4

Methodology

Approach

Approaches tested

A first attempt to assess the impact of the UNV programme was made by analysing the contents of the reports prepared by the Volunteers themselves on their assignments. The review covered 30 countries in Africa, Asia, the Pacific, Latin America, the Caribbean, the Middle East, and Europe. Unfortunately, even if a report should have been completed by each Volunteer, in many cases either a report had not been submitted, had been misplaced and could not be found, or had been submitted but was incomplete.[1]

In the case of Nepal, for instance, reports could be retrieved for only 59 of the 110 UN Volunteer assignments that had been completed during the period covered by the study (1987–96). This corresponds to 54 per cent, which was not considered a sufficient basis for an analysis of the impact of the UNV programme.[2]

Another even more serious problem with the review of the reports of the Volunteers as an approach is that little relevant information on the outcomes of the work of the Volunteers could be found in the reports. The reports address many other issues, particularly of an administrative nature, but they contain little on what has been achieved by the Volunteers. Still, a review of a total of 562 reports

was completed. The information contained in the reports was used to get an overview of the extent to which different kinds of training had been performed by the Volunteers, and the kinds of skills and knowledge that had been transferred by the Volunteers.

Another way to try to assess the impact of the UNV programme was to review almost 90 evaluation reports and assessments of the activities of the UN Volunteers that were available for the period covered by the study. An attempt was made to use meta-analysis to synthesize the findings of the 88 evaluations and reviews shown in Appendix C.[3]

Meta-analysis was considered an appropriate approach to try to assess the impact of the UNV programme since the evaluation reports covered a fairly representative sample of the activities of the UNV programme from different countries during the period covered by the study. What could not be anticipated was how little the evaluation reports contain in terms of results or impact of the work of the Volunteers. Predominantly, the reports are what could be described as input focused, process oriented, and problem centred.[4]

Some of the evaluation reports do contain examples of results of the work of the Volunteers, such as high pass rates of students taught by UNV teachers or improvements in the living conditions of the beneficiaries, but these examples are few and far between. Most of the reports only describe the activities of the Volunteers, such as the provision of on-the-job training or the introduction of income-generating activities, but do not describe the results of these activities. None of the evaluations specifically looked at the impact of the activities of the Volunteers some time after the Volunteers had left.

What the reports do contain, and what could be used in this study, was information about issues that had affected the work of the Volunteers, usually negatively. These issues were included in the conceptual framework of the study. Beyond that, the reports provided little except general views of government and UN officials on the work of the UN Volunteers and the performance of the head office of the programme, and nothing or very little on the perceived value of the programme to the beneficiaries. What became clear was that it would be necessary to collect original data to be able to assess

the impact of the UNV programme and its perceived value to the users and beneficiaries of the programme.

Methodology adopted

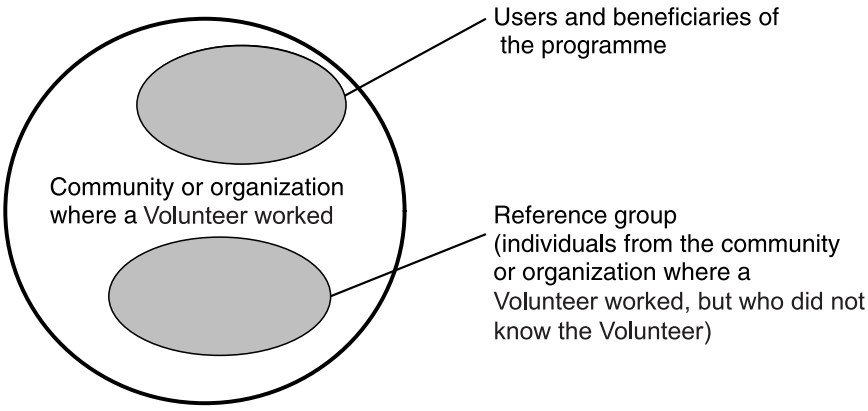
Even if the review of the reports prepared by the Volunteers and the evaluation reports yielded little information about actual results of the work of the Volunteers, these reviews provided the basis for the development of a survey to assess the impact of the UNV programme. Based on these reviews and the reports prepared by the Volunteers, a number of factors that could explain the outcomes of the work of the Volunteers were identified, and an inventory of issues to be included in a survey was developed.

The next step included interviews with government officials, representatives of non-governmental organizations, supervisors and co-workers of Volunteers, and beneficiaries of the UNV programme, in order to sharpen the focus of the study and to anticipate potential problems with the research design. The preliminary interviews were carried out during visits to Nepal, India, Bhutan, and Costa Rica.[5]

Following these exploratory visits, a decision was made to collect data through a mail survey addressed to former Volunteers and through interviews with individuals who had worked with or benefited from the work of the Volunteers. This included co-workers and supervisors of the Volunteers, and representatives of UN agencies, government departments, civil society organizations, and communities where UN Volunteers had worked. It was also decided to introduce a reference group in the study, in order to compare changes that could be attributed to the work of the Volunteers to changes that had taken place without any involvement of the Volunteers, that is, a "counterfactual situation".

