Evaluation of Development Interventions and Humanitarian Action

In: The SAGE Handbook of Evaluation

By: Osvaldo Feinstein & Tony Beck

Print pages: 537-558

©2006 SAGE Publications Ltd. All Rights Reserved.

This PDF has been generated from SAGE Research Methods. Please note that the pagination of the online version will vary from the pagination of the print book.
Evaluation of Development Interventions and Humanitarian Action

Introduction

Development evaluation is the evaluation of interventions such as projects, programs, policies, and processes whose general objective is to promote development. Thus, the evaluation of aid projects funded by external donors (the type of development evaluation that dominated development evaluation during several decades) is a specific type of development evaluation, which also includes the non-aid dimension of development, such as trade. There are also development evaluations focusing on clusters of projects/programs by sector or theme (e.g., agricultural projects). In addition, by the second half of the 1990s, development evaluations started to focus on country-level programs, and since 2000 there have been some new evaluation efforts addressing global programs.

Not surprisingly, whereas development evaluation has been a marginal topic for associations of evaluators based in developed countries, such as the American Evaluation Association, the practice of development evaluation is the main concern of national associations of evaluators based in developing countries (several of which were established in the late 1990s and at the beginning of the twenty-first century). However, the discussion on development evaluation as a discipline has been rather absent in the discussions of national and regional associations, and it became the main focus of the recently created “International Development Evaluation Association” (IDEAS).

Every day large-scale international humanitarian action is taking place in at least 20 and sometimes as many as 40 locations around the world. Occasionally, this hits the headlines, in particular where there are western foreign policy interests involved. Evaluation of humanitarian action (EHA) has become the main means of accounting for the effects of humanitarian action - to donors, executive boards of international agencies, and western publics. For example, the international response to humanitarian needs created by the 1999 conflict in Kosovo led to 16 formal evaluations (Wood, Apthorpe, & Borton, 2001); and the response in Afghanistan after 2001 has led to at least 15 evaluations.

Evaluation of humanitarian action tends to mirror humanitarian practice - it is often rushed, heavily dependent on the skills of its key protagonists, ignores local capacity, is top-down, and keeps an eye to the media and public relations implications of findings. In some ways, in particular in relation to the context in which much EHA takes place, it differs radically from mainstream evaluation. In other ways it resembles mainstream evaluation, for example an earnest attempt to improve the quality of programming combined with lack of attention to use of
evaluation findings. Its subject matter is usually controversial, and often sensationalized by the western media - famine in Southern and the Horn of Africa, Rwanda, Hurricane Mitch, Kosovo, and in the early twentieth-century Afghanistan and Iraq.

After a discussion on the differences and similarities between evaluation of development and humanitarian action, the first part of the chapter will review current practices of development evaluation, including its different types and the way in which its practitioners have organized themselves, concluding with the challenges faced by development evaluators. The second part of the chapter is organized as follows. First, a background section provides a general overview of both humanitarian action and EHA, pointing out main trends. In the following section findings of the successes and failings of humanitarian action, as reported in 215 evaluations carried out since 2000, are elaborated. After this there is a discussion of the quality of EHA, drawing on a meta-evaluation carried out on 127 evaluation reports between 2001 and 2003. The final section examines some recent and likely future trends in EHA.

**Differences and Similarities Between Development Evaluation (DE) and the Evaluation of Humanitarian Action (EHA)**

One important difference between DE and EHA is that whereas in development evaluation “sustainability” (i.e., the continuation of the effects of the development interventions after they are completed) matters, in humanitarian actions what is crucial is “effectiveness” or “efficacy” (the extent to which the objectives were achieved). Humanitarian actions have been more effective in achieving their short-term objectives, whereas longer-term objectives correspond more to development actions. EHA bears some similarities to development evaluation, in particular the use of procedures and tools such as document review, structured and semi-structured interviews, surveys and focus groups. While some agencies, for example, WFP and UNHCR, deal almost exclusively with emergencies, others such as UNICEF and many NGOs work in both emergencies and development situations, so that their evaluation offices cover both areas of work.

However, there are two substantial differences between EHA and DE. First, humanitarian agencies operate within a “humanitarian space” that may be constrained by deliberately restricted access (roadblocks, attacks on aid convoys and personnel) as well as access being restricted through poor infrastructure and seasonal climatic factors. The population being assisted may be subject to a range of human rights abuses and exclusion from accessing basic services by the armed forces and groups directly involved in the conflict or by others exploiting the breakdown of law and order. All of these areas need to be considered in EHA.
Second, key information on a range of matters of vital significance to evaluators is often unavailable. Whilst evaluators of development programs are also often faced with a lack of information on key indicators or decisions, such problems are considerably more serious in the case of EHA. Staff turnover, and the difficulties of accessing key data in areas such as mortality and morbidity, mean that EHA is at a considerable disadvantage when attempting to draw conclusions about interventions.

Current Practices of Development Evaluation

Purposes and Roles of Development Evaluation

There is a growing consensus that development evaluation should serve both an accountability and a learning role. Though, early on, nongovernmental organizations (NGOs) emphasized the learning role of evaluation, they have been under pressure to use evaluation also for accountability. To some extent this has also been the case with bilateral agencies (i.e., those corresponding to the aid agencies of developed countries’ Ministries of Foreign Affairs), whereas the purposes of evaluation in international financial institutions evolved in a rather symmetric way, from an initial emphasis on accountability to a more balanced approach combining evaluation for accountability and for learning.

It is worthwhile to note that there are different views as to the relation between the learning and the accountability function of evaluation in the development evaluation context. Thus, whereas there is a widespread perception that there are trade-offs between the learning and the accountability function, an alternative conception has emerged by which learning and accountability are seen as complementary. In this view, accountability provides the incentive framework for learning, if there is an “accountability for learning,” i.e., for using lessons learned from experience to improve the design and implementation of development interventions. An organizational solution to the “trade-off” (or dilemma) has been to leave the accountability function to auditors and the learning function to evaluators.

Different Types of Development Evaluations

For the purposes of learning and accountability different types of development evaluations are conducted. Development evaluations can be classified in terms of what is being evaluated (their “object”) and how the evaluation is carried out (their “approach”). According to their object, it is possible to distinguish between project or program evaluations (evaluations of single interventions, such as the evaluation of rural development projects), sector or thematic evaluations (i.e., evaluations of a cluster of interventions corresponding to a sector - e.g., transport - or a theme - for example, the environment), country evaluations (which can
correspond either to the evaluation of the country program supported by a development agency, or, to all development interventions in a country. The latter is a type of evaluation that very seldom has been carried out, and it could be led by the recipient country. A more recent type is the of global public policies and/or programs, which are evaluations of partnerships (e.g., the evaluation of agricultural research programs supported by several donors).

