Fellows described a wide range of skills and expertise including deep knowledge of medical areas (microbiology, toxicology, oncology, HIV/AIDS, laboratory, pharmaceuticals), and medical research (clinical trial procedures and management). They also described an array of management expertise, i.e. MBA degree, project management, situation analysis, strategic planning, labor forecasting, marketing, budget analysis, training programs, and pharmacy management, among others.

During the first evaluation, we found that Type 1 Fellowships (i.e. fellowships that did not result in any permanent change in capacity of the host organization) were ones that usually did not require a Fellow to use his or her professional skills and expertise. Therefore, in this evaluation we asked both the Fellow and the interviewer to assess how often the Fellow used prior training and experience in the Fellowship (given the Fellow’s list of skills, the SOW, and the described accomplishments). 100% of Fellows reported that they used prior training and expertise “a lot”, as opposed to the other possible responses (“a little”, or “not at all”). Similarly, interviewers judged that Fellows drew on their expertise all or most of the time in 94% of Fellowships examined.
Characteristics of Partner Organizations

On average, Partner Organizations in the sample were large, local NGOs (defined as 20 or more full time employees), five years of age or older, with experience hosting volunteers. Table 3 presents a detailed breakdown of these variables.

<table>
<thead>
<tr>
<th>Table 3: Characteristics of Partner Organizations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of Organization</td>
</tr>
<tr>
<td>International NGO</td>
</tr>
<tr>
<td>Large local NGO</td>
</tr>
<tr>
<td>Small local NGO</td>
</tr>
<tr>
<td>Government</td>
</tr>
<tr>
<td>Years in Operation</td>
</tr>
<tr>
<td>Less than 5 years</td>
</tr>
<tr>
<td>5 years or more</td>
</tr>
<tr>
<td>Prior Experience with Volunteers</td>
</tr>
<tr>
<td>Hosted prior GHF and/or other Vol.</td>
</tr>
<tr>
<td>Have not had volunteers</td>
</tr>
</tbody>
</table>

Role Clarity and SOW

During the first GHF evaluation, the research team noticed that low-impact fellowships often seemed associated with weak understanding of the Fellow’s role. We explored this issue through several questions in the present study. First, we asked an open-ended question “Please describe your expectations of the Fellow’s role” and recorded whether the respondent mentioned anything related to technical assistance or capacity building. We also recorded whether the respondent mentioned having trouble understanding the Fellow’s role, or described a role that was incorrect (e.g. fundraising for the placement NGO). By this measure, the interviewers judged that most Partner Organization respondents (85.7%) fully understood the role of the Fellow as being related to technical assistance and/or capacity strengthening.

Examples of Partner respondent answers that were judged as showing full understanding of the technical assistance/capacity strengthening role included the following:

- This organization needed help in expanding its [country] operations which built/contracted with PEPFAR for a private sector AIDS treatment network. The Fellow was to help them to develop an organizational strategy and tactics to manage this expanded program. (notes of interviewer)

- [The Fellow] had expertise in QA (Quality Assurance) and [extensive international experience]. The fellow’s role was to work on Manufacturing and Quality Control units within the pharmacy so that they could become fully licensed & function well.

Examples of answers judged as showing only partial understanding, or showing that the Partner understood only later in the process, included:

- When we heard that GHF was coming, we weren’t sure of what role he could have. We are service-based, and he is in marketing. But we thought, if we have
to attract customers, we can use him to help us. Let’s get a feel for what he has to offer.

When we asked for GHF we had heard that they could do a lot of things. So we had her work with us to learn how we work. Then we could see what she could do for us.

In addition, we asked questions in both questionnaires about the scope of work. About half the Partner Organization respondents and Fellows described having to make changes to the Fellow’s scope of work, even after the first two weeks (when the SOW is normally finalized), while the other half made no changes. As one Fellow stated “[It was] a very bureaucratic…agency. They had lots of ideas and knew the bottlenecks in processes, but they didn’t know how to address them specifically.” Another Fellow explained:

They asked me to be involved in eight of their activities, but I told them we needed to focus. I looked at their wish list, and narrowed it based on my qualifications, what could be accomplished in the timeframe, what resources they had access to, and what had already been done.

Eighteen out of 20 Partner Organization respondents (90%) reported that the final SOW was very or mostly clear and precise. Responses were similar for Fellows. However, while 90% of Partner respondents stated that the Fellow understood the SOW very well; however, only 42% of Fellows thought the Partner Organization understood the SOW very well. While one Fellow mentioned that the Partner organization staff were more interested in gaining connections and perhaps getting a job with Pfizer, most of the misunderstanding was due to a willingness to leave the details to the Fellow, changes in staff, or an underestimation of what was possible.

They knew what they wanted, but not really how to get it.

The scope of work was something they had never done before and was outside their normal deliverables…plus, the person who [got GHF involved] left his job the week [Fellow] arrived in country.

At first we tried to set a quantitative goal, but there was confusion between the management level and the operational level. The manager agreed to a 100% goal, but the operational staff didn’t think that would be possible…I gave up on the goal and decided to just develop an approach to the problem. Getting to know people was very important first, before goal setting.

They might have under-estimated what could be done.

Overall differences in understanding of scope of work between partners and fellows were not statistically significant (p=0.5048).

Finally, we asked questions about the priority areas which the Partner Organization perceived as most important for capacity building assistance during the Fellowship, and we probed specific responsibilities in the scope of work (e.g. whether they were quantifiable, whether the respondent’s answer matched the final scope of work on file at Pfizer). Regarding priority areas for capacity building, the top areas were management
functions (67%, or 14 out of 21 respondents) and research/clinical systems (14%, or 3 respondents). Within management functions were included finance, human resources, organizational behavior, marketing, strategic planning, IS, and operations management. Many Partner respondents found it hard to choose among the different options. Personnel development, including training, motivation, leadership development, and mentoring, was often a favored “second choice” for this question, but since we required the respondents to choose only one category, it did not rank high.

Fellows were asked the same question about priority areas for capacity building, with slightly different results: 47% mentioned management functions, and 36% mentioned personnel development.

The set of questions on specific responsibilities in the SOW was less successful. The process at Pfizer is that Fellowship assignments are described in a scope of work developed by the Partner Organization. The scope of work is then revised by the Partner and the Fellow after the Fellow’s first two weeks in country. The final scope of work specifies the objectives of the assignment and the specific tasks or responsibilities that the Fellow is being asked to undertake, as well as the necessary skills and experience needed, and the counterparts with whom the Fellow will work. Compared to the findings of the earlier GHF evaluation, scopes of work were much clearer and better defined. Yet, we found that there were reliability problems in how interviewers labeled objectives as “quantifiable”, and there were data problems in determining quantifiable indicators of performance according to the objectives. We also found that final scopes of work were sometimes missing or the version on file in the corporate office did not match the version shown by the Partner Organization respondent. Although the SOW is an important communication and planning tool, it appears to be less useful for evaluating accomplishments for the reasons mentioned above. We do not include questions related to SOW in the recommended evaluation toolkit.

**Fellow and Fellowship Characteristics as Perceived by Partner Organization**

<table>
<thead>
<tr>
<th>Table 4: Partner Perceptions of Fellow</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>How prepared?</strong></td>
</tr>
<tr>
<td>Very</td>
</tr>
<tr>
<td>Mostly</td>
</tr>
<tr>
<td>A little</td>
</tr>
<tr>
<td><strong>How often flexible?</strong></td>
</tr>
<tr>
<td>All the time</td>
</tr>
<tr>
<td>Most of the time</td>
</tr>
<tr>
<td><strong>How much interaction with staff?</strong></td>
</tr>
<tr>
<td>A lot</td>
</tr>
<tr>
<td>A little</td>
</tr>
</tbody>
</table>

During the first evaluation, we observed that Fellowship impact seemed to be related to characteristics of the Fellow and the Fellowship, including the Fellow’s level of preparation and flexibility. We therefore included such questions on this evaluation, as shown in Table 4. Most Partner respondents thought Fellows were very or mostly prepared. Data gathered from Fellows was very similarly distributed. These quotes illustrate some Partner perceptions about GHF preparation:

His preparation was excellent. He just moved right in and started doing the things [we] needed.
She was very prepared. She tried to learn the language, even. The equipment we had was very diverse, so she wasn’t familiar with it all. But she took the initiative to help us discover together what needed to be done.

It is hard for anyone to be completely prepared if they haven’t lived here before. She had worked in [less developed region] for many years and so she knew how to work in a resource-poor setting. She still needed some time to learn about the environment, but less than some other volunteers.

