

Re-thinking the Impact of Humanitarian Aid: Background Paper for the 24th ALNAP Biannual

1. Background and Context

The aim of this paper is to provide an overview of current debates concerning impact assessments and to help clarify key issues around their application and utilisation in humanitarian aid.

Our intention is to frame these issues so as to support discussion and debate at the forthcoming 24th ALNAP Biannual in Berlin.

It is hoped that the discussions at the Biannual, and subsequent work, will enable the ALNAP membership and the wider humanitarian community to develop a shared understanding of the limits and possibilities of humanitarian impact assessment, and to use this understanding to outline a practical vision for future work in this area.

2. Introduction

Prior to the 1990s, few humanitarian organisations even thought to measure the consequences of their actions, assuming that the mere provision of their assistance activities was evidence of their good results (Barnett, M. 2005; Crisp, J. 2004). As Barbara Harrell-Bond observed in 1986,

“Humanitarian work...is thought to be selfless, motivated by compassion, and by its very definition suggests good work...As relief is a gift, it is not expected that anyone (most especially the recipients) should examine the quality or quantity of what is given” (Harrell-Bond, B. 1986 cited in Crisp, J. 2004).

This attitude has changed dramatically over the last two decades. In a context of limited resources and growing concern about humanitarian performance, there is now much more analysis and scrutiny of humanitarian operations, demonstrated by the growing number of evaluations and accountability mechanisms being applied within the sector (Clarke, P. and B. Ramalingam, 2008). As Jeff Crisp notes, the evaluation of humanitarian action has become “big business”, attracting new levels of donor funding and agency commitment, as well as public and political interest (2004). However, the role of beneficiaries highlighted by Harrell-Bond twenty years ago remains at best a partial one in the planning, monitoring and evaluation of humanitarian aid (Kaiser, T. 2004).

Despite these developments, evidence about the impact of aid interventions (whether humanitarian or development) is not forthcoming and impact assessment, nominally the day-to-day work of evaluation departments, is held by some critics to be insufficiently rigorous in the sector (CGD, 2006; Forss, K. and S. Bandstein, 2008).

Calls for more evidence-based policy combined with methodological advances for assessing and attributing impact (White, H. 2007) have led to increased interest in impact assessment. This has been characterised by a shift in focus from *process* and on how funds are spent, to what *effects* interventions are having on the populations they claim to serve.

There are now both explicit and implicit commitments by aid agencies and convening bodies to assessing impact. In the wider aid sector, important work is being led by the OECD-DAC Evaluation Network, the Network of Networks on Impact Evaluation (NONIE), and the International Initiative on Impact Evaluation (3IE).

Within the humanitarian sector, there are a number of interagency initiatives aimed at monitoring and evaluation of impact. These include the Standardized Monitoring and Assessment of Relief and Transitions (SMART), the Health and Nutrition Tracking Service (HNNTS), the Tsunami Recovery Impact Assessment and Monitoring System (TRIAMS), the Fritz Institute Humanitarian Impact project; and the Collaborative for Development Action Listening Project. ALNAP is currently exploring the feasibility of a mechanism for assessing and reporting on the overall performance of the humanitarian sector, which will have a focus on understanding 'collective impact'.

Among the individual agencies, WFP, ECHO and UNICEF all include impact in their evaluation guidelines (Watson, C. 2008). Several NGOs have introduced impact assessment systems that aim to improve accountability at the organisational level, for example, ActionAid's Accountability, Learning and Planning System (ALPS) and Save the Children (UK)'s Global Impact Monitoring (GIM) (Hofmann, C.A. 2004). The ECB agencies have produced a widely disseminated 'Good Enough Guide' to impact measurement (ECB, 2008) while the DEC's new Accountability Framework also addresses the question of the impact of its member organisations.

But despite considerable progress, use of impact assessments has not become common practice. A great deal of confusion remains regarding the conceptualisation and definition of impact assessment, the range of approaches, tools and methodologies, and whose needs these meet.

The rest of this paper provides emerging findings from ongoing ALNAP research, synthesising the available literature and interviews to inform discussion and debate at the 24th Biannual.

Section 3 of this paper highlights the wider contextual factors contributing to the current interest in humanitarian impact assessment.

Section 4 outlines some of the issues surrounding the conceptual and definitional debates around humanitarian impact assessments.

Section 5 briefly considers the use of impact assessments and reviews current motivations for doing them by looking at different stakeholders and their interests in impact.

Section 6 considers impact indicators, baselines and data issues in emergency contexts.

Section 7 explores appropriateness of different methods and summarises related theoretical and ideological debates which influence the selection of methods.

Section 8 addresses the need for better engagement with affected populations and other stakeholders for improved analysis of outcomes and impacts.

Section 9 highlights some of the institutional challenges related to capacities and incentives for doing good quality impact assessment;

Finally, **Section 10** provides a brief summary and suggests some important areas for discussion at the 24th Biannual Meeting.

3. Humanitarianism ‘transformed’ and the emergence of impact assessment

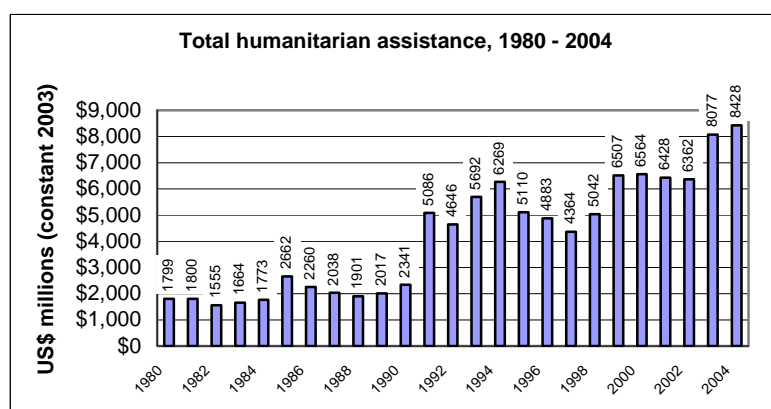
Contemporary interest in humanitarian impact assessment needs to be understood in the context of several changes that the sector has undergone since the late 1980s. Broadly speaking, these can be categorised into three trends driven by forces within and outside the sector: 1) the *politicisation* of humanitarian aid; 2) the *institutionalisation* of the humanitarian ‘sector’; and 3) the *changing nature of vulnerability* in humanitarian crises. In the following paragraphs, the implications of these trends for understanding humanitarian ‘impact’ will be explored.

1) Humanitarian purpose and principles have expanded and become politicised.

Whereas once humanitarian actors attempted to insulate themselves from the world of politics, today humanitarianism is increasingly implicated in global governance. Many humanitarian agencies now work closely with states and attempt to eliminate the root causes of conflict that place individuals at risk (Barnett, M. 2005). Nowhere is the politicisation more evident than in high-stakes political crises such as Afghanistan, the occupied Palestinian territories and Iraq (Donini, A. et al. 2008). A number of factors contributed to this shift:

- Funding of humanitarian action by Western states increased following the end of the cold war (see Figure 1); this led to the alignment of many of the major international NGOs, whether by institutional conviction or as a by-product of their funding sources, with the geopolitical interests and foreign policy objectives of donor governments (Barnett, M. 2005; Clarke, P and B. Ramalingam, 2008; Slim, H. 2006; Donini, A. et al. 2008; Foley, C. 2008)¹.
- The nineties witnessed an unprecedented number of large-scale humanitarian operations (notably Somalia, southern Africa, former Yugoslavia, Rwanda, Central America, Kosovo, and East Timor), many involving armed intervention in the form of peace enforcement or peacekeeping operations (ALNAP 2001). As part of ‘system-wide’ responses to these emergencies, relief and development workers began to occupy the same spaces and articulate a relief-rights-development discourse (Barnett, M. 2005).
- Humanitarian functions began to be taken on by military, paramilitary, development actors and commercial companies with a variety of mandates and operating principles (Donini, A. et al. 2008; Macrae, J. 2002).