During the exploratory visits it became clear that it would be impossible to identify respondents who in all other respects were similar to the users and beneficiaries of the programmes other than that they had not had any contact with the Volunteers. This was the reason that a control group, in the strict sense of the term, could not be introduced. The second-best, therefore, was to establish a refer-

Figure 4.1 Schematic illustration of the approach adopted



ence group consisting of respondents who did not know a particular Volunteer, but who knew the community or organization where the Volunteer had worked. The role of the reference group was to describe changes that had taken place in the communities or organizations during the time the Volunteer had worked there. A schematic illustration of the approach adopted can be seen in Figure 4.1.

The approach adopted entailed collecting information from specific communities and organizations where a Volunteer had worked. On the one hand, information was collected from individuals who had either personally benefited from the work of Volunteers or otherwise been in direct contact with the Volunteers (users and beneficiaries), and on the other hand, information was collected from individuals who had had no contact whatsoever with the Volunteers (reference group). In addition, information was collected from former UN Volunteers themselves.

Selection of Nepal as a case study

During exploratory visits to Nepal (November 1995), India (December 1995), Costa Rica (December 1995), and Bhutan (April 1996), the appropriateness of each country as a case study was also

assessed. The criteria used to determine which country or countries to select as case studies included: (i) the size of the Volunteers programme in the country; (ii) the availability of different categories of Volunteers; (iii) the variety in the assignments of the Volunteers; (iv) the role of the country as a host as well as supplier of Volunteers; and (v) the availability of surveyors to carry out the data collection for the study.

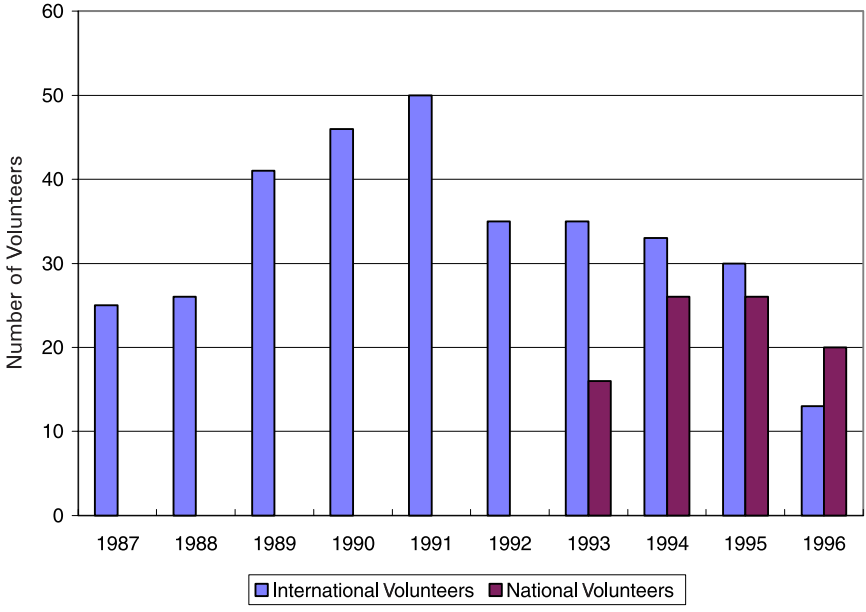
India was considered a potentially interesting case study because both international and national UN Volunteers had worked in the country and many Indians had worked abroad as Volunteers during the period covered by the study. However, the number of Volunteers who had served in India during this time, 44, was considered too small to justify selecting the country as a case study.

Similarly, the total number of volunteers who had served in Costa Rica between 1987 and 1996, 26, was considered too small, even if both international and national Volunteers had worked there. In contrast, Bhutan had one of the largest UNV programmes in the world, 250 international Volunteers between 1987 and 1996, and would have been suitable from the point of view of the size of the programme. However, the overall development context of the country was considered too special to make it suitable as a case study.

Nepal, on the other hand, was considered appropriate as a case study taking into consideration not only the length of the programme, which started in the mid 1970s, but also the size and availability of different categories of Volunteers in the country. During the period covered by the study, 97 international Volunteers, 50 UNV specialists and 47 UNV community workers, worked in Nepal. In 1987, the number of Volunteers working in Nepal was 25. During the late 1980s and early 1990s, the number of Volunteers grew and in 1991 the number reached 50, only to drop to 35 the following year and decline further to 30 in 1995. In 1996, another big drop brought the number of international Volunteers to 13. Meanwhile, national Volunteers were introduced in Nepal, with between 16 and 27 serving each year from 1993 to 1996 (see Fig. 4.2).

The fact that almost 350 Nepalese nationals served as Volunteers in other countries during this time also made Nepal attractive as a

Figure 4.2 Number of Volunteers serving in Nepal by year (1987–96)



case study, even if this aspect of the programme was not analysed in the end. Finally, it was also possible to identify and train a group of interviewers to collect data in Nepal at a reasonable cost.