During the first Pfizer GHF evaluation, the anthropologist on our team observed that the terms “preparation” and “flexibility” were understood more broadly in some settings. Similarly in this study, Partner respondents seemed to use these terms to refer to cross-cultural sensitivity or interpersonal skills in general. For example, in response to the preparation question one Partner respondent expanded, “She fit in from day one. Other people haven’t fit in. A fellow in another organization was shy, reserved, too sensitive to criticism. But this fellow wasn’t like that.” Another Partner respondent stated “She was adaptable and showed good judgement and diplomacy. She didn’t overstep boundaries, but respected [the project sites’] autonomy.” A respondent commenting on a GHF’s flexibility observed that “He has a way of getting people on his side.”

Fellows’ answers to this question were consistent with Partners’ responses: most felt they were very or mostly prepared. Some Fellows mentioned that they wished they had done a bit more work ahead of time to prepare, including more communication with the Partner organization. One Fellow said “I worked up until the time I left, and I didn’t prepare myself a lot…maybe I should have set up more communication with the Partner Organization.” Another mentioned that the NGOs “need to be better prepared to have the benefits of the Fellow. If they haven’t thought of their needs ahead of time, it takes too much time for the Fellow to sort it out…I should have asked more detailed questions to force them to think about what they wanted.” One Fellow followed exactly this advice, reporting that he spoke with his in-country supervisor over several months before he arrived:

“I called him every week…We started talking about family, then gradually about the mission, goals, objectives, deliverables and output timeframe. We agreed on the scope of work and then [he] took it to management before [the Fellow] arrived.”

We also included a question on the level of interaction with staff. As noted by Sherraden et al. (2006), the characteristic of working alongside a counterpart may be associated with greater impact in terms of skills transfer and technology transfer. Due to our small sample size, we didn’t see very much variability on this measure: almost all (20 out of 21) Partner respondents reported that the GHF interacted a lot with staff. Among Fellows, 16 out of 19 respondents said they interacted “a lot” and three respondents said they interacted “a little.”
Partner Organization Capacity

In addition to Fellow and Fellowship characteristics, we wanted to measure how Fellowship impact is related to the Partner Organization’s initial level of organizational capacity. Building capacity is the goal of most ICV programs and a way to enhance and expand an organization’s provision of programs and services. Yet, some organizations have greater capacity than others at the start, and the results of an ICV intervention can be better understood in light of this contextual information.

We asked Partner respondents to think about the area of the organization most concerned with the activities of the Fellowship, and to give their opinion on how well the area was functioning before the Fellowship began, and the amount and direction of any changes that occurred in organizational development during the period of the Fellowship. Responses to the baseline questions about development were on a five point scale ranging from “very high” to “very low.” We coded these as 1 for “very high” and 5 for “very low.” We later asked the respondents about the amount and direction of any changes that occurred in organizational development by the end of the Fellowship. The scale ranged from 1 “much better” to 5 “much worse.” Figure 1 presents the baseline perceptions.

Figure 1: Partner Perceptions of Organizational Capacity at Baseline

Overall, partner organization respondents in this sample felt they were strong in organizational development at baseline. Most felt they had a clear mission and goals and were pursuing appropriate program strategies at the time the Fellow arrived. In addition, most organizations felt they had moderate to high levels of staff training, motivation, and financial flexibility. Categories where more Partner Organizations perceived they had weaknesses included management systems, training, and sheer numbers of personnel, though even in these categories, many respondents thought they were moderately well off or better. All Partner respondents and 18 out of 19 Fellows felt that the Fellow’s scope of work was very or mostly appropriate, given the baseline organizational capacity.

Looking at Fellows’ perceptions of the baseline organizational capacity, we see some differences. Fellows seem to rate adequacy of staff time and numbers higher than Partner respondents do, while Fellows are more likely to perceive that the baseline level of staff training and financial flexibility are lower. These findings could be explained by

---

9 Organizational capacity and organizational development are considered to be synonyms and are used interchangeably in this report. The questionnaire used the term organizational development.
different standards or expectations related to productivity. Fellows may have higher expectations for what staff can accomplish, and are therefore less likely to see problems as due to lack of numbers of staff. Fellows may also have less knowledge of other tasks that staff are being asked to do, outside of the areas related to Fellowship scope of work.

**Index of Organizational Capacity**

We used principle components analysis to explore the data and to determine whether there were separate dimensions that would explain organizational capacity. Principle components analysis is designed to capture the variance in a data set by summarizing the most important (i.e. defining) parts. This analysis gave us two groupings. The first grouping included appropriateness of program strategies, financial flexibility, motivation of staff, strength of management systems, and clarity of organizational mission and goals. We called this first grouping “management”. The second grouping included level of staff training, and adequacy of staff number and time. We called the second grouping “personnel”. We also created a composite index which included all of the responses.

Partner organization respondents ranked their work area (the area or part of the organization most concerned with the Fellowship) with an average of 2.88 on the composite organizational ranking scale (1 being very high and 5 being very low), while Fellows ranked their organizational work area as 2.77. The management variables were scored by the Partners as 2.65 and the Fellows as 2.97, while the personnel variables were scored by the Partners as 2.53 and the Fellows as 2.71. Because the principle components analysis method assigns different relative weights for the population of Partners versus Fellows, these averages cannot be directly compared. However, we might have expected to see a correlation between the responses of individual Fellow-Partner pairs (i.e. in the responses of a Fellow and the staff member from the Partner Organization discussing the same fellowship). In fact, there was no correlation between the Fellows and Partners on the individual scores, meaning there was no broad agreement between the Fellows and Partners on the ranking of the organization.

**Accomplishments**

We asked questions to capture information on accomplishments (outputs), organizational development and changes in staff attitudes (impact) and results (impact). Key results areas which we queried included service expansion, efficiency, service quality, and leveraging of funds.

Looking first at accomplishments, we asked about needs assessments conducted by Fellows. Documentation of a problem or need is itself an important step toward capacity strengthening, and is often the first step before other outputs can be obtained. A needs assessment can also help build local support for change, something which is essential for sustainability. Fourteen Fellows (74%) said they had conducted assessments of staff training or learning needs, and the same number said they had conducted other types of needs assessments. Partner respondents gave virtually the same answers. Other needs assessments (as described by the Partners) included an organizational review leading to a restructuring, SWOT analysis, IT requirements for office operations, assessment of
quality control processes (clinical trials, drug manufacturing unit), and HR assessment. Almost all Partner Organization respondents (95%) stated that Fellows were very or mostly effective in identifying need for change. The number for Fellows was similar (87%).

We posed direct queries about other types of accomplishments by technical area (e.g. curriculum materials, training, management systems, etc.), using categories based on results from the first evaluation. The explicit prompts could have promoted some repetition (the same accomplishment being cited in more than one technical area) or over-reporting (reporting of relatively small accomplishments which might not otherwise have come up), compared to an open-ended question; however, we preferred this strategy to assure that no accomplishments were missed. Table 5 shows one Partner Organization respondent’s impression of a single Fellow’s accomplishments. Annex 2 presents illustrative accomplishments by area for all Fellows.

Table 5: Example of Accomplishments of One Fellow, according to Partner Respondent

<table>
<thead>
<tr>
<th>Area of Capacity Strengthening</th>
<th>Examples of outcome or accomplishments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improvements to curriculum or training systems</td>
<td>Updated the training SOPs (standard operating procedures) to be specific and more comprehensive, covering all topics even one-day trainings</td>
</tr>
<tr>
<td>Developed staff skills and competencies</td>
<td>Helped improve writing skills of staff; provided training in how to monitor clinical trials and how to follow SOPs</td>
</tr>
<tr>
<td>Leadership development or change in strategic direction</td>
<td>None mentioned</td>
</tr>
<tr>
<td>New or stronger partnerships, networking, or new business development</td>
<td>None mentioned</td>
</tr>
<tr>
<td>Operational management efficiency and systems improvements</td>
<td>Created guidelines and tools for monitoring feasibility study sites; what to document, how to disseminate reports. Put in place a system that allows cross site comparisons.</td>
</tr>
<tr>
<td></td>
<td>Created action plan for clinical trial site pre-audit, with steps outlined and place for comments, judgements on how to follow-up (used this with a new site in one country that had not done this type of trial before).</td>
</tr>
<tr>
<td></td>
<td>Created 40 SOPs for clinical trial sites, with level of detail required by external review agency. This was challenging because she did it with full participation of staff, and some of the current practices had to change.</td>
</tr>
<tr>
<td>Strengthening financial stability or improving revenue generation</td>
<td>None mentioned</td>
</tr>
<tr>
<td>Strengthening marketing or public relations</td>
<td>She made our organization look good as she performed her role in relation to the clinical trial sites and all the people involved.</td>
</tr>
<tr>
<td>Strengthening information systems or technology</td>
<td>None mentioned</td>
</tr>
<tr>
<td>Changes in research programs or systems</td>
<td>Improved existing program for initiating and implementing clinical trials and feasibility studies.</td>
</tr>
</tbody>
</table>

In addition to accomplishments in the technical areas mentioned, Partner respondents stated that about 86% of Fellows engaged in personal mentoring (similar response from Fellows). One Partner respondent was a senior hospital manager who had only clinical training and no management background. Speaking of the effects of the GHF’s mentoring, she stated “I learned how to interact with people, how to inform staff and keep them involved. People’s involvement is such an important issue. I’ve been using

---

10 Given the need to maintain confidentiality of respondents in this research study, some contextual details have been removed, i.e. type of organization, and department where the Fellow was assigned, etc...
this same procedure as new opportunities come up.” She described the GHF as a “core member of the management team,” involved in every meeting.