Figure 1: Total Humanitarian Assistance 1980 – 2004



Source: OECD DAC

¹ It is important to note that some NGOs have tried to resist co-option by these governments, notably MSF and the ICRC. MSF and Oxfam GB have refused Western government funds in wars where the donors themselves have been outright belligerents (Slim, H. 2006).

Combined, these factors fuelled increased interest in the sector from the media and the general public (Buchanan-Smith, 2003; Clarke, P and B, Ramalingam 2008), placed humanitarianism at the centre of the international policy agenda and opened up new opportunities for expanded meanings of humanitarianism (Barnett, M. 2005). Today, a longstanding philosophical debate about the meaning, ethics, politics and delivery of humanitarian assistance continues unresolved. This debate has largely centred on the question of where the core humanitarian principles of neutrality and impartiality end, and engagement with states and politics begin (Collinson, S. 2002; Donini, A. et al. 2008).

Box 1 below broadly illustrates the various positions held by humanitarian actors, ranging from those that remain loyal to a “Dunantist” ethos of providing neutral, impartial, life-saving assistance and protection, to those who go beyond this and actively engage more systematically with reconstruction, development, trade, conflict resolution, advocacy and the human rights agenda.

Box 1: A Typology of Humanitarian Actors

Principled. Some aid agencies affirm the continued relevance, indispensability, and centrality of humanitarian principles. This is particularly true of those with a long history of operating in conflict settings and with roots in the “Dunantist” tradition—i.e. based on the basic tenets of humanitarianism developed by Henri Dunant, the founder of the ICRC. In general, the proponents of principle-centred action argue for a narrower definition of humanitarianism limited to life-saving assistance and protection of civilians, based on core principles of neutrality, impartiality, and independence.

Pragmatist. Other agencies recognise the importance of principles but place a higher premium on action, even when this means putting core principles in jeopardy. Prominent among agencies of the pragmatist persuasion are “Wilsonian” agencies: those that identify broadly with foreign policy objectives of their home government, whose funds they often use.

Solidarist. A third path, embraced by some organisations on both sides of the Atlantic, goes beyond the provision of assistance and protection to address the root causes of conflict, which are political at the core. Wider than the traditional humanitarian brief, their anti-poverty thrust and social transformation agenda mixes elements from humanitarian, human rights, and developmental world views, with heavy emphasis on advocacy.

Faith-based. The world’s major religious traditions, western and non-western alike, embody humanitarian affirmations and obligations. The Christian tradition, expressed for example in missionary work, affirms the core values of compassion and charitable service. Islam embraces similar core values and has created organisations to express them that are in some respects similar to western religion-rooted organisations. In addition to international faith-based agencies that do not usually engage in proselytising, there is a wide variety of religious organizations at the local level. Faith-based entities may themselves embody principled, pragmatist, and/or solidarist features.

Source: Donini, A. et al. (2008) *The State of the Humanitarian Enterprise*. Feinstein International Center, Tufts University.

Where in this typology an institution positions itself has important implications for its conceptualisation of *intended* impacts, as well as its approach to impact assessment.

2) The humanitarian sector has become institutionalised. In response to the increasing politicisation of aid and the critiques of humanitarianism (e.g. de Waal, A. 1997) that it fuelled, humanitarian agencies began to examine the ‘dark side’ of what they were doing (Roche, C. 2008; Slim, H. 1997). Most notably, “Do No Harm” project focused on learning more about how humanitarian assistance interacts with the dynamics of conflict and how aid can be used and misused to pursue political and military advantage. The Do No Harm “Analytical Framework” provided a tool for understanding and mapping the complex interactions between

assistance and contextual factors in order to anticipate and prevent the possible negative impacts of interventions (Anderson, 1996).

Greater efforts were also put into improving accountability and performance of humanitarian action. Significantly, the system-wide evaluation of the international humanitarian response to the genocide in Rwanda (JEEAR, 1996) provided impetus and momentum for a number of initiatives aimed at increasing the quality and accountability of humanitarian aid and improving learning. These included the Sphere Project, ALNAP, the Ombudsman project, the Humanitarian Accountability Project², and People in Aid (Barnett, M. 2005; Slim, H. 2006; ALNAP, 2003). Of these initiatives, ALNAP has focused on improving evaluation quality and utilisation as a key objective. More recently, UN Reform and the Good Humanitarian Donorship (GHD) initiative have focused on changing the architecture of humanitarian assistance in order to improve accountability and predictability of aid.

It is important to recognise that the above developments within the international humanitarian system took place against a backdrop of fundamental change in the management of national and international public sector organisations (ALNAP, 2001). A central element of this reform was the shift from an input–output management model towards a greater emphasis on results. A culture of setting targets, measuring performance and assessing achievements in quantifiable terms permeated through to the international humanitarian system. Results-based management is now a dominant approach to assessing performance and impact (Hofmann, C.A. et al. 2004; *ibid*).

By creating more regular interactions amongst humanitarian aid agencies, increasing information and knowledge sharing, and developing abstract rules to guide standardised responses (Barnett, M. 2005), the institutionalisation of humanitarianism has created a climate and context in which humanitarian impact can be debated, discussed and operationalised.

3) The number and scale of emergencies has increased dramatically (Clarke, P and B. Ramalingam, 2008). Moreover, the Global War on Terror (GWOT), processes of globalisation, environmental degradation and climate change are changing the nature and the geography of human suffering (Donini, A. et al. 2008). The crises of tomorrow are arguably more likely to be urban than rural; related to weak governance structures than to traditional forms of armed conflict; and about access to and distribution of dwindling resources (*ibid*).

Meeting these humanitarian challenges will require good knowledge about what works and what doesn't. While calling for a better understanding of impact, they also simultaneously pose further conceptual, methodological and practical challenges to an already complicated field.

From the literature review and interviews conducted so far, a number of key themes emerge as especially relevant for further discussion and analysis, each of which is explored in turn in the rest of this paper.

4. Defining humanitarian impact assessment

One of the many challenges facing humanitarian impact assessment is the lack of clarity surrounding its definition and purpose. At its simplest, humanitarian impact assessment is about understanding the changes in people's lives brought about by an intervention. Yet, as Howard White observes, the term 'impact assessment' has taken on different meanings in the aid sector, including the following:

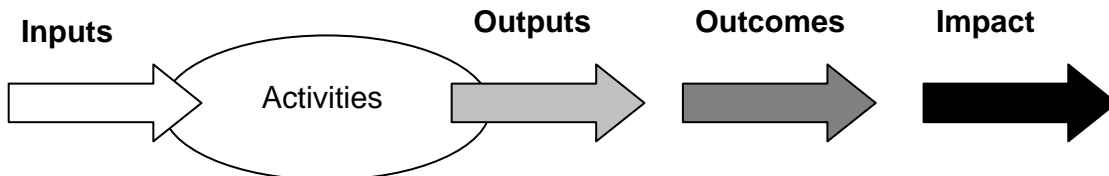
- beneficiary assessment, named as beneficiary or participatory impact assessment,
- a study focusing on final welfare outcomes,
- an evaluation carried out some time after an intervention has finished,
- a study encompassing a whole country or sector,

²Now the Humanitarian Accountability Partnership International

- a study employing counterfactual analysis, and
- studies with a specific focus, such as 'environmental impact analysis' (White, H. 2007).