The idea of using more than one case study was seriously considered, but in the end rejected in order to be able to go into sufficient depth in analysing the data from Nepal. Information and data from other countries visited during the course of the study was instead used to validate the methodology used and the results from Nepal.

Conceptual framework

At the time of the study, the UNV programme did not have an explicit model to explain how resources and activities were expected to

produce specific outcomes. A conceptual framework was therefore developed to analyse how the results were expected to be achieved. This was done using information obtained from reports and interview results, and by drawing on general evaluation literature and several different fields of study, including economics, organizational theory, sociology, anthropology, and political science.[6]

The outcome variables in the conceptual framework were derived from the six specific objectives of the study:

1. the impact of the programme in terms of human capital
2. the impact of the programme in terms of social capital
3. the impact of the programme in terms of changes in job opportunities, poverty, women's lives, and the environment
4. the perceived value of the programme in terms of the relevance of the work of the Volunteers
5. the perceived value of the programme in terms of the performance of the Volunteers
6. the perceived value of the programme in terms of the results and sustainability of the work of the Volunteers.

Even if the UNV programme may not have worked with explicit theoretical assumptions to guide its work at the time of the study, guidelines for the use of Volunteers in development cooperation did exist. These guidelines were used as a starting point for identifying factors or variables that could explain different outcomes.[7]

A logical framework was another starting point for the conceptual model of the study, based on the fact that most of the projects on which the Volunteers worked were designed using a logical framework.[8] The logical framework links the design of a project to its implementation by requiring the specification of inputs and activities that are expected to produce outputs and outcomes. A logical framework also identifies important assumptions and includes verifiable indicators of success, and may therefore help explain if an outcome was not achieved because the right kinds of inputs were not provided, or because important assumptions were not satisfied.[9]

From the point of view of the UNV programme, the Volunteers are the key inputs that determine the outcomes of the activities of the programme. This was also recognized by the users and beneficiaries of the programme, who during the exploratory visits frequently mentioned the qualifications, experience, skills, and motivation of the Volunteers as factors determining the outcomes of the activities of the programme. These have, consequently, been included as explanatory variables in the conceptual framework. Other background variables that could conceivably influence the performance of the Volunteers, and which have been included in the framework, are the age, gender, nationality, and family situation of the Volunteers.

In their reports, the Volunteers themselves also identified a number of factors that affected their work, usually in a negative way. These included the rules, regulations, and procedures of the UN, which were considered too cumbersome by many of the Volunteers. Other issues that were identified and included in the framework as possible negative or positive influences are the remuneration, entitlements, conditions of service, status, and placement of the Volunteers.

A number of other issues, which were identified in previous studies, relate to the implementation of a project. These are the recruitment, orientation, briefings, training, and support provided to the Volunteers during their assignment, the length of the assignments, and the coordination with other international organizations. These have also been included in the framework as possible explanatory variables, in the category of variables that the UNV programme normally can influence.

Several other variables also relate to the design and implementation of a project, which in the past normally have been the responsibility of the UNDP, another UN agency, and the government ministry or institution where a UN Volunteer worked. From this category, variables that have been identified and included in the framework are the design of the project, the availability of coworkers, inputs and support from the government, work planning, management and supervision of the Volunteers, and coordination with government agencies and non-governmental organizations.[10]

Research related to World Bank projects shows that assistance to certain sectors has been more effective than assistance to other sectors in a number of countries. Although the projects where UN Volunteers have worked are usually very different from most World Bank programmes and projects, the sector in which a Volunteer worked was also included as an explanatory variable in the framework of the study.[11]

Finally, a number of contextual variables also exist that the UNV programme cannot influence, but that may affect the impact of the programme. Factors that have been identified and included in the framework are the national legislation, the administrative procedures of the government, the location of the assignment of the Volunteer, the climate, and the security situation, the economic situation, the social situation, the cultural situation, and the political situation in the country.[12]

A summary of the different variables identified that could influence the impact and perceived value of the UNV programme are shown in Figure 4.3 (page 54). In this framework, factors that the programme can influence and those that it has little or no control over have been separated for purposes of clarity.

Data collection and analysis

Questionnaires, which primarily contained close-ended questions, were developed and used in the study. Before the questionnaires were used, they were pre-tested, revised based on the pre-test, translated into Nepali, and field tested in Nepal. Different questionnaires were developed and used to collect data from former supervisors and co-workers of the UN Volunteers, beneficiaries of UNV assistance, the reference group, and the former Volunteers themselves.