Although 71% of Partner respondents and 74% of Fellows said that Fellows engaged in responsibilities outside their SOW, it often turned out that these tasks had already been mentioned in the SOW. This points out the confusion over what is and what is not contained in the “final scope of work.” Such ambiguity is normal and probably not important, as long as accomplishments are being captured somewhere.

Figure 2 shows perceived change in level of organizational capacity by the time the Fellow left the Partner Organization, according to Partner respondents. The respondents were asked to assess how much the different aspects of organizational capacity had changed by the end of the Fellowship. Caution must be used when interpreting these results, as changes might be due to the Fellow or could be due to other circumstances that were occurring at the same time. In most cases, level of organizational development stayed the same or improved. In the case of motivation, two respondents noted a change for the worse, though in each case the change was attributed to financial or other issues that were not related to the Fellowship.11

Figure 2: Partner Perceptions of Changes that had occurred in level of Organizational Capacity by the end of the Fellowship

Generally, Fellows had a less optimistic assessment of the changes in organizational development by the end of the Fellowship. Fewer Fellows rated change as “much better” and more rated the situation as the same or worse, as shown in Figure 3.

11 Following the questions on changes in organizational capacity by item, we asked a general question about attribution: “For any of the changes above that was not specifically related to the Fellow” please explain why the change occurred.” In a few cases we were able to relate this back to the specific item.
Figure 3: Fellow Perceptions of Changes that had occurred in level of Organizational Development by the end of the Fellowship

Organizational Impediments to Change
Another important issue related to Fellowship outputs and impact is the Partner Organization’s ability to absorb and continue with any improvements introduced by the Fellow. We explored possible organizational impediments which could affect performance in this regard. For example, we asked whether the Partner had sufficient time, financial resources, and policy flexibility to implement any changes the Fellow had recommended.

Ninety percent of Partner respondents and 72% of Fellows said the organization’s staff had sufficient time to execute changes, while 81% of Partner respondents felt that they had the time to sustain those changes over time. About three-quarters of the Partner organizations stated they had sufficient financial resources to execute the changes, and 81% had the resources to sustain changes over time. The slightly higher figure for resources to sustain change may indicate that Partner Organization’s were selective in implementing only those changes that they had the resources to sustain. Fellows had similar opinions about financial resources to execute changes.

Over three-quarters of Partner respondents and Fellows felt that the Partner organization had adequate policy flexibility to execute changes. Examples of policy or structural issues which did appear to impede progress included changes that required involvement of many departments, or approvals from headquarters (i.e. lack of decentralized authority to adopt own management systems). Although most Fellows (75%) felt that the Partner Organization understood what the Fellow was recommending or trying to do, about half of the Fellows felt that their counterparts were not at the right level to facilitate execution of the changes that were needed.

12 We did not ask Fellows their opinions on sustainability of change, as it would require an ongoing relationship with the Partner Organization to be able to answer.
Besides time, money, and policy flexibility, respondents noted some other factors—both positive and negative—that affected their ability to execute changes recommended by Fellows. One Partner respondent stated “the continuum of GHF activities [several Fellows over time] in building financial systems” was a helpful factor, for example. Some of the negative factors mentioned by Partners included changing priorities, lack of office space, lack of technical know-how, and waning enthusiasm after the Fellow departed.

Fellows described several negative factors affecting the Partner Organization’s ability to implement changes, including staff politics and operations management considerations.

“Politics [is a factor]. The managers of the organization (the ones on top) weren’t always in agreement. There were huge communication barriers and questions about who was in charge. It wasn’t very stable.”

“It was hard to get a problem-oriented meeting held with all the staff present. You have to make sure that the people who are needed to make the decisions come to the meeting. Management should know about the real problems, but sometimes they don’t hear the issues. Staff thinks management won’t change, but they could…you just have to get them to hear. But people need to secure their positions. They are afraid to propose change.”

“People have loyalty to certain people, even though they don’t do their jobs.”

“Nobody can be fired; unqualified [people with connections] can be employed.”

“The structure of their day-to-day operations [was a barrier]. I think they under-utilized my skills by not having me work on other projects during my down time.”

In addition, several people mentioned cultural issues, work ethic, and incentives as barriers:

“The work ethic or energy [was a factor]. Sometimes I don’t see that they have the passion or the energy to move forward…”

“The slow pace of work was not driving [changes]. [People] don’t work to timelines and deadlines. Performance review makes no difference, because they are promoted based on seniority not on accomplishments. It is a different motivation.”

“The organization is beholden to [higher level], so policy changes don’t happen fast. Even though you recognize something needs to change, it doesn’t mean you will [change it], and that doesn’t mean you don’t want the change. Remuneration of staff is low. It’s important to understand things like this when trying to think about change.”

Accomplishments by Technical Area
Figure 4 looks at the proportion of Partner respondents who reported at least one accomplishment in a given technical area. This ranged from one-third of respondents who mentioned any accomplishments in the areas of marketing or research, to 90% of
respondents who said the Fellowship had developed staff skills. More Fellows than Partner respondents perceived accomplishments in management systems and partnerships/networking, while more Partner respondents than Fellows perceived accomplishments in curriculum design, financial stability or revenue generation, and information systems. These differences were not statistically significant, however.

Figure 4: Percent of Partner respondents mentioning any GHF accomplishment, by area

Both Partner respondents (25%) and Fellows (31%) said that management systems improvement was the most important area of accomplishment for their organization.

Fellow’s Performance
Finally, we asked both Partners and Fellows a question, similar to one used in Pfizer employee performance evaluation system, which asked them to rate the Fellows’ performance on a 5-point scale, from did not meet responsibilities (unsatisfactory) to significantly exceeded responsibilities. Table 6 shows the results. Eighty-six percent of Partners and 84% of Fellows thought GHF performance had exceeded or significantly exceeded the responsibilities outlined in the scope of work. Partner respondents were more likely than Fellows to report that GHF performance had significantly exceeded the SOW, though again, this difference was not statistically significant.

Table 6: Rating of GHF performance in relation to scope of work

<table>
<thead>
<tr>
<th>Rating of Fellow's Performance by:</th>
<th>Partner</th>
<th>Fellow</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Significantly exceeded responsibilities</td>
<td>6</td>
<td>29%</td>
</tr>
<tr>
<td>Exceeded some &amp; met all others</td>
<td>12</td>
<td>57%</td>
</tr>
<tr>
<td>Consistently met responsibilities</td>
<td>3</td>
<td>14%</td>
</tr>
<tr>
<td>Met some but not all responsibilities</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Unsatisfactory</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>100%</td>
</tr>
</tbody>
</table>
Evidence of Impact

This section discusses how accomplishments may translate into impact. We asked Partner respondents and Fellows questions about several types of impacts (see Table 7).

Table 7: Impact Measures

<table>
<thead>
<tr>
<th>Impact Category</th>
<th>Measure</th>
<th>% who reported impact by Fellowship</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Partner</td>
</tr>
<tr>
<td>Organizational impact</td>
<td>Service expansion</td>
<td>38%</td>
</tr>
<tr>
<td></td>
<td>Efficiency of operation</td>
<td>57%</td>
</tr>
<tr>
<td></td>
<td>Service quality</td>
<td>47%</td>
</tr>
<tr>
<td>Attitudinal impact</td>
<td>Motivation &amp; pride in work</td>
<td>71%</td>
</tr>
<tr>
<td></td>
<td>Self-confidence</td>
<td>76%</td>
</tr>
<tr>
<td></td>
<td>Time management</td>
<td>86%</td>
</tr>
<tr>
<td></td>
<td>Strategic thinking</td>
<td>67%</td>
</tr>
<tr>
<td>Leveraging impact</td>
<td>Financial resources</td>
<td>33%</td>
</tr>
<tr>
<td></td>
<td>Material resources</td>
<td>43%</td>
</tr>
</tbody>
</table>

For organization impact, includes responses "a little" and "a lot"
For attitudinal impact, includes responses "moderate", "high" and "very high"
For leveraging impact, includes "yes" responses.
% Diff = (%Fellows – %Partners) / % Partners

At the level of organizational impact, we asked whether the Fellowship had had an impact on service expansion, efficiency, and quality of service. Looking at impact on staff who worked with the Fellow, we asked about how the Fellowship had affected motivation, self-confidence, and work attitudes. Finally, we asked about whether the Fellowship had been able to leverage additional funds and resources which the organization might not otherwise have had access to. For most of the organizational and attitudinal impact measures, Fellows reported higher impact than Partner respondents, except in the work attitude of “time management” where Partner respondents perceived greater impact than Fellows. The last column of Table 7 shows this trend as the percent difference between Fellow and Partner responses, all of which are positive except for the Time Management impact category. The sample size is too small for us to conclude that these are statistically significant differences.