The aid evaluation community has traditionally viewed 'impact' as the final step in a linear results 'chain', which begins with an intervention's inputs, which lead to activities, which generate outputs, and then outcomes (often called effects or results), and which ultimately lead to impacts (see fig 2).

Figure 2: The Impact Chain



Source: Roche 1999: 26

While this model is a useful tool, its linear logic has been subject to criticisms, a number of which are pertinent to understanding humanitarian impact. It has been argued that it oversimplifies the causal pathway. In reality the relationship between cause and effect may be more fluid or cyclical, meaning that it is not always possible to establish linear causality between a specific intervention and its impact (Pawson, R. 2003, cited in Hospes, O. 2008; Watson, C. 2008). Moreover, this logic forms the basis of 'results-' or 'objectives'-based approaches to analysing impact. This attempts to analyse project or programme effectiveness according to a set of *pre-defined* institutional objectives (Hospes, O. 2008; Bolton, J. et al. 2007). This approach has benefits for policy, evaluation and financial audit departments, but also poses a number of problems.

Defining impact in an emergency context involves the need first to define humanitarian aid itself. This is not straightforward given rapidly changing humanitarian contexts and the lack of consensus regarding the goals and objectives of humanitarian aid, highlighted in the previous section. While there is no accepted definition, it is frequently defined by the humanitarian community as:

"The systematic analysis of the lasting or significant changes - positive or negative, intended or not – in people's lives brought about by a given action or series of actions" (Roche, C. 2000).

This definition was initially adopted by Oxfam GB in an attempt to overcome the shortcomings of development sector definitions of impact, which emphasised 'lasting' change and failed to capture the shorter-term and immediate nature of some humanitarian impacts. (Roche, C. 2000). By focusing on both intended and unintended impacts, the above definition recognises that processes of social change are rarely solely the product of a managed process undertaken by humanitarian agencies, but rather, the result of wider, complex historical, social, political, economic and environmental factors (ibid).

Humanitarian agencies arguably have greatest control over key decisions and events at the inputs, activities and outputs end of the 'chain', for example, project budgets and design, choice of partners, location, timing, etc. However, once the programme is underway, the balance of influence should begin to change, and wider contextual factors beyond the control of an individual agency become more prominent. Taking these wider factors into account is considered particularly important in emergency contexts in order to ensure that the minimum standard of 'do no harm' is maintained (Watson, C. 2008).

In this light, some commentators have suggested that linear 'results-based' approaches to impact assessment should be viewed with caution and complemented by approaches that situate intended impacts in the broader environment (Hofmann, C.A., et al. 2004). The ALNAP EHA Guide usefully distinguishes between 'effectiveness', which examines whether

the intermediate objectives of the intervention have been achieved, and 'impact' which is perceived as going beyond intervention-planning documents to consider the part an intervention plays in wider socioeconomic and political contexts. However, the goals and objectives of many interventions are often either unclear or overambitious, which complicates the process of anticipating or making judgements about *effectiveness* or *intended impacts* (Watson, C. 2008; NORAD, 2008).

This linear results chain also lies at the heart of a longstanding debate about attribution; namely, the idea that we can know that a particular project or set of inputs has brought about a lasting or significant change in the environment or well-being of a number of targeted beneficiaries. Methodologically, this requires isolating the key factors that caused a desired result and attributing it to a particular agency or set of activities (Smutylo, T. 2001).

While some argue that failure to attribute impact risks wasting public resources or even harming participants (CGD, 2006), others believe that attribution in humanitarian contexts is neither possible, nor desirable given the complex, non-linear reality of humanitarian aid described above. A growing number of analysts are recognising that in many cases it might be useful and liberating to assess the *contribution* that an agency makes to a given process (Roche, C. 2008; Hofmann, C.A. 2004; Watson, C. 2008). This approach involves looking at the overall collective impact of a set of interventions, and then working to establish what the contribution of a specific agency or set of agencies was to that collective impact. This principle is at the heart of the IDRC's Outcome Mapping approach, and has also found support from agencies within the humanitarian sector.

However, this approach requires collective agreement about objectives, indicators, and methods used for data collection. As the ALNAP Study on monitoring suggests:

"...All humanitarian agencies retain distinctive mandates, different funding lines, ways of operating, and organisational cultures. And while there may be pressure for more collective action there is equally strong opposing pressure due to the competition for funding and profile between agencies... Future research projects should study in more detail the costs and benefits of sustained collective assessments throughout a jointly implemented programme..." (ALNAP 2003a)

It is also important to consider the scope and timing of humanitarian impact assessments. Efforts may focus on the impact of a particular humanitarian intervention or the aggregate impact of an organisation, country-wide programme, or the sector as a whole (Chapman, J. and A. Mancini 2008; Hofmann, C.A. et al 2004; ALNAP, forthcoming 2009). Timing depends on a number of factors including the stage of the emergency, beneficiary calendars (harvests, migration etc.), and whether or not the findings are intended to have an influence on the intervention being assessed. Some impacts only become visible many years after an intervention has ended (Oakley, et al. 1998 cited in Watson, C. 2008). However, recent evidence suggests that often, the timing of impact assessments and evaluations depends more on the decision-making needs of the agencies rather than on the requirements to conduct reliable research (Forss, K. and S. Bandstein, 2008).

Key Questions around 'Defining humanitarian impact assessment'

- How can aid agencies define and analyse intended and unintended impacts in order to ensure that they do more good than harm?
- How can impact assessments deal with the inherent complexity of the real world?
- Is it possible, or desirable to attribute humanitarian impact to specific interventions? Or is analysis of contribution a more realistic approach?
- What is the appropriate scope and timing of humanitarian impact assessments in different contexts and for different kinds of emergencies?

5. Stakeholder interests in and needs for humanitarian impact assessment

The question of whether humanitarian aid works, why it works and what can be done to make it work better is of interest to a broad range of stakeholders. These include UN agencies, donor governments and publics, recipient governments, international and national NGOs, the media and most importantly, those affected by disasters.

There is a wealth of evidence to suggest that evaluations (including impact assessments) are more likely to be used if the design and process fits the purpose and the needs and interests of the end users. Planning for follow-up and making adequate time for meaningful participation of end users, credible evidence, accessible and relevant recommendations are all essential to ensure relevance, ownership and systematic consideration of findings (Patton, 1997; Sandison, P. 2006). It is therefore helpful to ask who wants to know about humanitarian impact and why?

To date, the objectives of impact assessments have predominantly related to the two institutional priorities of accountability and learning (Kaiser, T. 2004).

Accountability may be in two directions: most commonly 'upwards' to headquarters and/or donors; and less frequently 'downwards' in terms of reporting and accounting to beneficiaries and other stakeholders in the recipient community (Watson, C. 2008).

Learning, on the other hand, involves the use of impact assessment findings by donors and aid agencies to generate information and perspectives on what strategies and approaches are more or less effective in different contexts in order to improve current and future decision-making and initiatives (Watson, C. 2008; CGD, 2006). This can take place internally, or across organisations.

Importantly, the diverse needs of different stakeholders may not be reconcilable and achievable in a particular impact assessment exercise. Sandison notes that part of the problem facing humanitarian evaluations is that too much is expected by too many different stakeholders. In her chapter on Evaluation Utilisation in ALNAP's *Review of Humanitarian Action* she observes that, "if we continue to expect evaluation to cover most of the accountability needs of the sector, we will be disappointed" (Sandison, P. 2006). The same could arguably be said about impact assessment, which after all is just one type of evaluation.