All of the questionnaires included the main areas of focus of the study: human and social capital accumulation, changes in the availability of jobs, the level of poverty, women's lives, and the environment (objectives 1–3 of the study). Questions on factors that may

Figure 4.3 Conceptual framework for analysing the impact and perceived value of the UNV programme

Factors expected to influence outcomes

Volunteer characteristics[a]
qualifications, experience, skills, motivation, age, gender, nationality, and family situation of Volunteers

Terms and conditions[a]
rules, regulations, procedures, remuneration, entitlements conditions of service, status, and placement of Volunteers

Type of assignment[b]
sector, host, executing agency

External environment[b]
national legislation, administrative procedures, location, climate, security, economic, social, cultural, and political situation

(CONTEXT)

UNV contribution[a]
recruitment, orientation, briefings, training, support, determination of length and preparation of post description for the assignment

Counterpart contribution[b]
availability of and interaction with supervisors and co-workers, work planning, management, and supervision

Links and support[b]
link to and support from government agencies, international organizations, and non-governmental organizations

(INTERVENTION)

Outcomes

Impact of the programme
1. Human capital, measured through acquisition of new skills or knowledge.
2. Social capital, measured through changes in values and attitudes, motivation, cooperation, and participation.
3. Changes in job availability, poverty, women's lives, and the environment.

Perceived value of the programme
4. Relevance of the activities.
5. Performance of the Volunteers in comparison with other alternatives.
6. Sustainability, in terms of use of skills and knowledge, continuation of activities, and long-term benefits.

(RESULTS)

[a] Factors that can be influenced to a large extent by the head office of the UNV programme.
[b] Factors that can be influenced to a lesser extent – or not at all – by the head office of the UNV programme.

have affected the performance of the Volunteers and other aspects of the perceived value of the programme were included in the questionnaires to former Volunteers and respondents who had interacted significantly with the Volunteers (objectives 4–6 of the study).

Background information on how often the respondents had met with a Volunteer and how much time they had spent with the Volunteer was also included in the questionnaires. Information about the age, gender, and educational background of the respondents was also collected. For the Volunteers, the necessary background information about their age, gender, qualifications, experience, etc. could be obtained from the UNV head office and was therefore not included in the questionnaires.

At the time of the development of the questionnaires, a decision was made to use close-ended rather than open-ended questions. The reason was that it had been possible to identify the key issues and a range of answer options based on the reports reviewed and the exploratory visits. Open-ended questionnaires could have provided more in-depth information than close-ended questionnaires, but the responses would have been difficult to analyse for a large number of respondents. In the key area of human capital accumulation, however, respondents were able to list skills and knowledge that they had learned. The survey instruments also provided opportunities for the respondents to comment on other issues related to the UNV programme.[13]

In parallel with the development of the survey instruments, sampling and identification of respondents took place. All former Volunteers who had worked in Nepal during the period of the study and whose addresses could be retrieved were included in a mail survey. For the other respondents, however, a sample had to be drawn in order to keep the logistics of the study manageable. It was decided that limiting the number of interviews through sampling was preferable to reducing the number of different categories of respondents from whom information could be collected. The sampling of respondents for the interviews was done independently from the mail survey and was not influenced by the responses to the mail survey.

To select the interviewees for the study, a randomized, geographically stratified, and gender-balanced sample of 50 Volunteers was drawn. In addition, 15 replacements were identified. This was done in case it would not be possible to find supervisors, co-workers, or beneficiaries for some of the 50 Volunteers selected for the sample, or in case it would not be possible to complete the interviews for the sampled Volunteers for some other reason. In the end, 9 replacements had to be made.[14]

For each of the 50 Volunteers, 6 respondents were selected, which brought the total number of interviewees to 300. In most cases, only 1 supervisor and co-worker could be identified, so for these categories of respondents, the issue of selection of respondents did not arise. With regard to the beneficiaries and reference persons, there was usually a lot more choice and the surveyors were instructed to select 1 beneficiary and 2 reference persons who would be fairly representative of the community or organization where a particular Volunteer had worked. Although there was no deliberate strategy to ensure that the beneficiaries and reference group respondents were randomly selected, there is no reason to believe that any bias which would have distorted the findings of the study would have crept in.

Out of the total of 300 respondents identified, the results of 298 interviews could be used: 169 of these represented users and beneficiaries of the programme, that is, people who had professional contact with a particular Volunteer, while 129 of the respondents did not know a Volunteer whose work was assessed and therefore constituted the reference group. Where the final selection of respondents did not correspond to the underlying population of Volunteers, a correction was made by a weighting of responses.[15]

Structured interviews in Nepali or English were carried out with supervisors and co-workers of the 50 former Volunteers sampled for the survey, employees of government ministries and departments and international organizations, and members of non-governmental organizations and communities where the Volunteers had worked, that is, beneficiaries and reference group respondents.[16]

Additional information from the Volunteers whose work was

assessed was received through a mail survey. Of the 97 Volunteers who had served in Nepal during the period covered by the study, addresses of 85 could be located, and these Volunteers were included in the mail survey. The total number of responses to the survey, following a reminder sent to those who did not initially respond, was 48, that is, 56 per cent, which can be considered satisfactory. In the UNV community worker category the response rate was 50 per cent and in the UNV specialist category it was 63 per cent.