In terms of evaluation approaches, a widely used approach in development evaluation is “objectives-based evaluation,” which takes into account the objectives of development interventions as the basis for the evaluation. It assesses the degree to which the objectives were achieved (the “efficacy” or “effectiveness” of the intervention), the relevance of the objectives, the efficiency in the use of resources and the sustainability of the interventions. A problem faced by this type of evaluation is that the objectives of interventions are not always appropriately defined, thus affecting their evaluability, i.e., the extent to which an activity or a program can be evaluated in a reliable and credible fashion. In order to cope with this problem, the logical framework or the results chain is sometimes reconstructed in order to identify the “results chain” and key indicators.

Another approach that has been used, which also helps to increase the evaluability of interventions, and that complements “objectives-based” evaluation, is the so-called “theory-based evaluation” (TBE). It makes explicit the assumptions of the intervention or its implicit “theory of change.” The example shown in Table 24.1 corresponding to an evaluation of social funds, shows the way in which TBE was used to identify the main channels through which social fund projects are expected to influence institutional development, and possible side-effects. By making explicit the assumptions the process of learning from experience is facilitated, as the evaluation serves to test the hypothesis and, therefore, to contribute to the stock of knowledge.

An alternative approach, not based on the objectives of the interventions, is “goal free evaluation” (an approach introduced in 1972 by Michael Scriven). In fact, one of the most influential development evaluations, Hirschman (1995) corresponds to this approach, which tries to identify patterns of critical success factors and hidden rationalities in sets of projects. Another approach, which sometimes complements those mentioned before, is “participatory evaluation,” which in the development evaluation field has been particularly used in rural and health projects and programs, facilitating the incorporation of the views of those that are evaluated and/or of the population for which the interventions are intended; furthermore, participatory evaluation facilitates the empowerment of participants (indeed, there is an
emerging field of “empowerment evaluation”), increasing the ownership (and thus the use of evaluations) by those that participate in the evaluation.

It is also worthwhile to point out that an approach that was thought not to be applicable in development evaluation, randomized control trials (RCT), has recently been considered the “gold standard” for evaluation, including impact development evaluation. Though there is some scope for randomization in development evaluation, there are also limits to its application in this area. It is at the micro or project level where this approach is best used, particularly in terms of generating knowledge on what does and does not work (but it does not addresses accountability). At this level there is a growing number of applications of RCT (see the chapter by Duflo & Kremer in Pitman et al., 2004). However, randomized designs face problems of external validity (given that the context of specific interventions are often critical for the outcomes, thus limiting the value of RCT for “scaling-up” (on this and on limitations to “internal validity,” as well as on other rigorous methods, see the comments by Martin Ravallion in Pitman et al., 2004). Furthermore, it is unlikely that RCTs can be applied at the macro level, such as in the evaluation of national policies, and in the evaluation of complex development issues where there are multiple changing factors that influence the object of the evaluation.

**EXEMPLAR 24.1 Application of theory-based evaluation: framework for analysis of institutional development impact**
Joint Evaluations

For several years developing agencies have been trying to carry out joint evaluations, in order to diminish transaction costs for developing countries and to gain additional insights in terms of what works and what doesn't, and the reasons for the observed results, from joint assessments. Furthermore, to the extent that the recipient country participates in the evaluation, it is more likely that there will be ownership of the evaluation by the recipient country, and, therefore, it is more likely that the recommendations and lessons learned will be applied.\(^{15}\)

However, the practice of joint evaluations has shown some difficulties in applying this approach. Nevertheless, progress has been made in some recent joint evaluations, particularly those in the area of basic education (under the leadership of the Evaluation Department of the Dutch Ministry of Foreign Affairs) and the evaluation of the “Comprehensive Development Framework,” which was modeled following the example of the 1996 joint Rwanda evaluation (described in Dabelstein & Rebien (2002) and by Borton & Eriksson (2004); see also subsequently in this chapter), involving in its design and implementation representatives of recipient countries, multilateral development banks, bilateral development agencies, international NGOs, and the

---

**EXEMPLARY 24.1 Application of theory-based evaluation: framework for analysis of institutional development impact**

<table>
<thead>
<tr>
<th>Theory</th>
<th>Assumptions</th>
<th>Data required to test the assumptions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main channels</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct effects</td>
<td>The social fund project includes a well-designed and effective institutional development component</td>
<td>Extent and quality of capacity building activities; observed changes in organizational behavior/capacities</td>
</tr>
<tr>
<td>Demonstration effects</td>
<td>There is knowledge about social fund approaches/procedures among government agencies</td>
<td>Evidence of government agencies embarking social fund procedures</td>
</tr>
<tr>
<td>Learning-by-doing effects</td>
<td>Social funds involve government agencies in decision making/planning</td>
<td>Extent of adoption of social fund approaches by government agencies</td>
</tr>
<tr>
<td>Demand effects</td>
<td>Communities increase their demand for government services</td>
<td>Level and nature of community demands</td>
</tr>
<tr>
<td><strong>Side-effects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resource-allocation effects</td>
<td>Social funds and other agencies compete for staff/resources</td>
<td>Movement of staff from government to social fund agency</td>
</tr>
<tr>
<td>Systemic budget effects</td>
<td>Social fund resources are off-budget and undermine accountability; reduce fiscal prudence and distort the budgetary allocation process</td>
<td>Procedures for budget accountability</td>
</tr>
<tr>
<td>Decentralization effects</td>
<td>Social fund strengthens or undermines investment planning/resource allocation</td>
<td>Relationship between the social fund agency and local government</td>
</tr>
</tbody>
</table>

Source: Carvalho & White (2004), table 2.\(^4\) This useful typology of effects could also be applied in other evaluation contexts.
private sector. A 30-member, multipartner steering committee and a five-member management group governed the evaluation.