Table 1 below, adapted from Hofmann et al., presents a range of stakeholders and the reasons why they might be interested in impact assessment. It is intended to be illustrative, rather than exhaustive, and can provide a starting point for clarifying the needs and interests around impact assessment.

Table 1: Who wants to know about humanitarian impact and why?

Scope of impact assessment	Who wants to know and why?
Impact of projects	<p>Aid agencies, in order to improve their work, demonstrate impact and make choices between projects</p> <p>Donors, to choose what to fund and to develop policy</p> <p>National governments, to guide disaster preparedness, planning and response</p>
Impact of organisations	<p>Aid agencies, to demonstrate success and raise money from the public and from donors</p> <p>Governance structures may require information about impacts in order to assess the strategic direction of the organisation.</p> <p>Donors, to choose between competing agencies or to make choices about whether to use NGOs or private contractors</p> <p>National governments, to choose who to register and work with as partners</p>
Impact at a sectoral level	<p>Aid agencies and donors, to build up the evidence-base for what works; to develop sectoral policies and best practice protocols and guidelines</p> <p>National governments, to put protocols in place</p>
Country level or for a particular crisis	<p>Donors, to know how many lives were saved</p> <p>National governments, to assess whether appealing for international aid was the right thing to do</p> <p>Agencies, to advocate for increases in levels of aid</p>
Impact of international engagement in a crisis, including, but not limited to, humanitarian aid	<p>Governments, to review their overall engagement with countries in crisis (diplomatic, political, military and aid)</p> <p>Aid agencies and humanitarian donors, to be clear about the role and scope of humanitarian aid</p> <p>Agencies, to advocate for greater political engagement in 'forgotten crises'</p> <p>Governments, to promote the coherence of political and humanitarian agendas</p> <p>Aid agencies, to maintain the neutrality and independence of assistance</p>
All levels	<p>Affected populations to ensure that the knowledge generated by impact assessments is accurate and used in ways that improve their lives; and to hold governments and agencies to account.</p> <p>Donor publics to know whether the money they donated made a difference and hold to account those in whose name funding is raised</p> <p>Campaigners and think-tanks in order to advocacy work aimed at influencing the formulation of policy and best practice guidelines for humanitarian programming.</p> <p>Academics, in order to further their research interests</p> <p>The media, to hold to account those in whose name funding is raised, or to sell stories.</p>

Key Questions around 'Stakeholder interests in and needs for humanitarian impact assessment'

- Whose needs are currently met by humanitarian impact assessments?
- Are the objectives of humanitarian impact assessments clear?
- Are the different needs and objectives realistically achievable in a given impact assessment exercise?

6. Indicators, Baselines and Data for Humanitarian Impact Assessment

Indicators

The identification and use of relevant indicators is a crucial part of determining the impact of an intervention. However, the question of what are the appropriate forms and means of impact measurement in relation to humanitarian action is highly debatable. Impact assessment inevitably involves *value judgements* about which kinds of changes are significant for whom (Roche, C. 2000). As already highlighted, different stakeholders – donors, NGOs, beneficiaries - may have very different motivations for impact assessment. This can translate into distinctive criteria for assessing positive or negative impacts. It is therefore important to involve as many stakeholders as possible in identifying appropriate indicators for impact.

Although the terminology sometimes varies, there are generally two types of indicator - those that relate to the implementation of programmes (input, process and output indicators) and those concerned with the effects of the programme (outcome and impact indicators). Table 2 highlights the different types of indicators that agencies use (from Hofmann, C. A. et al. 2004):

Table 2: Types of indicator: example of measles immunisation programmes

Implementation of the programme			Effect of the programme	
Input indicator	Process indicator	Output indicator	Outcome indicator	Impact indicator
No. of vaccines administered	No. of people trained	Percentage vaccinated	Measles cases decrease	Mortality decreases

Humanitarian agencies tend to use a mix of indicators, depending on their own monitoring and reporting systems and the particular function of the indicators collected (Hofmann, C. A. et al. 2004). Evidence highlights that humanitarian agencies tend to collect process rather than impact indicators. For example, the SPHERE indicators - probably the most comprehensive attempt to define standards and indicators for most areas of humanitarian aid (Sphere, 2004) - focus largely on process and are not designed to show impact. The distinction between outputs and outcomes and the challenges of moving beyond outputs are usefully illustrated in the citation from Michael Quinn-Paton in Box 2.

Box 2: Moving from outputs to outcomes

‘The familiar adage [“you can lead a horse to water, but you can’t make it drink”] illuminates the challenge of committing to outcomes. The desired outcome is that the horse drinks the water. Longer-term outcomes are that the horse stays healthy and works effectively. But because program staff know they can’t make a horse drink water, they focus on the things they can control: leading the horse to water, making sure the tank is full, monitoring the quality of the water, and keeping the horse within drinking distance of the water. In short, they focus on the processes of water delivery rather than the outcome of water drunk. Because staff can control processes but cannot guarantee attaining outcomes, government rules and regulations get written specifying exactly how to lead a horse to water. Funding is based on the number of horses led to water. Licences are issued to individuals and programs that meet the qualifications for leading horses to water. Quality awards are made for improving the path to water – and keeping the horse happy along the way. Whether the horse drinks the water gets lost in all this flurry of lead-to-water-ship. Most reporting systems focus on how many horses get led to the water, and how difficult it was to get them there, but never quite get around to finding out whether the horses drank the water and stayed healthy.’

Source: Patton, M. 1997:157-8

A positive development during the last decade has been the movement to introduce standards and measures into programme planning, monitoring and assessments. This has helped to shift the focus of evaluations from programme outputs (such as the number of tons of food shipped, the number of shelter tarpaulins distributed, or the number of latrines constructed), to outcomes and impacts (such as number of deaths per day; percentage of children under 5 years of age with acute malnutrition; and incidence rates of measles and acute watery diarrhoea) (Bolton, P. et al. 2007). However, in cases where there is a strong evidence-base that an input leads to a humanitarian outcome e.g. therapeutic feeding or vaccinations, process indicators may be sufficient proxy indicators.

A key message that emerges from the literature and interviews is that considerations of impact and related indicators should be linked to each stage of the project cycle, from planning and gathering of baseline data, to monitoring and evaluation. Failure to do so is one of the main causes of the lack of useful impact assessment findings (C.A. Hofmann et al., 2004; Centre for Development, 2006; Watson, C. 2008).

Even if indicators are identified, and they have been integrated across the project cycle, unstable situations, limited access to affected populations of humanitarian and lack of methodological rigour and expertise may lead to timely and accurate data being unavailable. Furthermore, what is available is not always reliable (Hofmann, C.A. 2004; Spiegel, P. 2004). A previous ALNAP RHA highlighted that:

“Reports were so consistent in their criticism of agency monitoring and evaluation practices that a standard sentence could almost be inserted into all reports along the lines of: It was not possible to assess the impact of this intervention because of the lack of adequate indicators, clear objectives, baseline data and monitoring.” (ALNAP, 2003a, p107)

In the *ALNAP Annual Review 2003*, an ALNAP study on monitoring identified three important issues with regards to data focus and quality: that monitoring systems tend to focus on collecting data at the input and output levels; that the data collected is mainly quantitative data; and that data quality is often both poor, and poorly analysed.