Service expansion. Because the goal of the Pfizer Global Health Fellows program is to improve health in developing countries, we wanted to see if we could find evidence that stronger organizational capacity actually increased delivery of services. In fact, only 3 Partner respondents and 3 Fellows claimed that they had a lot of evidence of a measurable link between the activities of the Fellow and service expansion. More than 50% of Partner respondents and three-quarters of the Fellows said that effects on service expansion are not measurable.

In interpreting these results, it is important to remember that while all the Fellowships were Round 5 or 6, the recall periods varied. For Fellowships that had just recently been completed, the timing might not have permitted any observed impact on these organizational measures.
In addition, the types of organizations varied in terms of their service orientation. Some Fellows were placed in organizations charged with conducting clinical trials of vaccines: this type of organization might define “service delivery expansion” differently from a hospital laboratory department, or a micro-credit organization. We asked each respondent to consider what service delivery meant to their organization, in light of their own organizational mission and goals. The understanding of efficiency and service quality may be more standard and comparable among organizations, though even these measures could be difficult to interpret out of context.

Efficiency and Quality. More than half of the Partner and Fellow respondents felt that the GHF activities had resulted in greater efficiency. For example, one Partner said “Things are more efficient now because of the schedule she left and the clear SOPs…We now know the refrigerators won’t break down, so we have more storage space available. Also we have less [work that cannot be done] due to machines breaking down.” Another mentioned how the accounting system had improved, enabling reports to be produced more quickly and easily. Forty-seven percent of the Partner respondents, and 63% of the Fellows thought service quality had been improved. One Partner respondent mentioned how financial reports are now more accurate, while another said the Fellow had helped them pass a higher-level licensing inspection and perform work with fewer errors. Annex 2 provides additional examples from Partner respondents.

The following quote from a Partner respondent gives a flavor for how one Fellow, working in a clinical trial management setting, helped to improve efficiency:

The GHF worked with a relatively new staff member, who didn’t know his job very well….After all the time he spent with the fellow and after receiving her mentoring, we’re very happy with this person’s performance….It is much easier to classify problems, address them, and follow them….She created new procedures [and log books]. She introduced the idea that the monitoring would start by looking at the last feedback letter, and making sure that those identified problems had been addressed and documented. If they hadn’t been, then the team would have a chance to address them while the monitor was present, but it had to be documented. It is very hard to give criticism in a way that causes action without resentment. But this is what the GHF was very good at, and we learned how the systems can support this.

A Partner respondent who is in charge of maintaining the computer hardware and software in her organization said that the GHF improved her productivity: “I get less requests for help now, which frees up my time to do other things. Also, people don’t have to wait so long for help anymore, and there are other staff who are trained to help them.” In another organization, the Partner respondent described time savings of several person-days per month because of a payroll system and fund accounting module that had been installed.

Some Fellows expressed uncertainty about improvements in efficiency. “I don’t know; unless I go there and measure it, I cannot say,” said one Fellow. Another Fellow wasn’t
sure if her work on financial systems had helped: “did I make a difference, or just generate spreadsheets?”

Several Partner respondents claimed quality had improved due to new quality control procedures implemented with the Fellow’s assistance. One thought there were fewer errors on review, however few could provide evidence of these improvements. One Fellow described improvements in information systems management goals of an organization:

“We disseminated this information and it has been transparent, which has helped to build trust [among development partners]. The effect of this is hard to quantify, but it is so important. We used everyone’s input to create higher quality relationships and better political rapport. And politics, well, that’s how you get things done.”

To provide some level of consistency and a check on social desirability bias, we decided to have the interviewer comment on his or her agreement or disagreement with the Fellow’s or the Partner Organization respondent’s assessment of impact, based on the availability of evidence to document the impact. In about ¼ to 1/3 of the cases, the interviewer disagreed or agreed only in part with the respondent’s rating of impact. Most of the time the interviewer felt the impact was lower, or that too little evidence existed to be able to say one way or the other. The quotes from interviewers, below, show how this question was sometimes difficult to answer:

The question’s focus on evidence [for impact] is a bit awkward. I agree with the Fellow’s answer that there is an indirect impact on service expansion. The Fellow helped to structure the roll-out plan and transformed it from chaotic to strategic. She said that the expansion would have happened without her, but it would have exploded. So the evidence of impact is in the improved structures and communications—the strategic thinking that lay the foundation for the expansion.

I agree with the Fellow’s answer…Since the clinical trials hadn’t started by the time she left, there was little or no evidence of service expansion or efficiency, but the potential was there. I would probably [disagree with the Fellow and] say there is little evidence for quality improvement, but she is basing her answer on knowledge of SOPs and counterpart skills that she has developed, rather than the outcome [the clinical trials].

The Fellow saw a little evidence for improvement in quality of service, but I think it’s very hard to make the connection because the fellowship had to do with internal personnel systems improvement, so the trickle down effect is hard to measure.

Leveraging Resources. About a third of the partners stated that Fellows had leveraged financial resources in the course of their time in country. Resource leveraging included the $500 to $1,000 Pfizer fund available to most Fellows for project-related activities.13

13 The enhancement fund is not normally available for Fellows working in multilateral organizations or Governments; however, as government-based fellowships are often organized and facilitated through an international NGO, there are occasions when the enhancement fund may be tapped. Some Fellows reported
Also some Fellows reportedly fund-raised among their friends and colleagues. One Fellow raised cash donations which were used to ship 2 cartons of specialized new books for a library. A sequence of two Fellows with management and finance expertise leveraged a grant from a Pfizer country office for $30,000.

But it is in the leveraging of material resources that Fellows achieved some surprisingly significant results. One Fellow noted that very few computers existed in her organization, so she leveraged the donation of 10 laptops (used) from a Pfizer division. Another Fellow managed to obtain donated used laboratory equipment and six laptop computers valued at $90,670. Other Fellows donated stationary, pens, umbrellas, and reference books. Two Fellows worked with orphans (outside their Fellowship responsibilities) and leveraged contributions from friends and family, including Christmas presents and clothing. Overall, according to the Partner respondents 9 Fellows (43%) leveraged material resources for Partner Organizations or other community needs. Responses from Fellows were consistent with these findings.

*Other impacts.* An open-ended question asked respondents to describe any other impacts from the Fellowship, not captured in the answers to the previous questions. One Partner respondent mentioned how the Fellow had helped to resolve a communication gap between two services, stating that “Sometimes it takes an outsider to bring people together.” Other insights included new ways of thinking, other viewpoints, or work style issues. For example, a laboratory technician described how the Fellow brought about an “appreciation of the role of preventive maintenance; that commitment to a program of preventive maintenance keeps equipment running and sustained.” Another said he “gained insight into the way a pharmaceutical company thinks, especially when developing a drug and about the patent and licensing process. We gained an understanding of the industry viewpoint.” Given the clinical research setting, this viewpoint is relevant to the Partner Organization’s work. Finally, a new manager who worked alongside a Fellow for three months struggled to capture in words the insights gained and applied: “If we couldn’t schedule a meeting when he wanted, he would compromise. But he wouldn’t forget! He would keep reminding us that we had to have the meeting.”

**Sustainability and Attribution**

The impact of an international corporate volunteering program also brings up questions of sustainability and attribution. How sustainable are the accomplishments achieved during the Fellowship, and how much of the impact can really be attributed to the Fellowship?

To examine sustainability, in addition to the earlier questions about potential impediments to implementation and sustaining changes, we asked partner organization respondents whether the personnel were demonstrating use of materials or skills imparted at the time the Fellow left. Almost all (94%) said that skills were being used at the time applying too late to receive funds through this source, or did not apply because they felt their scope of work was achievable without additional funds.
the Fellow left. Interviewers noted evidence to support the Partner respondent’s assessment in about half the cases (48%). When we asked Fellows the same question, 89% (17 out of 19) said that skills/materials imparted were being used at the time the Fellow left. According to interviewers, there was evidence to support this in most cases (14 out of 17).

The concern about attribution—how much change is really due to the Fellowship—arises because sometimes a Fellow provides an “extra pair of hands” to accomplish tasks that would have happened anyway. While this has a benefit, it is not necessarily worth the large expenses associated with this type of program. We asked questions to measure the importance of the Fellow in driving the resulting improvements or changes. These questions were asked only to the Partner organization respondents.