Though joint evaluation efforts have focused on producing single evaluation reports, an approach that is emerging is to place emphasis on joint evaluation processes rather than only on its final product or outcome. In this way opportunities for valuable joint evaluation work can increase (e.g., joint case studies or joint surveys) even if a final single evaluation report cannot be produced (given, among other reasons, the different timeframes and audiences that evaluation partners have to take into account, which in some cases may not allow for the production of a timely single report). 16

Capacity Building for Development Evaluation

One of the constraints for conducting joint evaluations that include recipient country representatives is the limited evaluation capacity available at the country level. Though in some cases there are potential evaluation capacities in developing countries that can be mobilized (a “potential supply”), generally there is need to level the playing field, and promote the development of evaluation capacities in developing countries. These capacities, not only to conduct but also to manage evaluations, are crucial in order to promote ownership of evaluations and to facilitate their use. An interesting case is Chile, where during three decades a public sector evaluation system has evolved, outsourcing evaluations to the private sector and to universities (through a system of competitive bidding). The Chilean evaluation system takes into account the demands from the Parliament in establishing the evaluation program, and it has become a model that some countries in the developing world are using as a source of inspiration.

It is worthwhile to mention the four-week international program for development evaluation training (IPDET), launched in 2000, designed originally by the World Bank’s Operations Evaluation Department, and since then delivered annually by an international faculty, at Carleton University, Canada. Furthermore, tools have been designed for systematic diagnosis of evaluation capacities at the country level, which are useful to identify suitable actions to address the critical constraints. 17 Finally, it should also be taken into account that, as in other areas, “learning-by-doing” plays an important role in developing evaluation capacities, and providing opportunities to do evaluations is therefore a way to provide opportunities to learn.
New modalities of development aid, such as those involving a common pool of resources by several donors (“basket funding,” such as in “sector-wide approaches” or SWAPs, which instead of focusing on individual interventions in specific sectors deal with sectors - such as education - as a whole, with a comprehensive approach), pose new challenges for evaluators. Thus, the attribution of results to a program funded by a single donor, which is never an easy task, is made even more difficult when several donors fund jointly the same program. An approach that has been suggested to face this challenge, and which is gaining increasing support, is to focus on “contribution” rather than on “attribution,” trying to identify the contribution of each development partner to the development results. However, this may be only a way to relabel the problem, as “contribution” may be equivalent to “marginal attribution,” and it does not actually solve the attribution problem. Another approach is to “dissolve” the problem, focusing on other issues such as the governance of partnership programs, and/or to focus on the results of the development aid system as a whole. Of course, harmonization of evaluation practices would facilitate this type of evaluation.

An additional challenge is to find new ways of involving, in a significant way, developing-country evaluators in development evaluations. So far the role of developing-country evaluators in carrying out evaluations of development interventions has been quite limited. The most frequent arrangement (for their rather infrequent participation) is as consultants in missions of bilateral or multilateral agencies. One important reason for the lack of evaluations of development aid designed and carried out mainly by developing-country evaluators is the perception (and sometimes the reality) of a lack of sufficient capacities for evaluation at the country level. However, there is a vicious circle in that the lack of opportunities for designing and implementing evaluations deprives developing-country practitioners of opportunities to learn through a learning by doing process (as indicated above). One of the consequences of this lack of opportunities is that the emergence of indigenous models of development evaluation becomes unlikely. An interesting and promising way of involving developing country evaluators has been tried in Tanzania, through an “independent monitoring/and evaluation/group,” composed of foreign and national experts (see Killick, 2004).

Furthermore, as indicated at the beginning of this chapter, development aid is only a subset of development interventions. There are other interventions and issues, such as trade and foreign direct investment, which have a significant impact on development, frequently being much more important than development aid, and therefore becoming very “relevant.” Generating a demand for, and a capacity to carry out, evaluations of these crucial issues is one of the important challenges for development evaluators. This type of evaluations would complement the more traditional evaluation of development focused on aid, requiring an
adequate division of labor with researchers who had been working on these issues for several decades.

Finally, the implications of alternative views of development pose evaluation challenges that are not always acknowledged. Thus, Amartya Sen's conception of “development as freedom” and his “capabilities approach” has implications for the assessment of the relevance of development interventions, shifting the focus to the effects of development interventions on people's freedoms and capabilities. A more general challenge is to make explicit the implicit assumptions and values used in development evaluation. For example, economists conducting development evaluations frequently assume that the greater the removal of “distortions” to free trade, the better will be the development results in terms of increasing welfare; however, the economic “theory of the second best” shows that this is not necessarily the case and that when all distortions cannot be removed, then the removal of some distortions may increase or decrease welfare. 19 The importance of making explicit the values in development evaluation has been forcefully argued by Des Gaster (2004), complementing the emphasis on making theories explicit (characteristic of the “theory-based evaluation” approach). In fact, an “Assumptions Based Comprehensive Development Evaluation Framework” (ABCDEF) would allow development evaluators, and those that are evaluated, to become fully aware of the implicit and explicit assumptions made in the design of development interventions (including counterfactuals) and in their evaluations, thus facilitating a constructive and critical discussion that can facilitate the process of learning from experience. 20

**Development Evaluation as a Profession**

Most development evaluation practitioners have a professional background in some social science discipline (such as economics, sociology, anthropology, political science, and history). However, through professional evaluation training such as that provided by the Master in Evaluation of the University of Costa Rica (which started in 1995) and the International Program for Development Evaluation Training (IPDET, at Carleton University, Canada, 21 as well as the formation of an association of development evaluation professionals, namely, the International Development Evaluation Association (IDEAS), development evaluation is becoming a profession with practitioners that are starting to consider themselves as development evaluation practitioners. Furthermore, there is a growing demand for this type of professionals, particularly for the evaluation of development aid interventions, thus creating opportunities for professionals trained in development evaluation. Another (complementary) way to cater to this demand is through the mobilization of a potential supply of development evaluators whose background is research and/or development, and who may be able to carry out development evaluations with
some guidance, thus developing as a by-product their competence as development evaluators. The involvement of locally based evaluators, who are fluent in local languages and fully familiar with the cultural, social, and political context, is one of the best ways to overcome the limitations of foreign evaluators.