In relation to the tendency to ignore the importance of qualitative data, a particular problem identified by the ALNAP study is that while, it is relatively easy to collect quantitative data and send it ‘up the line’, it is far more difficult both to define what qualitative data staff should collect on a regular basis, and to analyse such data when collected. Furthermore, the *ALNAP Annual Review 2003* highlights that the focus on quantitative monitoring is excluding a deeper understanding and hence decreasing opportunities for learning. As one interviewee in the ALNAP study of monitoring suggested: “...By providing masses of quantitative data in reports you could distract people from the key issue of whether funding for the reconstruction of housing was the right approach in the first place...” (ALNAP, 2003a). Importantly, collection and interpretation of qualitative data on an ongoing basis can be a key means of understanding causality and social process; something that is currently largely missing in evaluations of humanitarian action (see ALNAP, 2002:184).

The absence of baseline data is often used as an excuse for not assessing impact (Catley, A et al. cited in Watson, C. 2008). This has serious implications for humanitarian impact assessment. As Hofmann et al., describe,

“...Attempts to analyse the impact of humanitarian interventions are often handicapped by a lack of baseline data and knowledge about regular seasonal variations in key indicators. It is difficult to show that a humanitarian programme has had an impact without knowing the rate at which something was occurring before the intervention began, and after it was implemented. Likewise, when people are arriving in a new location or are returning home, it is often impossible to determine the baseline before their arrival... Programmes which keep mortality low or keep water and food provision high may be successful in meeting their objectives, but if they do not have a baseline

rate or a comparison group they may not be able to quantify the impact of the intervention...”

Yet, a number of methods exist for overcoming the challenge of obtaining baselines in emergencies including ‘retrospective’ baselines – which use beneficiary recall to reconstruct situations prior to emergency onset - or ‘rolling’ baselines’ – in which new beneficiaries of a programme provide background information which is compared to that of existing beneficiaries (Watson, 2008). It is also possible to construct a comparison group based on what is known about the project context and participant selection process (NORAD, 2008). Better use of national and local data sources, where relevant, also has potential to strengthen data quality.

Key questions around ‘Indicators, Baselines and Data for Humanitarian Impact Assessment’

- What kinds of indicators are needed for effective impact assessment and how can they be collectively agreed by different stakeholders with different needs and interests?
- What alternative approaches for establishing baselines are appropriate and relevant to the humanitarian context?
- How can good quality data be collected in highly fluid and unstable emergency contexts, wherever possible drawing on local and national data sources?
- What is needed to move beyond output-oriented monitoring towards monitoring outcomes and impacts?

7. Methodologies for humanitarian impact assessment

The key message for humanitarian agencies looking to identify appropriate methods is that any research design is shaped by both opportunities and constraints and must be viewed in the context of feasibility. It is therefore always useful to ask the question, *under which conditions are different approaches feasible and worthwhile?* In practice, impact assessment invariably involves selecting the best methods from the available tools based on sound analysis of the costs and benefits of each method. This is a far from straightforward process, and in many situations the best choice will be a mix of methods that can take into account, rather than dismiss, the complexity of assessing humanitarian impact (ESS, 2008).

The selection of methods will inevitably depend on what is being evaluated, who the key stakeholders are. There will always be trade-offs between precision and relevance; immediate and long-term learning; quantitative and qualitative data etc. The humanitarian sector needs to foster an appreciation of the strengths and weaknesses of these different options for particular research objectives. A key challenge is to select a research design and a range of tools and methodologies that are feasible, affordable and which fit the purpose of the impact assessment (Simanowitz, A. 2001).

There currently exist a wide variety of *quantitative* and *qualitative* tools and methodologies for assessing impact, drawn largely from the health sciences field and its epidemiological models, and from the development field and its program evaluation tools, and often divided between qualitative and quantitative. These include: documentary analysis; interviews; questionnaires (including recipient perceptions surveys); monitoring; ex-post evaluation; case studies; Participatory Rapid Appraisal (PRA) methods; experimental and quasi-experimental methods, including randomised controlled trials (RCTs).

Insights gained from interviews with key informants in the sector suggest that decisions about methodological approaches to impact assessment are highly ideological, and stem from a longstanding debate between two distinct research traditions - ‘logical positivists’ and ‘social constructivists’. This debate relates to different ways of knowing and understanding the world. While a comprehensive review of this debate is beyond the scope of this paper, the main points relating to impact assessment methodologies are summarised below:

Logical positivists seek to understand the social world by testing and falsifying hypotheses to uncover universal 'truths' about the material world, which can be observed, quantified and measured objectively (Prowse, M. 2007). From this perspective, quantitative methods such as surveys, experimental or quasi-experimental research designs, are best suited to assessing impact. Indeed, some proponents of these methods only define impact assessments as 'evidence-based' if they make explicit use of a counterfactual – i.e. what would have happened if a given intervention had not taken place - through experimental or quasi-experimental research designs (Forss, K. and S. Bandstein, 2008) (See Box 3).

Box 3: The Use of Randomised Controlled Trials to Assess Impact

One approach currently being strongly advocated as the only way to truly know about and attribute impact is that of randomised controlled trials (RCTs) – often referred to as a 'gold standard' in impact assessment. This method involves randomly choosing which individuals, families, or communities will be offered a program and which will not, similar to the way that the effect of a medical treatment is assessed in laboratories, by administering it to some members of a test group and not to others. Given a large enough sample size, RCTs ensure that impact measurements are not confounded by systematic differences between beneficiary and control groups (CGD, 2006; Nelson, J. 2008).

However, there is growing concern among members of the humanitarian community that assigning aid randomly is unethical, firstly because it contradicts principles about serving those in need, the most vulnerable and excluded, and secondly because it contradicts all we have learned about the importance of evaluation as a means of empowerment (Nelson, J. 2008). In response to this, proponents of RCTs argue that whenever funds are limited or programs need to be expanded in phases, only a portion of potential beneficiaries can be reached at any time. Choosing who initially participates by lottery is no less ethical (and perhaps even more so) than many other approaches (CGD, 2006; *ibid*). It has also been suggested that other methods do not produce knowledge that can be used to identify which interventions are more or less effective and why (Nelson, J. 2008).

Some have questioned the cost-efficiency of RCTs (Chapman, J. and A. Mancini, 2008; Roche, C. 2008 Forthcoming), others suggest they may only be suitable in contexts where a linear causal relationship can be established between the intervention and desired outcome that can be clearly defined; where it is possible to 'control' for context and other factors; when it can be anticipated that programmes under both experimental and control conditions will remain static for considerable period of time; and where it is possible and ethically appropriate to engage in randomisation (EES, 2008).

For more information on the use of RCTs in: <http://www.povertyactionlab.com/>

Social constructivists reject a positivist focus on the physical world and are sceptical of claims about universal 'truths'. For them, the social world is constituted by ideas, beliefs, languages and discourses and does not exist outside human consciousness. From this perspective, impact assessment is about the negotiation of different opinions and perspectives. Social constructivists are therefore more likely to employ qualitative, participatory methods to understand the opinions of different interest groups, for example, PRA techniques or the Most Significant Change Approach (Davies, R. 1996).