The data was coded and analysed using the Statistical Package for Social Sciences (SPSS).[17] The data analysis included one-way analyses of variance, and multiple linear and logistic regression analyses. The regression analyses were used to test if the interaction with Volunteers could explain differences in the responses of different respondents. Contact with Volunteers and four other explanatory variables were included in the analyses: the age, gender, and education of the respondents as well as the geographical location.[18]

Since the very beginning, the study faced two serious methodological constraints: the absence of any pre-intervention data, and the absence of a control group in the true sense of the word. A number of measures, therefore, had to be undertaken to strengthen the research design, to ensure the relevance and validity of the study, and to monitor the reliability and objectivity of the study continuously.[19]

Validity

From the outset, particular attention was given to ensure the internal validity of the study, that is, that the survey would capture and measure the main areas of focus of the study. Support for the validity of the conceptual model of the study was received through a factor analysis, which confirmed five of the seven key factors that had been identified in the conceptual framework of the study. The identified factors were: (i) the experience, competence, motivation, and origin of the Volunteers; (ii) the type of assignment of the Volunteers; (iii) the recruitment of the Volunteers; (iv) the link to and support from other organizations; and (v) the rules and conditions of the assign-

ment. The two factors that did not stand out in the factor analysis were the external environment and the counterpart contribution.[20]

To enhance the validity of the survey instruments in terms of their content, comments on the draft questionnaires were requested and received from staff of the UNV programme and former Volunteers. Furthermore, in order to assess the face validity of the questionnaires, that is, how they would really work in practice, pilot tests were carried out in Nepal.[21]

The correlation between items in the questionnaires that measured similar things was calculated to assess the internal validity of the survey instruments. For instance, respondents who said that a Volunteer did a job that no one else locally could have done, should in most cases also have considered the performance of the UN Volunteer compared to a Nepalese national as good or very good. This indeed was the case, as is reflected in a significant correlation between these two items, which supports the internal validity of the questionnaires.[22]

In order to determine the construct validity of the study, the extent to which the questionnaires were able to discriminate between different categories of respondents was measured. This was done using a questionnaire item that assessed changes with respect to peace, democracy, and/or human rights. As expected, and as is explained below, the responses of the users and beneficiaries of the programme differed markedly from those of the reference group.

None of the Volunteers in Nepal had specifically worked on issues related to peace, democracy, or human rights, and it was therefore not expected that the Volunteers would have had much impact on these areas. At the same time, in view of the political changes that Nepal went through in the early 1990s, it was expected that the respondents in the reference group would indicate significant changes with regard to peace, democracy, and human rights in the country. This indeed was the case, which supports the argument that the items in the survey instruments were able to differentiate between different groups of respondents.[23]

To ensure the validity of the statistical conclusions of the study, methods for analysing the data were used in a standard way and results were interpreted carefully. To obtain more certainty regarding the relative magnitude of the changes that could be attributable to the work of the Volunteers, multiple regression analyses were complemented by logistic regression analyses, which are usually considered more robust under model misspecification.

To assess the external validity of the study, the methodology that was used in Nepal was also tested in Costa Rica and Mozambique. The questionnaires developed for Nepal were used in English in Costa Rica, translated into Portuguese for Mozambique, and administered with very good results in both countries. The fact that the survey instruments worked very well in very different contexts, and that the results from Mozambique and Costa Rica, albeit from a small number of respondents, were very much in line with the findings from Nepal, suggests a broader applicability of the methodology and supports the validity of the findings.

Reliability

To increase the reliability of the survey, two pilot tests of the questionnaires were conducted. Two different groups of respondents in different parts of Nepal were used in order to obtain as much feedback on the questionnaires as possible. Re-administering the questionnaires to the same group of respondents would have been a way to test if responses changed over time, but this was not considered necessary. The reason was that responses in the actual survey were expected to be stable, since the interviews in all but three cases were conducted more than a year after the Volunteers had left. Re-administering the questionnaires to the same respondents in the reference group could, however, have been useful since this might have uncovered some of the weaknesses with using a structured format for this category of respondents.

Although some difficulties with the structured format of the one-

to-one interviews could be detected at the time of the pre-testing, it was still considered advantageous to collect data in a form where the results of the reference group interviews could be directly compared, item by item, with those of the supervisors, co-workers, and beneficiaries. An alternative, which was considered but rejected in favour of the structured interviews, was focus group discussions to learn about changes in the work environment of the UN Volunteers that had taken place without the involvement of the Volunteers.

Collecting information from a group of respondents against which to compare the assessments of the users and beneficiaries of the UNV programme was no doubt very good, but did pose some problems. The reference group consisted of employees of government departments and international organizations and members of communities who did not know a particular Volunteer. Since they could not be asked to assess the work of a Volunteer, they were asked more general questions about skills or knowledge they or other people in their community or organization had learned, changes in values and attitudes, the availability of jobs, poverty, women's lives, and the environment, etc. To several of the respondents these questions seemed somewhat vague or abstract, and the overall conclusion is that the survey instruments could have worked better in this group.

Another potential threat to the reliability of the study was the way the different respondents were selected. The reference group respondents, on the one hand, were selected specifically because they did not know a particular Volunteer; the users and beneficiaries of the programme, on the other hand, were chosen because they had professional contact with a particular Volunteer, and they knew this was the reason they were selected. If the selection procedure had an effect on the responses of the users and beneficiaries and not on the reference group, making comparisons between these two groups would have been problematic. This, however, does not seem to have been the case.