It should be noted that the “development evaluation architecture” consists of a set of partially overlapping evaluation associations and networks: the OECD/Development Assistance Committee (DAC) Network on Development Evaluation, whose members are the evaluation departments of bilateral development agencies of the OECD countries (including as observers the Evaluation Cooperation Group (ECG) members and the chair of the United Nations Evaluation Group - UNEG); the ECG, composed by the evaluation units of the International Financial Institutions (including as observers the chair of the OECD/DAC Network and the Chair of the UNEG); the United Nations Evaluation Group (UNEG), which includes all the evaluation departments of the UN system (with the chair of the OECD/DAC Network as observer, and the World Bank representing the ECG); IDEAS and the national and regional associations of development evaluation practitioners based in developing countries (including their coalition or federation, the International Organization for Cooperation in Evaluation (IOCE), which was launched in 2003). Furthermore, NGOs are also active in development evaluation; for example, the International NGO Training and Research Center (INTRAC) provides support to NGOs in evaluation (and other areas) and is a useful source of information about NGOs’ involvement in development evaluation.

Associations of Development Evaluation at the National, Regional and Global Level

Associations of evaluators in developing countries have been formed since 1990 at the national, regional, and global levels. The first of such associations was the Central American Evaluation Association, whose members have been mainly based in Costa Rica. It developed, jointly with the Canadian University of Hull, a masters program in evaluation, with several generations of graduates. Some of them have been participating in evaluations of program funded by the Central American Bank for Economic Integration. In the late 1990s national associations were formed in Asia, Latin America, and Africa. The African Evaluation Association was established in 1999 and held two regional conferences.22

A Latin American Program for Evaluation, PREVAL, was launched by the UN International Fund for Agricultural Development in 1995 (the program was designed in 1990). It has been organizing workshops on monitoring and evaluation, producing and disseminating evaluation documents, and it established a managed electronic network. Furthermore, the International Organization for Cooperation in Evaluation (IOCE), which is a federation of evaluation
organizations, including those in developing countries, was created in 2003. Finally, in August 2002 the International Development Evaluation Association (IDEAS) was launched, becoming the only global association whose main focus is development evaluation.

Thus, there has been an evolution from a period in which development evaluation was seen as a possible topic and/or “topical interest group” within established evaluation associations, such as the American Evaluation Association, to a situation in which the evaluation landscape has become populated with evaluation associations based in developing countries and with an international association, IDEAS, whose board includes a majority of members from developing countries. These associations are starting to play a role in promoting the demand for development evaluations, building an evaluation culture and, at the same time, are supporting the enhancement of capacities to undertake development evaluations with full involvement of developing-country practitioners and policy-makers. Thus, they are also promoting development through development evaluation.

The Evaluation of Humanitarian Action

Background - the History of Evaluation of Humanitarian Action

Funding from international donors to humanitarian crises in the 1990s has averaged around US$5 billion a year. This is some 10 percent of all overseas development funding, although if all humanitarian funding is taken into account the figure is probably closer to US$10 billion a year. Much of the actual work of humanitarian action is carried out by non-governmental organizations (NGOs) as implementing partners, with the UN, European Community Humanitarian Office (ECHO), and national donors as funders and monitors. The main UN actors are the World Food Program (WFP), the United Nations High Commissioner for Refugees (UNHCR), and the United Nations Children's Fund (UNICEF). A large number of NGOs are involved, in some high-profile emergencies more than 200, including internationally known organizations such as the International Federation of Red Cross and Red Crescent Societies (IFRC), the International Committee of the Red Cross (ICRC), Oxfam and Save the Children Fund. The multiple actors and funding mechanisms often make assessing humanitarian action complex.

Practitioners of humanitarian aid generally claim to anchor their work to the principles of neutrality and impartiality - for example, aid should be provided to those most in need, independent of political factors. As such it can be provided in any country or regime independent of its political or human rights record, and in this process is usually controlled by humanitarian agencies. In practice circumstances almost always intervene to problematize
ethical approaches (IFRC, 2003), and the key humanitarian principles have been under increasing pressure with recent interventions in Kosovo, Afghanistan, and Iraq, where humanitarian responses are seen as being subsumed under western foreign policy priorities, or as part of the US led “war on terror” (see, among others, FIFC, 2004; Hansen, 2004; Slim, 2004; Vaux, 2004). However, the humanitarian-military interface is not well covered in EHA.

EHA is a relative newcomer in the evaluation field, occurring systematically across the humanitarian sector only from the mid-1990s. A few evaluations were carried out after the 1983–5 famine in the Horn of Africa, but numbers were limited until 1993 (Wood, et al., 2001). The catalysts for increased attention have been uptake in recipient governments, the UN, donor governments and NGOs of results based management (see below); increased funding to humanitarian action since 1991; and the Joint Evaluation of Emergency Assistance to Rwanda (JEEAR, 1996).

The JEEAR has been a touchstone of EHA since the mid-1990s. This has been because of its subject - the genocide in Rwanda - and its scope. Recently, the Danida Evaluation Secretariat commissioned a follow-up study to the JEEAR to determine the extent to which its recommendations have been carried out (see Exemplar 24.1). JEEAR established what was possible in EHA, although not as yet repeated, despite major emergencies such as Hurricane Mitch, Kosovo, East Timor, and Afghanistan, which led to the disbursements of substantial amounts of humanitarian aid from multiple donors.

**EXEMPLAR 24.1 The Joint Evaluation of Emergency Assistance to Rwanda and Its Follow-Up**

The Rwanda genocide in 1994 resulted in the killing of at least 800,000 people during a 10-week period. Those killed were predominantly members of the minority Tutsi ethnic group, but a significant number of politically moderate Hutus were also killed. Inaction by the “international community,” including the US, and the evacuation of almost all non-Rwandan nationals, including the expatriate staff of most of the relief agencies, effectively cleared the way for the genocide. A massive refugee flow of over two million people to eastern Zaire (now the Democratic Republic of Congo - DRC) followed, provoking a large international relief effort and allocation of some US$1.4 billion between April and December 1994.

The enormity of these events led to an unprecedented collaboration to learn the lessons from the international response - the Joint Evaluation of Emergency Assistance to
Rwanda (JEEAR). JEEAR consisted of four major studies: on historical perspective; early warning; the effects of humanitarian aid; and rebuilding Rwanda. It was managed by a 38-member steering committee, including all major actors in the humanitarian system, chaired by the well-respected head of the Evaluation Secretariat at Danida. In scope, size, and influence, it dwarfs all previous and subsequent EHA.

The main finding of the JEEAR was that humanitarian aid had been used as a substitute for effective political and military action. Only political or military intervention would have stopped the genocide, but instead the international community relied on a belated humanitarian response as its main intervention. An eight-year follow-up to the JEEAR - also unprecedented in EHA in terms of follow-up to evaluations, and unusual in any evaluation sphere - found that little had changed in this respect, for example in Sudan and DRC.