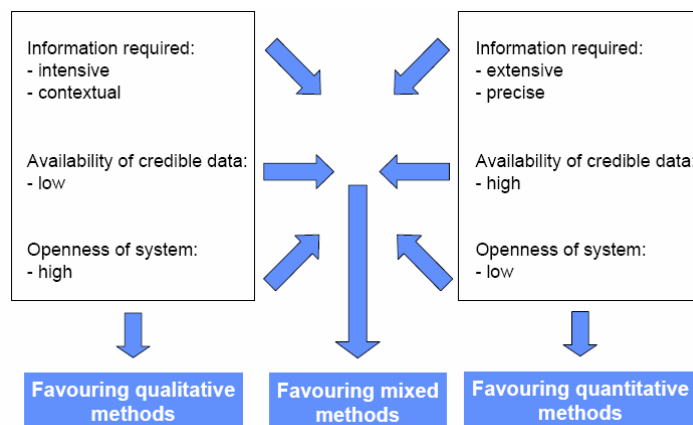
The positivist critique of social constructivist approaches and by association, qualitative methods, is that they lack objectivity and therefore policy-relevance. This is because they can lead to relativism - the notion that any model is as good as any other and there is no way to distinguish 'true' knowledge from inadequate knowledge (Jackson Chapter 11). From a social constructivist perspective, a positivist focus on what is measurable and quantifiable may make aggregation of information easier (Roche, C. 2008) and tell you about states and conditions (Ellis, F. 2000 cited in Prowse, M. 2007) but does not tell you very much about what those states and conditions mean, or how they came about.

Most agree that such an either/or position is not helpful, as both research traditions, and the methods they are most closely linked to, are suited to answering very *different types of research questions*. As Prowse notes, positivist methods such as RCTs are able to tackle 'what' and 'where' questions. Social constructivist approaches and qualitative methods, on the other hand, are better able to answer 'why' and 'how' questions, and are good at capturing processes (Prowse, M. 2007). A strict quantitative/qualitative divide therefore hides much more than it illuminates, and there is a very strong case for mixed methods approaches (Prowse, M. 2007), which can help to minimise tensions between quantitative and qualitative approaches, and compensate for weaknesses inherent in each approach.

It is worth noting that qualitative and quantitative methods compliment each other in a number of ways. One the one hand, qualitative methods can be used to develop quantitative measures and indicators or to explain quantitative findings. One the other, quantitative methods can be used to embellish a qualitative study or be used in parallel with qualitative data (Steckler, A. et al 1992).

One decision making framework which may be worth considering in the humanitarian sector is the contingency model. This suggests that different approaches are relevant in different situations, and that three important criteria for decision making are: the information required, the availability of credible data and the openness of the 'system' being evaluated (see Figure 3).

Figure 3: Contingency model for methodology selection



Source: William, K. 1999³

Key Questions around 'Methodologies for humanitarian impact assessment'

- How can agencies improve decision making around humanitarian impact assessment methodologies, taking account of strengths and weaknesses of different approaches?
- Could more joint work be done which looks at the relative value of different methodological approaches in different contexts?
- What mixed-method approaches are relevant and feasible in what situations?

³<http://www.cemcentre.org/Documents/CEM%20Extra/EBE/EBE1999/Kevin%20Williams.pdf>

8. Improved interpretation and analysis of data through engagement with affected populations and other stakeholders

Humanitarian interventions have historically been top-down in nature, hence participation by affected populations has not been a key feature of impact assessments. While many agencies speak of the importance of participation and incorporating beneficiary perspectives into evaluations and impact assessments, there remains a significant gap between rhetoric and practice. Insecurity in humanitarian contexts poses particular methodological constraints, which mean that women, older people and children are often not consulted (Roche, C. 1999). Participation is also time-consuming and often not required by donors who remain preoccupied by upward accountability. When participation does take place, it is usually informal and opportunistic, fuelling criticism that qualitative methods are 'unscientific'.

The work of the Emergency Capacity Building Project with the 'Good Enough Guide'⁴ is an important initiative to provide basic guidelines on how to be accountable to local people and measure programme impact in emergency situations (ECB Project, 2007). More recently, the Feinstein International Center has developed a detailed approach to carrying out project level Participatory Impact Assessments (PIA) of livelihoods interventions in the humanitarian sector. PIA came about partly as a response to the need to carry out impact assessment in complex emergency situations with difficult operational environments (Watson, C. 2008), and partly in response to the lack of international standards for measuring impact in areas other than water, health and nutrition (Catley, et al. 2008). PIA assists communities and NGOs to measure impact using their own indicators and their own methods. It can also overcome the weaknesses inherent in many donor and NGO monitoring and evaluation systems which emphasize the measurement of process and delivery, over results and impact (Catley, A. et al. 2008).⁵ There is also potential for national and local actors to play an active role in such collective processes. The idea of 'learning partnerships' for impact assessments is highlighted in Box 4.

Box 4: Learning Partnerships for Participatory Impact Assessments (PIA)

"...The experience of the Feinstein Center shows that where project participants are included in the impact assessment process, this can create an opportunity to develop a learning partnership involving the donor, the implementing partner, and the participating communities. The impact assessment process can create space for dialogue, and the results can provide a basis for discussions on how to improve programming and where best to allocate future resources.... a systematic, well designed PIA can assist communities and NGOs to measure impact using their own indicators and their own methods. It can also overcome the weaknesses inherent in many donor and NGO monitoring and evaluation systems which emphasize the measurement of process and delivery, over results and impact." (Catley et al, 2008, p8-10)

Many feel that impact assessment should not only be about providing more and better information, but also about making sure that the knowledge they generate is used in ways that improve the lives of affected populations. If impact is about beneficiaries, then the system needs to become more bottom-up. It has been suggested that one way to do this is through building partnerships between governments, NGOs and researchers in disaster-affected countries as well as affected populations (CGD, 2006; Catley, A. et al. 2008). While there is little evidence to suggest that disaster-affected populations currently constitute an external pressure or driving force behind impact assessments, the importance of making the voices and perspectives of disaster-affected populations central to impact studies is increasingly

⁴ The Good Enough Guide can be downloaded from:

http://www.oxfam.org.uk/what_we_do/resources/downloads/Good_Enough_Guide.pdf

⁵ For more information see:

<https://wikis.uit.tufts.edu/confluence/display/FIC/Participatory+Impact+Assessment+-+a+Guide+for+Practitioners>

being recognised. Initiatives such as the Fritz Institute's beneficiary perception surveys and the Listening Project are exploring these areas.

It is important to highlight the potential of impact assessments as a potential resource for affected populations. As Sandison observes, although beneficiaries are more likely now to be consulted during evaluations, it is not common practice to communicate the findings afterwards (Sandison, P., 2006). While there are numerous reasons why it may not be feasible or desirable to provide beneficiaries in disaster affected communities with the results of evaluations, the marked absence of accountability 'downwards' deserves further attention (ibid). However, this would require affected populations to be involved throughout the impact assessment process, from definition to generating indicators, baselines to playing an active role in judging impacts.

Key Questions around 'Improved interpretation and analysis of data through engagement with affected populations and other stakeholders'

- How to meaningfully engage affected populations in the process of humanitarian impact assessments?
- How to establish effective learning partnerships with affected populations, donors, implementing partners and national / local actors?
- How can impact assessments improve accountability to and information for affected populations?

9. Capacities and Incentives for improved humanitarian impact assessment

Many commentators have noted a lack of organisational capacity within the sector to carry out good evaluations and impact assessments, despite a wealth of tools and methods (Hofmann, C.A. et al. 2004; ALNAP, 2006; Watson, C. 2008). There is a wealth of evidence to suggest that evaluation TORs are often unclear; objectives are not defined clearly within the context of the intervention; stakeholder analysis is limited; timing relates to institutional priorities rather than humanitarian need; and that skills relating to impact assessment methodologies, in particular experimental and quasi-experimental design are lacking (CGD, 2006; Forss, K. and S. Bandstein 2008; ALNAP 2003b, 2004). Capacities are even more limited in developing country contexts (Hofmann, C.A. 2004).