All Partner Organization respondents stated that the GHF had helped to speed up changes, while 80% thought the GHF had contributed new ideas. About 70% thought the outcomes would probably not have happened if the Fellow had not come, while close to three-quarters of respondents reported that they would have had to pay someone else to do the work the Fellow did. It isn’t clear whether they would have prioritized the work or had resources to pay someone else; however, these findings do suggest the Partners assigned the Fellows to do work which was of significant value to the organization.

Finally, research team members shared observations based on the interviews conducted. Topics cover several areas, including the impact of multi-fellow commitments, length of Fellowships, and measurement of sustainability.

1. Multi-fellow commitment to organizations seems to be successful in helping NGOs grow in size and capabilities. For example, at the International AIDS Vaccine Initiative (IAVI), an international NGO, and Mothers to Mothers-to-be (M2M2B), an indigenous NGO, the Fellows were part of a multi-fellow commitment. At M2M2B, the succession of Fellows helped build the financial system that enabled the organization to grow from a very small NGO (with 2 employees) to one that now gets multi-million grants from the President’s Emergency Plan for AIDS Relief (PEPFAR). At IAVI, the Fellows moved into a similar environment to the ones they worked in at Pfizer. They also worked with Supervisors with similar backgrounds/similar job descriptions. This clearly improves communication, as the supervisors feel comfortable scoping tasks consistent with high levels of skill and experience in the Fellows. According to one respondent, IAVI Johannesburg was so impressed that they considered asking for six GHF in the next round.

2. In five Partner Organizations, including some of those who were most enthusiastic (M2M2B, IAVI, AORTIC), respondents were very clear that a Fellowship should be at least six months. Some (AORTIC) would have liked to have the Fellow even longer.
3. Assessing sustainability is challenging, especially if the evaluation takes place very soon after the Fellow has left the country. For example, one supervisor was very enthusiastic about the Fellow’s work---getting an IT application up and running. However, it is too soon to tell if this Fellow’s intervention will result in more rapid implementation. The Fellow’s work involved an internal system whose relation to increased output/efficiency is not clear, even though the intervention was done very well by the GHF.

Fellowship Types and Impact Scores

The first Pfizer GHF evaluation (2006) used these definitions for Fellowship types:
1. Fellow provided volunteer assistance not requiring his/her professional training and not building on expertise acquired through employment at Pfizer. Clients of the partner organization benefited from Fellow’s efforts, but no permanent change in the ability of the partner organization or its staff to deliver services
2. Fellow provided technical expertise or training, based on his/her professional training or Pfizer acquired expertise, which resulted in upgrading the skills of staff in the partner organization. Volume of service unaffected, but quality may be improved.
3. Using his/her professional skills and Pfizer acquired expertise, Fellow worked with counterparts to introduce an operational or managerial improvement that will result in expanding the output of the organization.

As shown in Table 8, using these typology definitions, about 33% of Fellowships were classified by Partners as being Type 2 and 67% were Type 3. This is interesting in light of the fact that only 15% agreed that the Fellowship had actually expanded services. There were no Type 1 fellowships identified by Partner Organization respondents.

Fellows’ ratings were somewhat similar, with a distribution slightly weighted toward lower impact assessment. As a check on reliability of classification, in interviews with both types of respondents we asked interviewers to assess the Fellowship based on their understanding of the Fellowship impact typology. The interviewer’s impression was sometimes not consistent with the classification category chosen by respondents.

<table>
<thead>
<tr>
<th>Table 8: Fellowship Types, Scored by Respondent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Type 1</td>
</tr>
<tr>
<td>Between 1 and 2*</td>
</tr>
<tr>
<td>Type 2</td>
</tr>
<tr>
<td>Type 3</td>
</tr>
</tbody>
</table>

*Criteria for Type 1 included a) assistance not requiring professional training, and b) leading to no permanent change in ability of Partner to deliver services. One Fellow respondent and one interviewer were reluctant to use the categories given, and elected to classify a Fellowship as “Between 1 and 2” indicated that Fellowship met one criteria for Type 1 but not both criteria. This classification option was not given to all respondents, however.

**Includes 3 respondents who felt the fellowship was “between a 2 and 3”. The reasons for not being able to classify fellowship as either 2 or 3 were not recorded.
The difficulty in categorizing a Fellowship as Type 1 is that several criteria must be met, including the use of non-expert assistance. We found two Fellowships that met some Type 1 criteria (i.e. clients benefited, but no permanent change in organizational capacity) but still used a Fellow’s professional expertise at least to some extent.

Another challenge encountered is how to assess impact type in organizations with a multi-fellow commitment. For example, at one organization (M2M2B), the combination of four Fellows over time clearly made possible the creation of a financial system to permit major growth. Should we score all contributing fellowships at the “bottom line”, which is a 3, or attempt to differentiate based on the preparation or contribution of individual fellows?

The lack of consistency in results and other difficulties in applying the typology suggested to the research team that further refinement was needed. We therefore created the Impact Index discussed below.

C. Impact Index

The 9 impact measures presented in Table 7 are one way to report on Fellowship success in strengthening the capacity of a Partner Organization. But in order to compare fellowships, we wanted to see if we could summarize the 9 impact measures further. To do this, we created a summary score for each of the three impact categories, and calculated an overall impact score for the fellowship as a whole.

We used two different techniques to do this. The first method was to simply sum the scores for each response for Partners and Fellows, while the second method used a statistical technique called factor analysis to create the composite impact measures. Each method has pros and cons, which is why we present both sets of results below. Because the results were similar, we recommend using the first method—the sum of scores—in the future, for transparency and simplicity.

Sum of Scores Method
To calculate this score, we added up the respondent’s answers to the questions within each of the three categories of impact. The overall range of possible scores using this method is 9 (best score, indicating highest impact) to 33 (lowest level of impact). Table 9 shows the range of scores overall and by category of impact.

For example, the organizational impact variables (service expansion, efficiency of operation, and service quality) were all scored on a 3-point scale with a 1 indicating “a lot of evidence”, a 2 indicating “a little evidence”, and a 3 for “no evidence”. The best score a fellowship could receive for organizational impact variables was 3 (most impact), and the lowest score was 9 (least impact). The attitudinal impact variables (staff’s motivation, self-confidence of staff, time management of staff, and ability of staff to think strategically) were scored on a 5 point scale with 1 representing “very high” impact of fellowship; 2 was “high” impact of fellowship, 3 was “moderate”, 4 was “low”
and 5 was “very low”. The best score for attitudinal impact was 4 (highest impact) and the lowest score was 20 (lowest impact). The leveraging impact variables (financial resources, and material resources) were scored as simple 1 is “yes” and 2 is “no”. The best score for leveraging impact was therefore 2 (highest impact) and the lowest score was 4 (lowest impact).

Table 9: Sum of Scores Method: Range of Possible Scores by Category

<table>
<thead>
<tr>
<th>Number of items measured</th>
<th>Organizational Impact</th>
<th>Attitudinal Impact</th>
<th>Leveraging Impact</th>
<th>Overall Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>Original Scale used</td>
<td>3-point scale:</td>
<td>5-point scale:</td>
<td>2-point scale:</td>
<td>NA</td>
</tr>
<tr>
<td>to measure individual</td>
<td>1=a lot of evidence</td>
<td>1= very high</td>
<td>1=yes</td>
<td></td>
</tr>
<tr>
<td>items</td>
<td>2=a little evidence</td>
<td>2= high</td>
<td>2=no</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3=no evidence</td>
<td>3= moderate</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4=low</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5=very low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range for total score on</td>
<td>3 to 9</td>
<td>4 to 20</td>
<td>2 to 4</td>
<td>9 to 33</td>
</tr>
<tr>
<td>all items (from high to</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High Low</td>
</tr>
<tr>
<td>low)</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Median value for this</td>
<td>6</td>
<td>12</td>
<td>3</td>
<td>21</td>
</tr>
<tr>
<td>index</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 10 presents the GHF program fellowship data, scored by this index.