Study 3 of the JEEAR, focusing on the humanitarian response, was made up of a core team of three supported by 14 technical specialists, including anthropologists, and health and nutrition and logistic specialists. Interviews were carried out with 480 donor, UN, NGO, and government personnel, and some 140 beneficiaries of assistance. In terms of the humanitarian response, JEEAR found that shortcomings included response capacity, coordination, the monitoring of the effectiveness of overall efforts, the professionalism of some NGOs, and accountability mechanisms in the sector generally. While there has been some progress since 1994, many of these shortcomings still dog the sector.

Sources: JEEAR (1996); Rieff (2002); Borton & Eriksson (2004).

How Evaluation of Humanitarian Action Has Been Structured and Carried Out

Evaluation of humanitarian action was marked through the 1990s by five main characteristics:

It tended to be atheoretical - that is, it did not during that period in general draw on conceptual thinking in mainstream evaluation, nor develop its own theory. This has meant that for the most part EHA has not been reflective as to the implications of methods selected, and the influence of these methods on evaluation findings (on which see Lipsey, 2000). The consideration of evaluator bias rarely arises.

Following from their atheoretical nature, EHA methods tend to be formalized and routine, with evaluation methods following much the same pattern: initial document review; an introductory round of meetings in national capitals; a two- or three-week visit to the affected country, with a focus on meetings at donor, the UN, and NGO offices in the capital and districts; usually
somewhat rushed visits to refugee camps or project sites; and a debriefing in-country and back at the national capitals (ALNAP, 2003).

There are of course important exceptions. For example, much EHA is structured around the Development Assistance Committee evaluation criteria (OECD-DAC, 1999). These are standard evaluation criteria such as impact, effectiveness, and efficiency, adapted for the humanitarian field. For the most part the DAC evaluation criteria have not been used creatively and have tended to foster a focus on accountability-based evaluation as opposed to lesson learning (ALNAP, 2004b). In other words, EHA has concentrated until recently on what happened (e.g. how much food aid was delivered, how many beneficiaries received aid, how many houses were built and for whom), rather than why things happened. This prioritizing of results over process has tended to hinder lesson-learning in the sector.

Much EHA consists of single-agency, single-sector evaluations, which tend to have a somewhat narrow focus on the effectiveness and impact of individual interventions (e.g. the WFP’s food aid program in the Great Lakes of Africa, or Oxfam’s water intervention in Angola). There have been very few joint evaluations and subsequently some of the wider political questions relating to human rights, protection, or the military-civilian interface, which could be better captured through joint evaluations, are inadequately evaluated.

The main purpose of EHA has been reporting to executive boards and donors. The impact of EHA is little understood, but in many cases evaluations are not changing practice in the ways intended. For example, a number of evaluations of housing programs in Bangladesh since 1988 have made similar recommendations, with little effect on practice (ALNAP, 2002).

However, partly because of concern regarding the lack of use of evaluation findings, and partly as a result of greater professionalization, over the last five years EHA has seen three encouraging trends:

- A move towards more experimental and participatory evaluation approaches, manifested mainly in real-time evaluation (see below for further details), now being used by a number of NGOs
- The move to more participatory evaluation is part of an attempt to achieve a better balance between accountability and lesson learning
- Again linked with the move to more participatory evaluation, many agencies have increased their focus on evaluation use, stemming from dissatisfaction with lack of follow-up and interest in evaluations, and using Patton’s work (1997).

EHA in Context

With much fanfare, most western government departments involved with aid and humanitarian
action announced the introduction of results-based management (RBM) in the early 1990s. This was partly a result of cuts to aid budgets and the persistent accusation that aid does not work. Where donors went, recipients (NGOs, the UN, and developing-country governments) had followed by the late 1990s, and a substantial RBM industry had developed. While RBM consisted of window-dressing to a certain extent - so that results-based mismanagement may better describe it in some circumstances - it did shift the focus from development inputs to outputs and outcomes.

For the most part RBM led to the establishment or strengthening of evaluation offices; these were usually internal departments within organizations, but at the same time intended to be independent (OESP, 1997). The implications of this organizational set-up - the self-censorship, negotiation, and compromise so central to evaluation practice - is rarely discussed in evaluation reports. How far evaluation offices are in fact independent varies.

One of the results of the JEEAR was the formation of ALNAP (Active Learning Network for Accountability in Humanitarian Action) in 1997. The genesis behind ALNAP was the need for more coordination of evaluation and learning in the humanitarian sector. It is an international interagency forum with 51 full members representing the main humanitarian actors from within the UN, bilaterals, and NGOs. Among other activities it holds a six-monthly meeting of members and publishes an Annual Review on EHA findings and quality, and maintains an Evaluative Reports Database, which is an online database collection of EHA.

The majority of EHA is carried out by sectoral specialists (e.g. in water, food aid), working individually or in teams. Evaluators in many cases have a technical rather than a social science or evaluation background. One of the consequences of this is that EHA is often evaluation with affected people left out - in other words a focus on food aid “pipelines,” nutritional data and operational agency staff opinions. This is perhaps the most problematic aspect of EHA as currently conducted. (In development evaluation this also happens to a considerable extent, though there is also an important set of practitioners of participatory development evaluation, and an increasing utilization of different ways to incorporate affected peoples’ perceptions.)

EHA also tends to be carried out by expatriates, often from the country providing assistance. While this may ensure there is adequate technical expertise on the evaluation team, it also smacks of tied aid, that is, the requirement in many donor countries that aid money should to some extent provide employment or services to their own citizens. A good example is the spate of evaluations carried out on interventions in Kosovo; of 55 evaluators involved, only three were from the region (ALNAP, 2001). The flip side of this reliance on western evaluators is that little has been or is being done to support evaluation capacity in crisis hit countries, an exception
being UNICEF's support to the African Evaluation Society (Nairobi M&E Network, 2002; AFREA is also supported in its development evaluation work by the African Development Bank, the World Bank, and other organizations such as the bilateral agencies of development assistance of Denmark, the Netherlands, and Norway). This despite the fact that there is evidence to suggest that EHA, which uses teams of national and expatriate specialists tends to be of higher quality, as this combines both local knowledge with international expertise (ALNAP, 2004a).

One of the ironies of this situation is that evaluation managers have noted that lack of well-qualified evaluators is probably the main constraint to improving the quality of EHA. Currently the “market” for evaluators operates like most markets with evaluation officers competing against each other. This means that those who have better knowledge, networks, and funds are at an advantage. While this may benefit individual agencies, it does not support improvement for the sector as a whole, particularly given there are so few joint evaluations.