The high staff turnover and the lack of a learning culture and adequate resources are all contributing factors (Roche, C. 2008; Watson, C. 2008) to the lack of capacity. In their report, "When will we ever learn?", the Center for Global Development suggest that part of the difficulty is that too many different tasks are implicitly simultaneously assigned to evaluation. These include building knowledge on processes and situations in recipient countries, promoting and monitoring quality, informing judgment on performance, and measuring actual impacts (CGD, 2006). While these different dimensions of evaluations are complementary, there are strong arguments to suggest that, for effectiveness and efficiency reasons, they should be carefully identified and organised into separate 'systems' demanding specific methodologies and capacities.

One area that has to date been insufficiently explored is the potential for sector-wide initiatives to strengthen capacity for humanitarian impact assessments, or for collective within specific crisis contexts in order to undertake impact assessments. There have been system-wide initiatives focusing on the development sector, such as NONIE and 3IE, but no such mechanism has developed in the humanitarian sector. ALNAP may be well placed to convene such a mechanism, should it be seen as necessary and relevant.

The capacity issue is crucially linked to incentives within aid agencies. There are a number of ways in which incentives within humanitarian agencies and, in some cases, the sector as a whole, can promote or hinder the development of effective impact assessment (Watson, C. 2008). In the extreme, institutional incentives can override humanitarian ones, with decision-

making shaped more by organisational priorities than by humanitarian outcomes or working effectively (Roche, C. 2008).

The literature suggests that an evaluation gap exists because there are **too few incentives** to conduct good impact assessments (CGD, 2006; Roche, C. 2008; Watson, C. 2008). According to the Center for Global Development, donors do not prioritise understanding impact and hence do not put pressure on implementing agencies to carry out good impact assessments (2006). Furthermore, impact evaluation has immediate costs while the benefits are felt only well into the future. Good quality impact assessments can also be very expensive, while budgets for evaluation are usually far less than 1 per cent of overall budgets and the option that they could increase radically does not seem likely (Forss, K. and S. Bandstein, 2008; Chapman, J. and A. Mancini, 2008). As a public good, the benefits of impact assessment findings have the potential go well beyond the organisation in which they are generated. However, the incentives for any organisation to bear the costs are much lower than the full collective and social benefits would justify (ibid).

What's more, many have noted that a results-based approach to impact can create **perverse incentives**, which serve to discourage individuals and organisations from admitting error, undermine attempts to innovate or build on local needs and contexts (Chapman, J. and A. Mancini 2008). In some cases, these incentives can lead agencies to downplay the role of others in achieving results (Roche, C. 2008).

Since impact evaluations can go either way – demonstrating positive, zero or negative impact – they can work to the detriment of learning and transparency. Staff and organisations are more likely to keep quiet about the challenges they face and work that is not going well (Chapman, J. and A. Mancini 2008). Moreover, policymakers and managers have more discretion to pick and choose strategic directions when less is known about what does or does not work (CGD, 2006). A recent scoping study, which looked at improving impact evaluation coordination and uptake, found that only a limited number of impact evaluation findings - in both development and humanitarian sector - are published, and tend to be biased towards those which contain favourable results (Jones, N. et al. 2008).

A number of humanitarian actors are concerned that results-based management approaches to impact assessment, which favour log frames and quantifiable results, can undermine a central element of the humanitarian ethic: a desire to protect and demonstrate solidarity with victims, and to restore their dignity (Barnett, M. 2005; Robertson, D. et al. 2002). Barnett asks, *can such non-quantifiable values as compassion and caring for others be operationalised within such a model?* Others noted that advocacy work also falls into this category of humanitarian work that is hard to measure (Hofmann, C.A. 2004).⁶ Focusing on what is measurable, they argue, risks reducing humanitarian aid to a technical question of delivery, rather than a principled endeavour in which the process, as well as the outcome, is important (ibid). RBM can also lead to a focus on outputs (as opposed to outcomes or impact), as these are easier to attribute to particular interventions.

A number of cultural barriers and biases that hinder good quality humanitarian impact assessment have also been identified including: the tendency to value action over analyses (ALNAP, 2002); and the tendency to maintain previously held beliefs and neglect evidence that might conflict with those beliefs (Roche, C. 2008); and increasing aversion to risk, reflecting donor cultures and growing competition for resources (Watson, C. 2008).

Key questions around 'Capacities and Incentives for humanitarian impact assessment'

- Can humanitarian agencies work to maximise capacities and resources for and benefits of impact assessment?
- Is there potential for a cross-sector initiative for strengthening capacities for impact assessments and to facilitate joint impact assessments?

⁶ For a wealth of resources on measuring the impact of advocacy see http://www.innonet.org/?section_id=3&content_id=601.

- How can incentives be adapted so that organisational and humanitarian priorities are better aligned and impact assessments are positioned to help address both sets of priorities?
- How can organisations support their staff to reflect critically on their work, including learning from things that did not go well?

10. Summary and Key Questions

“...Taken as a whole, the humanitarian system has been poor at measuring or analysing impact, and the introduction of results-based management systems in headquarters has yet to feed through into improved analysis of impact in the field. Yet the tools exist: the problem therefore seems to be that the system currently does not have the skills and the capacity to use them fully. This suggests that, if donors and agencies alike want to be able to demonstrate impact more robustly, there is a need for greater investment in the skills and capacities needed to do this. Given the large (and rising) expenditures on humanitarian assistance, it is arguable that there has been significant under-investment in evaluation and impact analysis. [Improvements in impact assessment] would have wider benefits beyond simply the practice of impact assessment: greater emphasis on the participation of the affected population, the need for clearer objectives for humanitarian aid, more robust assessments of risk and need and more research into what works and what does not would be to the advantage of the system as a whole...” (Hoffman C.A. et al, 2004)

Our review gives little indication that there has been much movement from the position above, articulated in the HPG Report of 2004. It is clear that the humanitarian sector faces a number of challenges in relation to impact assessment. Specifically, the six areas outlined here are:

- Defining humanitarian impact assessment
- Stakeholder interests in and needs for humanitarian impact assessment
- Indicators, Baselines and Data for Humanitarian Impact Assessment
- Methodologies for humanitarian impact assessment
- Improved interpretation and analysis of data through engagement with affected populations and other stakeholders
- Capacities and Incentives for improved humanitarian impact assessment

These areas and related questions which are highlighted throughout the paper provide the basis of further discussion and debate at the ALNAP Biannual, as well as for future investigations and research into humanitarian impact assessments.

By working together to think through and address these and other issues raised at the Biannual, it is hoped that the ALNAP membership and wider humanitarian sector can establish a **realistic vision** for humanitarian impact assessment, and a series of **practical steps** to achieve this vision. The focus should be on how the sector and agencies within it can move towards a **more systematic, considered and collaborative approach to analysing intended and unintended, positive and negative humanitarian impacts**. It is to be hoped that in this important area, the sector will see more change in the next four years than it has done in the last four.

This paper was written by Karen Proudlock and Ben Ramalingam, and is an initial output of an ongoing ALNAP Study on Humanitarian Impact Assessment which will be published as part of the 8th Review of Humanitarian Action in 2009.