Table 10: Sum of Scores Impact Index by Fellowship, Scored by Partners and Fellows

<table>
<thead>
<tr>
<th>Fellow ID</th>
<th>Partners</th>
<th>Fellows</th>
<th>Attitudinal Impact</th>
<th>Leveraging Impact</th>
<th>Overall Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>4</td>
<td>9</td>
<td>7</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>5</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>C</td>
<td>5</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>D</td>
<td>7.5</td>
<td>8</td>
<td>10.5</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>E</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>F</td>
<td>7</td>
<td>6</td>
<td>11</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>G</td>
<td>6</td>
<td>4</td>
<td>11</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>H</td>
<td>7.5</td>
<td>4</td>
<td>12.5</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>I</td>
<td>5</td>
<td>7</td>
<td>11</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>J</td>
<td>6</td>
<td>6</td>
<td>12</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>K</td>
<td>7</td>
<td>6</td>
<td>16</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>L</td>
<td>7</td>
<td>6</td>
<td>15</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>M</td>
<td>7</td>
<td>7</td>
<td>16</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>N</td>
<td>9</td>
<td>6</td>
<td>16</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>O</td>
<td>8</td>
<td>7</td>
<td>18</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>Average</td>
<td>6.7</td>
<td>6.6</td>
<td>11.9</td>
<td>8.2</td>
<td>3.2</td>
</tr>
</tbody>
</table>

Note: We have not included fellowships where the Partners’ data were incomplete. This results in 15 observations. Blank data could be due to answer of “NA” on some measure, item left blank, or if Fellow respondent was not available for that Fellowship. Overall impact score cannot be calculated if some data are missing. Data include two separate records for one Fellow who worked half-time with two distinct departments within a hospital: each of these experiences was separately evaluated by a different Partner respondent.
To illustrate how the data can be interpreted, we can examine the first fellowship in Table 10, with Fellow ID “A”. Note that for each category of impact, there is a separate column for Partners and for Fellows. This indicates whether the data come from the questionnaire administered the Partner respondent, or the questionnaire answered by the Fellow.

Considering just the Partners’ column for the moment, the Organizational Impact score for Fellow ID “A” is 4 on a scale of 3 (highest) to 9 (lowest). This means that the partner respondent reported “a lot of evidence” for 2 of the 3 categories and “a little evidence” for the 3rd category. This could be interpreted that the partner considered the fellowship to have had an impact on the organization. The partner respondent for Fellow ID “A” also gave a score of 7 for attitudinal impact. There are 4 attitudinal impact variables and the highest (best) score for attitudinal impact overall would be a 4, lowest would be 20. In this case, it is impossible to determine whether the partner gave three “Very highs” and a “Moderate” or 1 “very high” and 3 “high”. Yet, we can still infer that a score of 7 is well above the median for this category.

Leveraging impact has only 3 possible scores: 2, 3, or 4. A score of 2 represents a fellowship which leveraged both funds and material resources, while a 4 represents a fellowship which did not leverage either funds or material resources. A score of 3 represents a fellowship which leveraged either funds or material resources, but not both. For Fellow ID “A”, the Partner respondent score is 2, which indicates this fellowship leveraged both funds and material resources. (Note that there are some differences in perception between the Partner respondent and the Fellow, regarding the extent of leveraging.)

The overall impact score given by the Partner respondent for Fellow ID “A” is 13, out of a range of 9 (highest impact) to 33 (lowest impact).

By summing these we were able to develop a score which is easy to calculate and easy to interpret: a low number would represent greater impact than a higher number. The disadvantage to this ranking system is that it automatically provides a weighting system based upon the scale of the questions asked. For example, if a fellow did not leverage any resources then they would be given a score of 4 for the leverage impact index; however, if they had “very low” impact on all of the attitudinal variables, then they would be given a score of 20. The 20 would pull down the total score much more than the 4 would raise it. This example illustrates that the weighting of these variables by the sum of the scores implicitly weights the attitudinal variables the most, followed by the organizational variables and finally by the leveraging variables due to the number of the variables making up the index and the possible scores.

It must also be noted that this method for evaluating impact does not account for the primary goals of the fellowship, or how well the organization was already functioning in any of these categories. For example, a fellowship may not have been designed to work on expanding services, so there was no impact on this dimension. It also might have been the case that self-confidence of staff was already very high, which would mean that the
fellows could not show much impact on this variable. This is not accounted for in the index, and is a limitation to this methodology.

**Factor Analysis Method**
To try to avoid the weighting problem implicit in the sum of scores methodology, we also used a different method for calculating indices. This method, using the statistical technique of factor analysis, calculated indices for each of the three impact categories and for the impact measures overall, and then compared the rankings of the fellows and the partners as shown in Table 11.

**Table 11: Factor Analysis Impact Score Index by Fellowship Scored by Partners and Fellows**

<table>
<thead>
<tr>
<th>Fellow ID</th>
<th>Partner Organizational Impact</th>
<th>Fellow Organizational Impact</th>
<th>Partner Attitudinal Impact</th>
<th>Fellow Attitudinal Impact</th>
<th>Partner Leveraging Impact</th>
<th>Fellow Leveraging Impact</th>
<th>Partner Overall Impact</th>
<th>Fellow Overall Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.3</td>
<td>4.2</td>
<td>1.1</td>
<td>1.0</td>
<td>2.3</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>B</td>
<td>1.9</td>
<td>3.2</td>
<td>1.4</td>
<td>2.7</td>
<td>2.3</td>
<td>1.0</td>
<td>1.3</td>
<td>2.5</td>
</tr>
<tr>
<td>C</td>
<td>1.9</td>
<td>.</td>
<td>1.1</td>
<td>.</td>
<td>2.4</td>
<td>.</td>
<td>1.5</td>
<td>.</td>
</tr>
<tr>
<td>D</td>
<td>3.8</td>
<td>3.9</td>
<td>1.8</td>
<td>3.0</td>
<td>1.0</td>
<td>3.1</td>
<td>2.0</td>
<td>3.2</td>
</tr>
<tr>
<td>E</td>
<td>3.4</td>
<td>4.2</td>
<td>1.2</td>
<td>.</td>
<td>2.4</td>
<td>3.1</td>
<td>2.2</td>
<td>.</td>
</tr>
<tr>
<td>F</td>
<td>3.9</td>
<td>2.6</td>
<td>2.2</td>
<td>2.7</td>
<td>3.7</td>
<td>2.1</td>
<td>2.2</td>
<td>2.8</td>
</tr>
<tr>
<td>G</td>
<td>2.4</td>
<td>1.6</td>
<td>2.1</td>
<td>2.2</td>
<td>3.7</td>
<td>1.0</td>
<td>2.3</td>
<td>2.2</td>
</tr>
<tr>
<td>H</td>
<td>3.8</td>
<td>1.6</td>
<td>2.3</td>
<td>2.1</td>
<td>1.0</td>
<td>1.0</td>
<td>2.3</td>
<td>2.1</td>
</tr>
<tr>
<td>I</td>
<td>1.0</td>
<td>3.6</td>
<td>2.3</td>
<td>.</td>
<td>3.7</td>
<td>3.1</td>
<td>2.5</td>
<td>.</td>
</tr>
<tr>
<td>J</td>
<td>1.6</td>
<td>2.3</td>
<td>2.3</td>
<td>.</td>
<td>3.7</td>
<td>3.1</td>
<td>2.6</td>
<td>.</td>
</tr>
<tr>
<td>K</td>
<td>2.2</td>
<td>2.3</td>
<td>3.5</td>
<td>.</td>
<td>1.0</td>
<td>.</td>
<td>3.4</td>
<td>.</td>
</tr>
<tr>
<td>L</td>
<td>2.2</td>
<td>2.6</td>
<td>3.3</td>
<td>4.5</td>
<td>3.7</td>
<td>3.1</td>
<td>3.4</td>
<td>4.5</td>
</tr>
<tr>
<td>M</td>
<td>2.1</td>
<td>2.9</td>
<td>3.5</td>
<td>1.3</td>
<td>2.4</td>
<td>3.1</td>
<td>3.8</td>
<td>1.3</td>
</tr>
<tr>
<td>N</td>
<td>3.4</td>
<td>2.3</td>
<td>3.5</td>
<td>.</td>
<td>2.4</td>
<td>.</td>
<td>3.9</td>
<td>.</td>
</tr>
<tr>
<td>O</td>
<td>2.8</td>
<td>3.3</td>
<td>4.2</td>
<td>4.2</td>
<td>2.4</td>
<td>1.0</td>
<td>4.4</td>
<td>4.1</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>2.5</strong></td>
<td><strong>2.9</strong></td>
<td><strong>2.4</strong></td>
<td><strong>2.6</strong></td>
<td><strong>2.5</strong></td>
<td><strong>2.2</strong></td>
<td><strong>2.6</strong></td>
<td><strong>2.6</strong></td>
</tr>
</tbody>
</table>

Note: Fellowships where Partner respondent information was incomplete are not included. Blank data could be due to answer of “NA” on some measure, item left blank, or if respondent was not available for that Fellowship. Overall impact score cannot be calculated if some data are missing.

When factor analysis is used, the scores are adjusted to a five-point scale ranging from 1 to 5, where the 1 would represent the highest impact and the five the lowest impact. For both the Partners and Fellows, the Attitudinal Impact score was most closely correlated with Overall Impact. This means that Attitudinal Impact was seen as the most significant impact of the fellowships. The overall impact scores of the Partners and Fellows show no correlation.