Findings of Evaluation of Humanitarian Action

Part of ALNAP's work has been a synthesis of findings of EHA (ALNAP, 2001–2004). Over the last four years some 215 evaluations have been included, including 20 synthesis reports. This is probably the most systematic attempt to date to determine the strengths and weaknesses in the humanitarian sector. Several findings can be highlighted from this synthesis.

Humanitarian action is in general successful in meeting its short-term objectives of saving lives, feeding hungry populations, providing short-term health care, water and sanitation, and shelter. This in itself is a remarkable achievement given the often harsh operating environments. On the other hand, humanitarian action has failed to support livelihoods, or to establish a bridge between relief on the one hand and rehabilitation/reconstruction and development on the other. As such, humanitarian action is usually dealing with symptoms rather than causes. Part of this failure to build longer-term solutions has been a lack of attention to capacity development, although there have been pockets of success. The failure to hand decision-making and responsibility for humanitarian action to those who appear to do much of the work - that is, national organizations and staff - is an ongoing theme (e.g. Minear, 2002). Surprisingly little is known about what makes for positive capacity development, and much more is known about what undermines capacity, for example, competition between agencies and in-fighting about control of budgets.

Staff, and in particular national staff, are key to the success of humanitarian action. A large number of dedicated staff work in the sector, often on short-term contracts and in dangerous or stressful environments. At the same time these staff are often not supported or valued by reactive bureaucracies (People in Aid, 2004).
A number of commentators have claimed that humanitarian actors, in particular since the Kosovo conflict in 1999, have been co-opted by western foreign policy priorities (Rieff, 2002; FIFC, 2004; Duffield & Macrae, 2001). UN agencies and NGOs, in particular US NGOs, have been willing to follow their government policy and funding, despite the fact that their governments are belligerent, hence apparently going against the key humanitarian principles of impartiality and neutrality. For example, many humanitarian actors worked in Iraq because of pressure from their governments, the accessibility of funding, and the need for profile, despite the fact that only “pockets” of humanitarian need existed in Iraq. NGOs were also hesitant about criticizing human rights violations, for example by US and coalition Afghan partners in northern Afghanistan (ALNAP, 2004a).

Over the last decade increasing attention has been given to the development of needs assessments to improve aid effectiveness and impact. However, the lack of learning about and from affected populations when all agencies in their policies strive for a participatory approach seriously undermines the credibility of humanitarian action. And while more is known about livelihoods of beneficiaries and how these might be supported than 10 years ago, the mechanisms for translating this knowledge into practice are still underdeveloped (Hofmann et al., 2004). At the same time there are several institutions which have taken learning from communities as a central feature of their work—for example, the Disaster Mitigation Institute in India, Groupe URD, and the Tufts University project on livelihoods in Afghanistan (e.g., Lautze et al., 2002). Constraints to learning from affected populations should not be underestimated, particularly in complex emergencies. Security considerations may make access difficult and curfews may keep contact brief. On the other hand, even in relatively stable situations such as refugee camps, learning from primary stakeholders is often limited.

A central finding from evaluations over the last four years, as well as much other literature (e.g., Minear, 2002; Rieff, 2002) is that the priorities of individual donors and agencies take precedence over a coordinated response, with subsequent loss of effectiveness. However, there has been some limited progress recently. A recent external review of the Inter-Agency Standing Committee concluded: “These [coordination] tools are significantly more developed than they were five years ago. There is evidence that field level coordination has improved, at least among the UN system of agencies and with a sub-set of the major international NGOs” (OCHA, 2003, pp. iii-iv). However, it also found: “much less evidence of progress on solving perennial problems of mandate gaps, capacity gaps, or system-wide problems, or in handling such issues as the ‘transition from relief-to-development’, IDPs, the military-humanitarian interface, etc.”
The Quality of Evaluation of Humanitarian Action

A meta-evaluation of 127 evaluations of humanitarian action produced between 2000 and 2003 as part of the ALNAP Annual Reviews reveals that the overall quality of EHA, as measured against mainstream evaluation standards, is generally inadequate, undermining the credibility of evaluation findings. In addition, a considerable proportion - probably somewhere between 30 percent and 50 percent - of humanitarian aid is allocated not to the emergency or relief phase, but to rehabilitation and reconstruction. However, the most often used evaluation criteria relate to the emergency phase, in particular: the number of lives saved, nutritional and morbidity levels, and numbers who have received emergency shelter. If indicators relevant to rehabilitation were used the performance of humanitarian action would consequently be seen to decline.

Figure 24.1 Three year totals of EHA report ratings, assessed using ALNAP’s Evaluation Quality Proforma.

Twelve areas key to EHA quality are covered in this section. The three-year averages of ratings for these areas are shown in Figure 24.1.

Overall EHA exhibits some strengths, in particular in evaluation of human resources, coordination, and the DAC criteria, but problematic areas remain. The main ones are opaqueness of evaluation methodologies; failure to meet good practice standards in use of methodologies; lack of systematic consultation with primary stakeholders; and failure to use agency policy to assess interventions. Insufficient attention to evaluation use and users remains common. A more detailed assessment of individual areas in the Quality Proforma follows.
Evaluation reports were found to be weak in both of these areas. In general, reports did not specify adequately the key methodological tools that evaluators should use. It was also rare for terms of reference to outline clearly the intended use of evaluation reports, thus adding to the likelihood that the findings of reports will not be fully used.

The first area refers to the theoretical approach drawn upon in the evaluation. As noted above, most EHA is atheoretical. For the second area a satisfactory or higher rating is given where a qualitative and quantitative multi-method approach has been used. To one of us (TB) this seems very controversial, as it appears to imply that mixed method evaluations will always be better quality. OF’s interpretation of this point is that what is being argued here is that a multimethod approach is more satisfactory than one that relies only on one method. This does not seem controversial, as it is one way to “triangulate” in a complex reality. The following have all been adequately addressed: application of the DAC criteria; reference to international standards including international law; consultation with key stakeholders to inform findings, conclusions, and recommendations; triangulation for verification purposes; and gender analysis. As can be seen, a small minority of evaluations fulfill all of these requirements.