If indeed the responses of the users and beneficiaries, on the one hand, and the reference group, on the other hand, were different because of the way the respondents for the two groups were selected,

one would have assumed differences between the two groups to be relatively consistent. However, once the respondents were divided into two groups based on their geographical location (Kathmandu and other areas of Nepal), a different pattern of responses emerged.

In Kathmandu, in terms of new skills or knowledge learned, for instance, there was not a significant difference between the users and beneficiaries, on the one hand, and the reference group, on the other hand. In areas outside Kathmandu the difference was significant. Among the users and beneficiaries, the difference between Kathmandu and other areas of Nepal was significant, while the difference in the reference group between Kathmandu and other areas of Nepal was almost significant.

Similar analyses relating to changes in social capital, jobs, poverty, women's lives, and the environment were also performed, with very similar results. Based on the results of these analyses, a proposition that the differences in the responses of the users and beneficiaries, on the one hand, and the reference group, on the other hand, would be a result of the way the respondents were selected must be rejected. This, thus, provides support for the overall reliability of the results of the study.

Another potential threat to the reliability of the survey could have been a tendency among respondents to select the middle option among the available answer options. In this study no such tendency could be detected while reviewing the responses to individual questionnaire items, which provides further support for the reliability of the survey.

The interviews were carried out by a total of 14 surveyors. It might seem that using a smaller number of surveyors would have been better in order to ensure consistency in the way the information was collected. It was, however, considered that the advantages of using a rather large number of surveyors did outweigh any potential loss in terms of accuracy or consistency in the way the interviews were conducted. Particular emphasis was given to the training of the surveyors to ensure that they all knew exactly what to do and what was expected from them. In the end, the performance of the

surveyors was very good, and the results of all but 2 out of 300 interviews completed could be used.

Using a smaller number of surveyors would have increased the time needed to complete the survey considerably. Including travel time, the 14 surveyors needed 221 days in all to complete the 300 interviews, which corresponds to 37 weeks, or more than 9 months, of work for one person using a six day working week. The average number of interviews completed by individual surveyors was 21, and ranged from 49 in the capital Kathmandu to 12 in the least accessible parts of western Nepal. The number of surveyors could have been reduced somewhat, but the capacity and availability of the surveyors also partly determined their number. Other factors and practical considerations that also drove the data collection included cost and weather. The data collection was timed to take place during the winter months when the weather would be pleasant and travelling would not be made difficult by monsoon rains.

Of the 97 Volunteers who had served in Nepal between 1987 and 1996, 85 Volunteers whose addresses could be located were included in the mail survey. Responses were requested within 5 weeks from the date the questionnaires were mailed. A total of 42 responses (49 per cent) were received within 2 months from the time the questionnaires were mailed. At this time a reminder, along with a copy of the questionnaire, was re-sent to all Volunteers who had not responded. This resulted in 6 additional responses, bringing the total number of responses to 48 (56 per cent).

While 56 per cent of the Volunteers did respond to the survey, it is still possible that these respondents are not entirely representative of the Volunteers who served in Nepal during 1987–96. Of the Volunteers who responded, most seemed to have had a positive experience. Not all, however, had a positive experience, and of the four Volunteers included in the sample whose contracts were prematurely terminated, two responded to the mail survey. This may be seen as an argument against any bias in the findings that could have resulted if only the views of the Volunteers who had a positive experience were included in the findings of the study.

Objectivity and relevance

One reason for recruiting and training a group of local surveyors was to be able to complete the data collection in as short a period as possible. Another reason was to try to ensure the neutrality and lack of bias of the survey. It was thought that using surveyors from Nepal, who spoke the local language and who were not associated with the UNV programme in any way, would be able to obtain more frank and honest responses than one or several external interviewers who in one way or another would have been perceived as linked to the UNV programme.

Another way to try to minimize any bias and increase confidence in the findings of the study was to use data from as many different sources as possible. These included previous studies, reports prepared by Volunteers, interviews with users and beneficiaries of the programme, external referees, and a mail survey to former Volunteers.[24]

Although some of the respondents who were interviewed indicated that certain items in the questionnaires were not applicable or relevant to them, the research questions and focus of the study remained relevant throughout the evaluation. While the study focuses on one country, it provides insights that should be relevant to the UNV programme as a whole. The results of the study and the lessons learned are also expected to be of interest and relevance to the UNDP and other funds, programmes, and agencies of the UN system as well as a more general audience interested in the work of the UN.

Notes

1. For a presentation of content analysis as a method, see: Krippendorff, Klaus. *Content Analysis: An Introduction to Its Methodology*. Thousand Oaks, CA: Sage, 1980.

2. 11 Volunteers completed two assignments in Nepal, and one Volunteer completed three assignments during the period covered by the study. Even if the number of Volunteers who served in Nepal (97) is used to calculate the percentage of end-of-assignment reports available, 61 per cent is still a low response rate on which to base any analysis of the outcomes of the work of the volunteers.