The advantage to using factor analysis in the construction of the indices is that it avoids the problem of “double counting”, or the implicit weighting which occurs because some categories of impact have more questions than other categories. Double counting could occur in the summation index that we created above if we have two closely related questions which would have similar responses. For example, if we found that everyone who leveraged funds also reported that they leveraged resources, and vice-versa, then by
summing the responses, we would essentially be double counting. While none of the variables in this data are perfectly correlated, there are a number of variables that are closely correlated. Using factor analysis to create the index reduces the problem of double counting. On the other hand, factor analysis has a number of problems: it is not easy to apply and requires a statistician, and the factor that it develops (which is used for the scaling) is unique for the specific data set that was used. Finally, factor analysis is not transparent and may not be easily understood by the people who want to use the index to draw conclusions about program effectiveness, or to plan for program improvements.

Recommended Method: Sum of Scores
In comparing the factor analysis index to the sum of scores index, we found that the two indices for the partners were extremely close. They had a correlation of 0.97, meaning that there was essentially no difference between the two. For the two separate indices developed using the fellows data, there was still a high degree of correlation (0.80) although the lower number demonstrates that there was greater disagreement between the two indices using fellows data than when using partners data.

Given that there is such a high degree of agreement between the indices developed by two separate methods, we chose to use the sum of scores index. Aside from closely agreeing with the index developed using factor analysis, this index has the advantage of simplicity and transparency in its application.

For application in the Pfizer GHF program, we propose using the scale shown in the first two columns of Table 12 to rank the impact of a fellowship. As shown in Table 12 and Figure 5, when we apply this ranking to the data we obtained during this study, fellowships rated by Fellows fell into higher impact categories than the fellowships as rated by the Partners, with no Fellows-rated fellowships in the low impact category, while 3 Partner respondent-rated fellowships fall into the low impact category.

<table>
<thead>
<tr>
<th>-ranking of fellowships by Partners and Fellows</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ranking of Fellowship</td>
</tr>
<tr>
<td>--------------------</td>
</tr>
<tr>
<td>Very High Impact</td>
</tr>
<tr>
<td>High Impact</td>
</tr>
<tr>
<td>Moderate Impact</td>
</tr>
<tr>
<td>Low Impact</td>
</tr>
</tbody>
</table>
VI. Discussion

This research study collected and analyzed information from Fellows and Partner respondents to evaluate the impact of Pfizer’s ICV program. Our study looked only at one aspect of impact evaluation, i.e. the impact on the recipient institution’s organizational capacity, including its success in expanding delivery of high quality, efficient services.

In developing our ICV Evaluation Toolkit, we considered several alternatives for measuring impact. One alternative was to measure **performance in comparison with objectives** of the Fellowship. While this method can work in some cases, it runs into problems when the scope of work is changed over time in order to be responsive to needs identified later in the Fellowship. Adaptation of scopes of work results in better technical assistance, but it also complicates the evaluation process.

Similarly, we found it hard to measure impact by **documenting accomplishments with quantifiable indicators**. This method was not able to capture all types of impact. For example, some tasks could be measured by indicators such as increased number of HIV+ patients on antiretroviral drug treatment, a direct impact measure. More often, though, the scope of work involved tasks which were less directly related to impact yet still important for capacity strengthening, such as the development of a plan. This type of task is measured by a binary (yes/no) indicator, which says little about the quality or usefulness of the plan.

The focus on quantifiable indicators appeared to be limiting in a cultural sense as well. One Partner respondent resisted the questions about quantifiable objectives, stating that this approach was “very American…we don’t really think that way.”
Finally, we tested the reliability of asking respondents to categorize Fellowship impact according to the typology developed in the previous Pfizer GHF evaluation. With these Fellowships, as with the first Pfizer evaluation, most Fellowships were rated a Type 2 or Type 3 (Type 2 showing impact on quality or efficiency, and Type 3 showing impact on service expansion); however, we found that Partner Organization respondents and Fellows had different opinions as to the Type which should be assigned to a Fellowship. We also found somewhat inconsistent results in assignment of type by the Interviewer compared to either the Partner respondent or the Fellow.

Given the problems we encountered with other models, we found that the most feasible approach was to evaluate impact using a process model with multiple measures. We adapted the LBG model (input ➔ output ➔ outcome) to include more contextual information, as well as measures of sustainability and attribution, or how much of the impact was due to the Fellowship. Figure 6 presents this model and the different elements measured.

Some issues in using this type of model for impact evaluation include reliability, quantification, and unit of analysis. To overcome reliability concerns, we used triangulation to gain a better sense of impact, collecting data from both the Partner respondent and the Fellow and asking the interviewer to give his or her opinion based on the data collection. Comparison of Fellow and Partner information, in particular, provides a fuller picture, although due to the small sample in this study we were not able to assess reliability using statistical tests.

Applying this evaluation model in the future, corporations should make an effort to probe the reasons for differences between Fellow and Partner informant perceptions. One way to do this is to create a small working group of program participants, to review the findings for possible misunderstandings or biases. In the model presented in Figure 6, we suggest what we feel is the best source for the data if it is not possible to collect both.

A second issue in using this model is quantification of impact. As shown in Tables 10-11 and Figure 5, we propose using a sum of scores index to capture impact in a single measure.

In addition, the methodology for evaluation also uses descriptive information on accomplishments which don’t easily fit into frequency tables or statistics. Although not easy to summarize or compare, we feel that descriptive data on accomplishments and impacts are still important and can help in the design of future fellowships. This type of information is already contained in the end-of-service report that Fellows complete for Pfizer, although collecting it from Partner respondents may highlight what the receiving organization felt were the most important outputs.

Finally, regarding unit of analysis for the evaluation, it is important to define the boundaries of the organizational unit. In other words, who or what is the “Partner organization”? Whose capacity is being strengthened through the Fellowship?
In one Fellowship, the Fellow was assigned to a local NGO whose mission was to assist small business development. The Fellow helped a single entrepreneur develop a business plan related to an agricultural-sector product (indirectly linked to health as it relates to food security and poverty reduction). Although the Partner respondent interview was held with the Fellow’s supervisor from the local NGO, the “organization” strengthened was really the entrepreneur and his business. In another example, there was some initial confusion over whether the “Partner organization was the hospital as a whole, or the particular department where the Fellow was assigned. All respondents and the interviewer must have a common understanding of the unit of analysis in order to respond reliably to questions such as assessment of baseline organizational development, changes in staff motivation and work attitudes, etc.
Figure 6: Input-Accomplishments-Impact Model for Evaluating International Corporate Volunteering Programs

**Contextual Factors**
- Baseline Level of Development of Partner Organization
- Age/Type of Partner Organization
- Prior Experience with Volunteers

**Inputs**
- Fellow Characteristics
  - Flexibility
  - Preparation
  - Experience
  - Skills
- Fellowship Characteristics
  - Length
  - Level of interaction
- Partner Characteristics
  - Understanding of SOW
  - Time, financial resources, and policy flexibility to execute changes

**Accomplishments**
- 9 Technical Areas
  (% with any accomplishment; description)
- Needs Assessment
  (% having done one or more; description)

**Impact**
- Change in level of organizational development of Partner
- Service Expansion
- Increased efficiency
- Higher quality
- Leveraging of resources/cost savings
  - Financial
  - Material
  - Would have had to pay if no Fellow
- Change in attitudes of staff
  - Motivation/pride in work
  - Self-confidence
  - Time management
  - Strategic thinking

**Sustainability and Attribution:**
- Still using skills/materials imparted at time Fellow left
- Fellow contributed new ideas, sped up change
- Would not have happened without Fellow
- Interviewer assessment: evidence of sustainability?

Assessed by:
- ◇ = Partner and Fellow
- ■ = Partner only
- ☐ = Fellow only
VII. Recommendations

A. Use process model to understand ICV program impact.

An evaluation strategy which includes multiple measures to examine different dimensions of performance is likely to be more useful for making decisions regarding program improvement. Multiple measures also allows for the collection of contextual information which is important for interpreting impact measures, and for measuring sustainability and attributing how much of the impact is due to the Fellowship. The Boston University ICV Evaluation Toolkit, provided under separate cover\textsuperscript{14}, includes questionnaires which capture many different dimensions of performance.

B. Use impact index and rankings to track progress over time.

Using the composite impact index which we have proposed allows ICV programs to summarize program impact according to measures that are weighted to reflect program priorities. It also provides a way to compare the perceptions of Fellows and Partner respondents, and to monitor the progress of the ICV program over time. The Boston University ICV Toolkit includes a spreadsheet for calculation of the composite impact index.

We believe that the weights implicit in our composite index are appropriate for the Pfizer Global Health Fellows Program. For example, organizational impact (3 measures) and attitudinal impact (4 measures) are weighted more heavily than leveraging impact (2 measures), which accurately reflects the relatively low amount of emphasis which the Pfizer program puts on the latter dimension of performance. The rankings should be interpreted with caution, however, as they do not take into account the baseline capacity of the organization. Organizations with better capacity to begin with may show a lower impact from the fellowship.