Quality of the Evaluation of Agency’s Management and Human Resource Practices

% satisfactory  51
% unsatisfactory49

This area is a strength in EHA, with over 50% of evaluations rated as at least “satisfactory” each year. Evaluators have consistently examined issues such as staff turnover, HQ-field communication, and security. However, other areas such as pre-departure briefing, training, and learning are less well covered.

The cross-cutting theme that consistently scored well was coordination, which is related to the ability of evaluators to cover institutional factors. In the other four theme areas, reports performed consistently poorly except in the case of consideration to the vulnerable and
marginalized where performance was somewhat better.

Much of the humanitarian system, including most of the UN operational agencies, and many NGOs, have been moving to what has been termed a human-rights-based approach to programming over the last few years. This means that human rights, as defined by international conventions such as the Convention on the Rights of the Child, provide an overarching framework for interventions. EHA, however, has not kept pace with changes in humanitarian programming, although it should also be mentioned that many agencies have adopted a rights-based approach as policy but this has yet to be turned into practice. The link between the cross-cutting themes of international standards, protection, and gender equality is that they deal with rights-based issues that are often controversial; these are the issues that are most often left out of evaluation terms of reference and with which evaluators appear to have the least skills. As noted, the single-agency, single-sector focus of much EHA tends to lead to a technical, apolitical focus. A good example is protection, referring to protection of the affected population from harm, which has been at the center of much discussion of humanitarian action in the last few years. The fact that this is so poorly covered in EHA is a central gap, and EHA is clearly a long way away from integrating a rights-based approach into a wider evaluative process.

Quality of Consultation with and Participation by Primary Stakeholders

% satisfactory 13
%unsatisfactory 87

Evaluation reports are expected to outline the nature and scope of consultation with, and participation by, beneficiaries and non-beneficiaries within the affected population in the evaluation process. This is a further area of weakness. Despite some good practice, EHA could rightfully be accused of systematically ignoring the views and perspectives of primary stakeholders in favor of those of institutional actors, particularly agency staff. This undermines its credibility and continues in the vein of humanitarian aid itself treating primary stakeholders as passive recipients of aid rather than active participants in their own recovery. This agency-centric perspective will only change if commissioning agencies insist on adequate primary stakeholder consultation and participation. But an equally important constraint would appear to be the structure of evaluation missions, which are usually short forays by foreign-based evaluators, with a focus on national capitals.

Application of the DAC Evaluation Criteria

% satisfactory 50
Application of the DAC criteria is one of the stronger areas of EHA. Overall it is possible to conclude that evaluators have had reasonable success with their application and that they have become EHA's central evaluative tool. Some of the DAC criteria, in particular efficiency and coherence, have proven more problematic for evaluators.  

The Future of EHA

As noted, EHA has in the last few years been incorporating new methods to overcome some of its inherent problems. Real-time evaluation, pioneered by UNHCR (2002) and used as well by WFP, and UNICEF (2003), is increasingly being used to attempt to feed lessons learned into ongoing programs. As currently undertaken, real-time evaluation has some overlaps with monitoring, and some similarity to utilization-focused approaches in that it works with program stakeholders on assessment of key intervention issues during the course of the intervention, attempting to come up with evaluation recommendations which can be utilized immediately. This focus on iterative lesson learning is to be welcomed, and has had some immediate impact. Exemplar 24.2 presents details of a real-time evaluation in Southern Africa undertaken by WFP.

EXEMPLAR 24.2 Real-Time Evaluation: Good Practice Case of WFP's Southern Africa Evaluation

Recent definitions of real-time evaluation (RTE) have stressed the importance of its timing, its interactivity, and participation of key users of information. As such it bears some similarity to utilization-focused evaluation, which emphasizes the evaluation process and the nature of the interaction between evaluation users and evaluator (Grob, 2003).

WFP piloted the RTE approach in impressive fashion in the evaluation of its response to the Southern Africa crisis (WFP, 2003). The evaluation included three missions between July 2002 and May 2003, with a summary and recommendations being submitted to the WFP Executive Board in October 2003.

In terms of staff time and committed resources, this was one of the most comprehensive evaluations undertaken by WFP, taking about 13 weeks in total. The core team for all three missions were a socio-economist, a nutritionist, and a needs assessment specialist. In addition, two separate logistics specialists were engaged for the second and third
RTE was a new departure for WFP. Its previous evaluations were often of high quality, but used a mix of standard EHA techniques, with occasional reports being strong on consultation with primary stakeholders. The Southern Africa evaluation, on the other hand, organized its three missions to coincide with three phases of the project cycle: the inception period; the start of the most intensive implementation phase; and the closing stage of the Emergency Operation (EMOP), and preparation for the new EMOP. While experimenting with the RTE approach, however, the evaluation sometimes reads like a traditional *ex post* evaluation, and its focus appears at times to be as much on accountability as on feeding lessons back on an ongoing basis into the WFP system.

The longer evaluation period and the greater length of time spent in country appeared to facilitate more in-depth consultation with primary stakeholders. This took two forms:

- Focus group discussions with beneficiaries at Final Distribution Points in each country. This included observing distribution methods, discussions with distribution committees, and on-site briefings from NGOs.
- Household visits - at least two (in most countries) at each food distribution point visited - where in-depth semi-structured interviews with primary stakeholders were carried out. An ad hoc “sentinel site” study was also undertaken in Malawi and Zambia, where households were visited during each mission for an update on progress, and an insight into the impact of the operation.

The RTE (WFP, 2003, p. 3) notes that: “one great advantage of the RTE was being able to visit project sites as they were in full operation and with key staff still present,” a point also emphasized in the UNICEF (2003) desk review of RTEs. This is particularly important given that many ex post evaluations are seriously constrained by key staff having moved on.

One other area of strength of the evaluation is its unusually detailed contextual analysis. Ex post evaluations sometimes have difficulty retracing key events. The WFP evaluation on the other hand, included a very detailed assessment of key events vis-à-vis the timing of the intervention. As the terms of reference note, RTE can help build a lasting description or chronicle of the intervention, which can later be used as a model for identifying lessons and good practice.

Agencies are in many cases already committed - through their policies and evaluation guidance
- to covering adequately a number of the weaker areas in EHA identified above, and through its meta-evaluation ALNAP continues to support improvement in evaluation practice, including through ongoing dialogue with agencies. Some of the weaknesses in EHA may more easily be improved than others, for example, lack of transparency in detailing methodology. Other weaknesses are likely to be more intractable, for example: lack of analysis of social process and gender, and weak attention to rights-based issues such as protection; limited consultation with beneficiaries; and inadequate focus on utilization. EHA will need to be strengthened considerably both in its theoretical underpinnings and practice if its findings are to be considered credible.