3. For a more detailed discussion on meta-analysis as an approach, see: Valadez, Joseph and Michael Bamberger, eds. *Monitoring and Evaluating Social Programmes in Developing Countries*, EDI Development Studies. Washington, DC: The World Bank, 1994, pp. 64–65.
4. A similar conclusion can be found in a study commissioned by the Development Assistance Committee of the OECD, which tries to synthesize evaluations on the impact of development projects and programmes of non-governmental organizations. This report raises serious questions about the quality of the evaluation reports, which constitute the basis for the synthesis study. See: Riddell, Roger C. and others. *Searching for Impact and Methods: NGO Evaluation Synthesis Study*, Ministry for Foreign Affairs of Finland, Department for International Development Cooperation, Report 1997:2. Helsinki: Hakapaino, 1998, pp. 11–13.
5. Through discussions with members of communities, organizations, and government departments where Volunteers had worked, the research design and the conceptual framework of the study were developed further. For a discussion on the benefits of exploratory work as part of the development of survey, see: Valadez, Joseph and Michael Bamberger, eds. *Monitoring and Evaluating Social Programmes in Developing Countries*, EDI Development Studies. Washington, DC: The World Bank, 1994, pp. 296–297.
6. Carvalho, Soniya and Howard White discuss the different emphases of different disciplines in the assessment of programmes and projects in *Implementing Projects for the Poor: What Has Been Learned?* Directions in Development. Washington, DC: World Bank, 1996, pp. 5–8.
7. United Nations Volunteers. "The Appropriate Use of Volunteers in Development". *United Nations Volunteers Thematic Series*, Programme Advisory Note. Geneva: UNV, 1991.
8. Logical frameworks, often referred to as a "LogFrames", have been a tool used for many years, in various adaptations, by the UN organizations and other agencies, for planning, managing, and evaluating development projects. For an overview of different logical frameworks, see: MacArthur, John D. "Logical Frameworks Today – Increased Diversification of the Planning Format". In: Kirkpatrick, Colin and Jon Weiss, eds. *Cost-benefit Analysis and Project Appraisal in Developing Countries*. Brookfield, VT: Edward Elgar, 1996, pp. 128–143.
9. For a sample of the logical framework that has been used by the UNDP and several other UN agencies, see: United Nations Development Programme. *How to Write a Project Document*. New York: UNDP, 1990, pp. 2–8.
10. A description of how political considerations, rather than the priorities of local people, affect the allocation of resources can be found in Blakie, Piers, John Cameron, and David Seddon. *Nepal in Crises: Growth and Stagnation at the Periphery*. Oxford: Clarendon Press, 1980.
11. Carvalho, Soniya and Howard White. *Implementing Projects for the Poor: What Has Been Learned?* Directions in Development. Washington, DC: World Bank, 1996, p. 6. See also Israel, Arturo. *Institutional Development: Incentives to Performance*. Baltimore: Johns Hopkins Press, 1987, for a discussion on differences in specification of benefits and competition as explanations as to why World Bank assistance to certain sectors such as industry, telecommunications, and utilities

may have been more effective than assistance to other sectors, such as education and services.

12. For a discussion on the importance of the social and cultural context for the effectiveness of development projects, see: Cernea, Michael, ed. *Putting People First: Sociological Variables in Rural Development*, 2nd edn. New York: Oxford University Press, 1991.

13. At the time of the design of the survey instruments it became evident that some of the questions would not be relevant to all Volunteer assignments, e.g., those of a primarily administrative nature. It was, however, considered better not to exclude certain categories of Volunteers in order not to manipulate the sample, and to have uniform survey instruments for all respondents so as to be able to compare the responses of all respondents. This was considered important in order to be able to present a more accurate picture of the impact of the UNV programme in Nepal – including assignments that may have had little or no impact because a Volunteer had done little else except administered fellow Volunteers, a UN programme, or an office. During the actual interviews, however, only a few of the respondents indicated to the surveyors that some of the questions were not relevant to their situation.

14. In a country such as Nepal, which is characterized by a difficult terrain and poor communications to many parts of the country, the geographical location of the assignments of the Volunteers was considered important. Therefore, the sample was divided over the five administrative regions of Nepal, the Eastern, Central, Western, Mid-Western, and Far-Western regions, and the capital city of Kathmandu. The next criteria that was used in the sampling was gender balance. The original balance between male and female Volunteers in the two sub-populations (approximately 3:2 for UNV specialists, and 4:1 for UNV community workers) was to remain in the sample. Finally, it was decided that an individual Volunteer should appear only once in the sample even if he or she had completed more than one assignment and could technically have been included a second time in a randomized, geographically stratified, and gender-balanced sample.