Bibliography:

- ALNAP (2001) *Humanitarian Action: Learning from Evaluation*. ALNAP Annual Review.
- ALNAP (2002) *Humanitarian Action: Improving Performance Through Improved Learning*. ALNAP Annual Review.
- ALNAP (2003a) *Humanitarian Action: Improving Monitoring to Enhance Accountability and Learning*. ALNAP Annual Review.
- ALNAP (2003b) *Meta-evaluation*. In: Review of Humanitarian Action.
- ALNAP (2004) *Meta-evaluation*. In: Review of Humanitarian Action.
- ALNAP (forthcoming 2009) *The Humanitarian Performance Project: Concepts, Contexts and Challenges*.
- Anderson, M. (1996) *Do No Harm: Supporting Local Capacities for Peace through Aid*. Cambridge, MA: The Collaborative for Development Action, Inc.
- Barnett, M. (2005) *Humanitarianism Transformed*. University of Minnesota.
- Buchanan-Smith, M. (2003) *How the Sphere Project Came into Being: A Case Study of Policy Making in the Humanitarian Aid Sector and the Relative Influence of Research*. Working Paper 215, London, ODI.
- Catley, A. et al. (2008) *Participatory Impact Assessment: A Guide for Practitioners*. Feinstein International Center, Tufts University.
- Center for Global Development (2006) *When Will We Ever Learn? Improving Lives through Impact Evaluation*. Report of the Evaluation Gap Working Group.
- Chapman, J and A. Mancini (2008) *Impact Assessment: Drivers, Dilemmas and Deliberations*. Prepared for Sightsavers International.
- Clarke, P and B. Ramalingam (2008) *Organisational Change in the Humanitarian Sector*. In: *ALNAP 7th Review of Humanitarian Action* ODI, London.
- Collinson, S. (2002) *Power, livelihoods and conflict: case studies in political economy analysis for humanitarian action*. HPG Report 13, London, ODI.
- Crisp, J. (2004) *Thinking Outside the Box: evaluation and humanitarian action*. In: Forced Migration Review 8
- Davies, Richard (1996), *An Evolutionary Approach to Facilitating Organisational Learning: An Experiment by the Christian Commission for Development in Bangladesh*. Centre for Development Studies, Swansea, UK
- de Waal, Alex. 1997. *Famine crimes: Politics and the disaster relief Industry in Africa*. Bloomington: Indiana University Press.
- Donini, A. et al. 2008 *The State of the Humanitarian Enterprise*. Humanitarian Agenda 2015: Final Report, Feinstein International Center, Tufts University.
- ECB Project (2007) *Impact Measurement and Accountability in Emergencies: The Good Enough Guide*.

Ellis, F. (2000) *Rural Livelihoods and Diversity in Developing Countries*. Oxford: Oxford University Press, cited in Prowse, M. (2007) *Aid effectiveness: the role of qualitative research in impact evaluation*. ODI Background Note. ODI, London.

European Evaluation Society (EES) (2008) *ESS Statement: The Importance of a methodologically diverse approach to impact evaluation – specifically with respect to development aid and development interventions*. EES December 2007

Forss, K. and S. Bandstein (2008) *Evidence-based Evaluation of Development Cooperation: Possible? Feasible? Desirable?* NONIE Working Paper No. 8

Fritz Institute (2005) *Recipient Perceptions of Aid Effectiveness: Rescue, Relief and Rehabilitation in Tsunami Affected Indonesia, India and Sri Lanka* San Francisco: Fritz Institute

Fritz Institute (2006) *Hurricane Katrina: Perceptions of the Affected* San Francisco: Fritz Institute

Harrell-Bond, B. *Imposing Aid: Emergency Assistance to Refugees* Oxford University press, Oxford. pp xi-xii

Hofmann, C.A. et al. (2004) *Measuring the Impact of Humanitarian Aid*. HPG Report 17. ODI, London.

Hospes, O. (2008) *Evaluation Evolution? Three Approaches to Evaluation*. In: The Broker Issue 8, June 2008.

Jones, N. et al. (2008) *Improving Impact Evaluation Coordination and Uptake*. A Scoping Study commissioned by the DFID Evaluation Department on behalf of the NONIE.

Kaiser, T. (2004) *Promise and practice: participatory evaluation of humanitarian assistance*. Forced Migration Review 8

Macrae, J. (2002) *The New Humanitarianisms: A Review of Trends in Global Humanitarian Action* HPG Report 11. ODI, London.

Patton, M. (1997) *Utilisation-Focused Evaluation: the new century text*. Sage.

NORAD (2008) *The Challenge of Assessing Aid Impact: A Review of Norwegian Evaluation Practice* Study 1/2008. Norwegian Agency for Development Cooperation. Oslo, Norway.

OECD/DAC (2002) *Glossary of Key Terms in Evaluation and Results Based Management*, DAC Working Party on Aid Evaluation.

Pawson, R. (2003) *Nothing as practical as a good theory*. *Evaluation*, 9:471-490.

Prowse, M. (2007) *Aid effectiveness: the role of qualitative research in impact evaluation*. ODI Background Note. ODI, London.

Prowse, M. (2008) *Impact Evaluation and Interventions to address Climate Change*. ODI, London.

Robertson, D. et al. (2002) *What kind of evidence do we need to justify humanitarian medical aid?* in *The Lancet* Vol. 360: 330-333.

Roche, Chris (1999) *Impact Assessment for Development Agencies*, Oxfam GB, Oxford.

Roche, C. (2000) *Impact assessment: seeing the wood and the trees*. in *Development in Practice*, Volume 10, Numbers 3 & 4, August 2000.

Roche, C. (2008) *The Seeming Simplicity of Measurement, Paper for the Ethical Questions for Non-Governmental Organisations*. Workshop, 18-20 July, Melbourne, Chris Roche, Oxfam Australia

Sandison, P. (2006) *Evaluation Utilisation*. Ch.3 in ALNAP (2006) Review of Humanitarian Action.

Slim, H. (1997) *Doing the Right Thing*. Disasters, vol. 21/3, September

Slim, H. (2006) *Global welfare: A realistic expectation for the international humanitarian system?* ALNAP 6th Review of Humanitarian Action ODI, London.

Simanowitz, A. (2001) *From Event to Process: Current Trends in Microfinance Impact Assessment* Imp-act Guidelines No. 5. The Institute of Development Studies, University of Sussex

Smutylo, T. (2001) *Crouching Impact, Hidden Attribution: Overcoming Threats to Learning in Development Programs* Evaluation Unit, International Development Research Centre, Ottawa, Canada. Draft Learning Methodology Paper, Block Island Workshop on Across Portfolio Learning 22-24 May 2001

SPHERE (2004) *Humanitarian Charter and Minimum Standards in Disaster Response*.

Steckler, A. et al. (1992) *Toward Integrating Quantitative and Qualitative Methods: An Introduction* in Health Education Quarterly Vol. 19 (1):1-8.

Teller, C.H. (2008) *Are We Learning About What Really Works? Lost Opportunities and Constraints to Producing Rigorous Evaluation Designs of Health Project Impact* NONIE WORKING PAPER NO. 9

Watson, C. (2008) *Literature Review of Impact Measurement in the Humanitarian Assistance Sector*. Paper submitted to the Feinstein International Center, Tufts University for the Bill and Melinda Gates Foundation Project: Impact Assessment of Innovative Humanitarian Projects in Sub-Saharan Africa.

White, H. 2007 *Evaluating Aid Impact* MPRA Paper No. 6716, posted 13. January 2008 / 01:34 Institute of Development Studies, University of Sussex

Williams, K (1999) *Mixing qualitative and quantitative evaluation tools: a pragmatic approach*. Centre for European Evaluation Expertise (C3E).