Conclusion

The practice of evaluation of humanitarian action (EHA) and development evaluation (DE) initially focused almost exclusively on single agency, single sector/project evaluations; though this type of evaluation still dominates both EHA and DE, joint development evaluations involving several agencies started to become more frequent in the early twenty-first century.

Whereas EHA tended to be atheoretical, DE draws on social sciences and evaluation approaches, and both make use of procedures and tools such as document and literature reviews, surveys, focus groups, and structured and semi-structured interviews. Also both types of evaluations are structured around the DAC evaluation criteria.

The concern with the lack of sufficient use of evaluation findings led to innovations in the way in which these evaluations are conducted (with more emphasis on participatory evaluation, which promotes empowerment and capacity building), and new ways to communicate findings (e.g., through the press and through workshops).

Networks of evaluators in the north and in the south involved in EHA and/or DE were formed at the end of the twentieth century, and national and regional evaluation associations, most of whose members are involved in DE and/or EHA, emerged in developing countries. These “communities of practice,” that are already playing a role in developing evaluation cultures in different countries and regions, and in building evaluation capacities, may also promote a further convergence (within the limits imposed by their different contexts) between the evaluation of humanitarian action and development evaluation.

Notes

1. Osvaldo Feinstein has been responsible for the part of the chapter on development evaluation and has completed some general revision work. Tony Beck has written the
humanitarian evaluation parts of the chapter.

2. The use of the term “development” in this chapter corresponds to the broad field of practice and studies that focuses on poverty reduction in “developing countries”

3. IDEAS was launched in 2002, and its Board has a majority of developing-country members. See references for website.

4. This section is based on OECD-DAC (1999).


6. See DAC Working Party on Aid Evaluation (2001). This publication also includes a valuable discussion about feedback and the use of evaluations for learning and accountability.

7. For example, the ex-post impact study of the Noakhali Rural Development Project in Bangladesh, conducted by Danida (2002).

8. See DAC Working Party on Aid Evaluation (2002). This glossary was prepared by a task force which included evaluation experts of bilateral and multilateral development agencies. See also DAC (1998), Cracknell (2000) and Molund & Schill (2004).

9. The “logical framework” is a tool used to improve the design of interventions, most often at the project level. It involves identifying strategic elements (inputs, outputs, outcomes, impact) and their causal relationships (the “results chain”), indicators, and the assumptions or risks that may influence success failure. See DAC (2002).


11. The original table in Carvalho & White (2004) includes “competition effects” among the “main channels,” as well as some additional assumptions and types of data that for the sake of simplification were omitted here.

12. Albert Hirschman, one of the most eminent development economists, published two books that are goal-free evaluations (though he never considered himself an evaluator). In one of them he evaluated projects of the Inter-American Foundation (see Hirschman, 1984), and the other one (Hirschman, 1995) is an evaluation of projects supported by the World Bank.

13. Participatory methods in development (including participatory monitoring and evaluation) have been pioneered by Robert Chambers, (see, e.g., Chambers, 1994; Feuerstein, 1992. In this handbook Chapter 9 by Stevenson & Thomas also makes reference to participatory
methods.

14. See the discussion of this approach in Pitman, Feinstein, & Ingram (2004), particularly the notion of “opportunistic randomization,” for example where the resources of an intervention have to be rationed, and randomization/lotteries are used as a rationing device, thus allowing RCT without creating ethical issues that otherwise may arise. By the way, the use of the analogy with the “gold standard” by the strong supporters of randomization is rather ironic, given the acknowledged failures of the “gold standard.”

15. See Dabelstein & Rebien (2002), which summarizes the different types of joint evaluation.

16. See Feinstein & Ingram (2003) for this sort of limited approach to joint evaluations (given the difficulties experienced in implementing full-fledged joint evaluations) emphasizing processes, including examples and a set of hypotheses on costs and benefits of joint evaluations.

17. For information on IPDET and on diagnostic tools for evaluation capacity development, see www.worldbank.org/evaluation/ecd, which also provides information on short evaluation courses.


19. See, for example, Sodersten & Reed (1990), p. 322.

20. The use of explicit counterfactuals in the evaluation of program aid is well presented in White & Dijkstra (2004). This book also includes an important and clear discussion of fungibility and its implications for the evaluation of aid, which due to space limitations cannot be treated in this chapter. On A. Sen's approach, see Sen (1999).

21. There are other evaluation training programs in Chile, Barbados, and Mexico. There are many more training programs that focus on “project appraisal,” though frequently they are presented as courses on project evaluation (indeed, their focus is on ex-ante evaluation).

22. Details on the national and regional evaluation organizations that have been formed in the last years can be seen at www.mande.co.uk.


24. This section draws heavily on the work of TB for the Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP), and the four years of the ALNAP Annual Review. For further details see www.alnap.org. This section deals with
humanitarian aid to developing countries; there is an extensive literature on evaluation of emergency aid in developed countries (e.g., Quarantelli 2000), which is not covered here.


26. This section is drawn from ALNAP (2003). The meta-evaluation was performed using the ALNAP Quality Proforma, a meta-evaluation tool developed from good practice in EHA, the American Evaluation Association Program Evaluation Standards, and following Stufflebeam (2001). The Quality Proforma set out specific criteria against which evaluations are to be assessed on a four-point rating scale of: poor, unsatisfactory, satisfactory, and good. As slightly different rating systems were used over the three years covered, analysis has been carried out using “satisfactory” and “unsatisfactory” ratings only. A satisfactory or better rating was achieved when evaluations met good practice standards. For further details see ALNAP (2004b).

27. For definitions see ALNAP (2004b).

28. One example is the UNHCR Afghanistan real-time evaluation, where the evaluators had an audience with the UNHCR Executive Board immediately on return from Afghanistan.

References

Relevant Websites

www.ecgnet.org
www.euforic.org
http://evaluation.wmich.edu/jmde/JMDE_Num001.html
www.ideas-int.org
www.intrac.org
www.mande.co.uk
www.oecd.org/dac/evaluation
www.undp.org/eo
www.worldbank.org/evaluation
www.alnap.org
http://www.unhcr.ch/cgi-bin/texis/vtx/research?id=3b850c744

http://dx.doi.org/10.4135/9781848608078.n24