C. Adapt tools for each program’s design and priorities.

The BU ICV Toolkit can be adapted for use by companies running or considering international corporate volunteering programs. Depending on the goals of the program, the factors and weights which are currently used in the index can be changed. These changes will affect comparability of data between companies, and should be explicitly noted.

\textsuperscript{14}Also available by contacting the lead author, Taryn Vian (tvian@bu.edu)
Some implementation issues which should be considered by both Pfizer and other corporations include adaptation of questions for each corporation’s specific program design; data collection strategies and the choice of on-site versus telephone interviews; and data analysis plans which incorporate statistical analyses and use of qualitative data.

1. Adaptation of Tools

The questionnaires have been calibrated using the data from the Pfizer Global Health Fellows Program. These tools are adaptable and may need to be adjusted for a corporation’s specific program design. For example, if scopes of work are less flexible than Pfizer’s program allows, the areas of technical accomplishment measured can be reduced.

For Pfizer, the tools should be reviewed with GHF staff, former Fellows, and representatives of Partner Organizations. Several issues that should be clarified have already been mentioned earlier in this report, i.e. definition of the unit of analysis (who is the “Partner organization”), meaning of certain terms used, such as “interaction with staff”, “flexibility”, or “preparation”, and what role the interviewer should play in confirming respondent perceptions by reviewing evidence of impact.

2. Data collection strategies

The pilot test of the BU ICV Evaluation Toolkit methodology used face-to-face interviews with Partner respondents, and phone interviews with Fellows. For Fellows, some of the data we obtained through phone interviews could actually be collected from routine reports submitted at the end of the Fellowship, although this assumes that information systems exist which can compile the different sources of data.

While on-site interviews seemed to work well for Partner respondents, this data collection strategy has cost implications, especially if the company’s employees are volunteering in a wide range of locations. In some cases, it may be possible to arrange phone interviews with Partner respondents. We interviewed one Partner respondent over the phone (three phone calls, one for each of the three Fellows he had supervised during this time period). That worked well, but it may not work in all cases. This issue should be discussed with a few key informants from Partner organizations to assess feasibility.

Other ways to economize on costs would be to use corporate field staff or regional placement NGO staff to collect data from Partner NGO informants.

3. Data analysis

In the BU ICV Evaluation Toolkit, we have included copies of the two questionnaires (Fellows and Partner respondents), as well as a scoring spreadsheet which can be used to calculate the overall fellowship rankings using the Sum of Scores methodology. For full
analysis of the Partner respondent and Fellow data, a statistical package may be needed (e.g. SAS or SPSS) to create frequency tables. If sample size is large enough, the analysis plan should measure whether differences in Partner and Fellow perceptions are significant. Other relevant factors which could be measured include correlations between perceptions of impact and Fellow characteristics (prior experience, preparation, flexibility) or Fellowship characteristics (length, level of interaction). As shown in the report, the qualitative data shed important light on the factors that might be barriers to achieving high impact Fellowships and are also important for capturing Partner respondents’ perceptions of accomplishments.

VIII. Conclusion

This study was designed and pilot tested as an evaluation toolkit to assess the impact of Pfizer’s ICV program on partner organizations. With adaptations to adjust for different program designs, the evaluation framework and tools can be useful to many public-private partnerships interested in documenting this aspect of program impact. The new evaluation toolkit found results that were quite similar to the more in-depth partner interview methodology used in a prior evaluation (with up to five key informants per Fellowship site). In both evaluations, we found that most Fellowships achieved many accomplishments leading to high levels of perceived impact on service quality, efficiency, and service expansion. This new toolkit provides a streamlined set of measures which allow for comparison of results between types of respondents and across Fellowship sites. This information is essential for continuous quality improvement of international volunteering programs so they can achieve maximum effectiveness and continue to strengthen the capacity of organizations to respond to the key public health challenges in global communities today.
Annex 1: Interview Contacts

FSG Social Impact Advisors, Boston, MA

Gregory J. Hills, Director, greg.hills@fsg-impact.org
Adeeb Mahmud, Consultant, adeeb.mahmud@fsg-impact.org

TNT International (Partnership TNT-WFP), Amsterdam, Netherlands

Henriette van Eeghen, henriette.van.eeghen@tnt.com

Innovations in Civic Participation, Washington, DC

Katherine Hutter, Projects Coordinator, hutter@icicp.org
Charlotte McDowell

Washington University at St. Louis

Dr. Margaret Sherraden, Center for Social Development
sherraden@umsl.edu
## Annex 2: Fellow Accomplishments by Technical Area, according to Partner respondents

<table>
<thead>
<tr>
<th>Area of Capacity Strengthening</th>
<th>Examples of outcome or accomplishments</th>
</tr>
</thead>
</table>
| Improvements to curriculum or training systems | Developed a clinical pharmacy training module as part of post-graduate diploma program graduating 6 pharmacists per year, plus schedule for continuing education for 120 pharmacy staff including pre- and post-tests (#16)  
Created a manual for troubleshooting IT problems in the office (#20) |
| Developed staff skills and competencies | Conducted a workshop on how to do clinical trial research, attended by 20 medical staff from eight departments in the hospital (#40)  
Staff competencies in using new lab instruments and in preventive maintenance of equipment for diagnosis and monitoring of infections in a research laboratory (#37)  
Conducted many training sessions related to topics on manufacturing, quality assurance, team building, and more. Also organized a study tour for staff to see other manufacturing companies and their QA processes. (#18)  
Training in how to prepare financial reports. Result is that they are more accurate, according to Finance Director (#9) |
| Leadership development or change in strategic direction | Helped think through the implications of a new strategy of approaching major philanthropists for donations to support institution (#17)  
The Fellow helped with IT design and a new procurement strategy. We choose equipment while thinking ahead to maintenance, now, not just price. (#20)  
The GHF brought leadership books which we read and discussed. This really helped us through the difficult times that we were experiencing. (#36)  
The GHF wrote the hospital strategic plan. Through the implementation of this plan, we will be able to transform to a semi-autonomous status and can offer better quality services than before. (#46) |
| New or stronger partnerships, networking, or new business development | The work done by the GHF was very appreciated by the international community here. We are now showing how we share work, and this has generated interest in what we do. (#33)  
Leveraged a Pfizer grant to one site ($30,000) (#9)  
Contacted organizations with similar interests; made preliminary contacts with drug companies. (#30) |
| Operational management efficiency and systems improvement | The Fellow helped establish connectivity between the field offices with the main office. This has improved communication and now everyone is well informed. (#20)  
We have staff meetings more regularly now. Also, the program has stronger collaboration and communication with other departments in the hospital. We didn't have this before. (#16)  
Things are more efficient now because of the schedule she left and the clear SOPs. For example, now we know the refrigerators won't break down so we have more storage space available. Also, we have less excused work due to machines breaking down. (#37)  
He advised on the organization of our referral network, i.e. that we needed to identify catchment areas for the different facilities, and then perhaps have pricing schedule to encourage people to seek care at the appropriate level. (#46)  
Wrote job descriptions, SOPs, and a site master file which resulted in upgraded license after inspection. The inspectors found no fault at all. He helped other departments preparing for ISO certification also. (#18)  
Number of patients enrolled on ARVs increased from 2,500 to 5,000. Could not have managed this growth without the new organizational and staffing plan. (#10) |
<p>| Strengthening financial | New plans for cost recovery and how we will set prices, who will be exempted. Also other sources of revenue for the hospital were identified. (#46) |</p>
<table>
<thead>
<tr>
<th>Area of Capacity Strengthening</th>
<th>Examples of outcome or accomplishments</th>
</tr>
</thead>
</table>
| stability or improving revenue generation | Based on work of the GHF, we decided to close our IV fluid manufacturing; this will result in cost savings and less wastage. (#18)  
Indirectly. The GHF helped improve revenue generation, because once people know their job description, job output can improve. (#39)  
Put in place a fund accounting system to enable organization to receive external money, report to donors, and pass audits (#28) |
| Strengthening marketing or public relations | The Fellow's work involved a market survey to determine factors motivating people to donate blood. This information has informed our choice of strategies and will help target programs to make volunteer recruitment more effective. (#19)  
The GHF helped us with internal relations, with the image we have with other departments. He was a very visible presence in the hospital (#18)  
Created new web site and helped with 'branding' of organization. (#30) |
| Strengthening information systems or technology | Strengthened our library information system; she also saw that we really had not enough computers, and she got donations of 10 laptops. (#36)  
Although not "IT" systems-related, the GHF helped to strengthen information about policies and procedures. This was important for communication and for productivity. (#18)  
Improved accounting system (#30) |
| Changes in research programs or systems | New clinical procedures for trial sites (#29)  
Fellow installed a resource-sharing (library) software that is used in many of the organization's country field offices for research. This helps us to share and draw on more resources for our projects. (#20)  
We are now able to do drug utilization reviews (#16)  
We got a baseline study report, and it was better quality because of what the GHF did. (#24) |
References