15. In order to prevent any estimation bias that could have resulted from the replacements that had to be made during the course of the study, a weighting of responses was introduced at the time of the analysis of the data. The correction that had to be made was needed to address the over-representation of female Volunteers and the under-representation of male Volunteers in the sample. Weights in the form

$$w_q = 1/q$$

were introduced to make the necessary corrections. In these expressions q equals the number of (either male or female) Volunteers sampled over the number of Volunteers who should have been sampled from the two sub-populations of UNV specialists and UNV community workers. For instance, out of the total 25 community workers who were sampled, 17 were males. The correct number would have been 20, therefore, $q = 17/20$ and $w_q = 1.18$. This means that the responses of each of the 17 male community workers have been weighted up by

a factor of 1.18, thus producing a combined result that is equivalent to that of 20 unweighted responses of male community workers, which is correct.

16. As was expected, some respondents, particularly those who worked in government ministries and UN agencies in Kathmandu, felt somewhat constrained by the structured format of the questionnaires and perceived that they could not enter into a dialogue with the surveyors. Because of the way the study was designed, however, it was never intended to encourage dialogue between surveyors and respondents, and it was in fact anticipated that some respondents would feel somewhat frustrated by the structured format. The advantages and disadvantages of the methodology chosen, and the questionnaires used, had, however, been weighed against using semi-structured interviews or another similar approach that would have required setting up the research differently in terms of the number, qualifications, and experience of the surveyors as well as the training and supervision provided to them.

17. For one of the outcome variables, changes in human capital, interval level data were obtained and used. For the other outcome variables, ordinal level data were obtained and used once responses had been coded from five-point scales such as: "very negative", "negative", "none", "positive", and "very positive". In addition, arithmetic averages of four ordinal level questions were used as indicators of specific outcomes: changes in social capital or changes in UNDP's priority areas. For a discussion of potential problems with using a summation of ordinal level data to create an index, such as the one used to measure social capital development, see: Valadez, Joseph and Michael Bamberger, eds. *Monitoring and Evaluating Social Programs in Developing Countries*, EDI Development Studies. Washington, DC: The World Bank, 1994, p. 202.

18. Human capital was measured using the number of skills learned and areas of new knowledge acquired that the respondents indicated. Social capital was measured using an unweighted average of the responses to four questions that assessed changes in (i) values and attitudes, (ii) motivation, (iii) cooperation, and (iv) participation of people in local affairs; in addition, the individual components of social capital were examined. Changes with respect to UNDP's priority areas were analysed using an unweighted average of the responses to four questions dealing with changes in (i) job availability, (ii) poverty, (iii) women's lives, and (iv) the environment, as well as through an analysis of the responses to the four individual questionnaire items. The other outcome variables related to the performance of the Volunteers and the relevance, results, and sustainability of the activities of the Volunteers were measured using individual items from the questionnaires. The explanatory variables (see Fig. 4.3) included quantitative variables (e.g., age of the respondents), qualitative variables with multiple categories (e.g., very poor, poor, OK, good, very good), as well as variables with only two values or levels (e.g., gender, or contact with the Volunteers). To facilitate the analysis, qualitative variables with multiple categories were where possible recoded as "dummy variables", i.e., with only two values. In the case of geographical location, for instance, the two values used were the capital city and other parts of Nepal.

19. See: Caudle, Sharon L. "Using Qualitative Approaches". In: Wholey,

Joseph S., Harry P. Hatry, and Kathryn E. Newcomer, eds. *Handbook of Practical Program Evaluation*. San Francisco: Jossey-Bass, 1994, pp. 84–93.

20. The total number of issues identified that could influence the work of the Volunteers for which data could be obtained was 64 (see Fig. 4.3). To determine which individual items to include in the factor analysis, bivariate correlation coefficients between the individual explanatory variables and the outcome variables (changes in human capital, changes in social capital, and changes in the UNDP's priority areas) were calculated. All items which had a significant correlation ($p < 0.05$) with one of the outcome variables were included in the factor analysis. This reduced the number of explanatory variables from 64 to 28. Following an exploratory factor analysis to identify principal components of the 28 explanatory variables, a varimax rotation was performed using unweighted least squares. This resulted in eight interpretable factors, five of which corresponded to the seven key factors identified in the conceptual model of the study.

21. Cook, Thomas H. and Donald T. Campbell. *Quasi Experimentation: Design and Analysis Issues for Field Setting*. Boston: Houghton-Mifflin, 1979, break down the concept of validity into four components: internal validity, construct validity, statistical conclusion validity, and external validity.

22. The correlation between the two items that tested the internal validity of the questionnaire was statistically significant ($r = 0.35$, $p < 0.01$).

23. Throughout this chapter, a significant difference refers to a probability of a difference of more than 95 per cent ($p < 0.05$). For an overview of recent economic, political, and social changes in Nepal, see, e.g., Karan, Pradyumna P. and Hiroshi Ishii. *Nepal: A Himalayan Kingdom in Transition*. Tokyo: The United Nations University Press, 1996.

24. For a discussion on the systematic use and comparison of independent data collection methods, known as "triangulation", see, e.g., Valadez, Joseph and Michael Bamberger, eds. *Monitoring and Evaluating Social Programmes in Developing Countries*, EDI Development Studies. Washington, DC: The World Bank, 1994, pp. 224–225.