Towards enhancing responsibility and accountability in humanitarian action:

Understanding the subjective factors that influence evaluation of humanitarian actions and the implementation of the recommendations made.

Anthony Christopher Land

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy

Liverpool School of Tropical Medicine

March 2014

LSTM

LIVERPOOL SCHOOL OF TROPICAL MEDICINE

<u>Declaration</u>

The material presented in this thesis is my own work and has not been presented, nor is it currently being presented, whether in part or wholly as part of any other degree or another qualification.

<u>Dedication</u>

This work is dedicated to the survivors of natural disasters, conflict or both, who strive to rebuild their shattered lives

and

To my colleagues and friends, the humanitarian practitioners, who seek to do their best, for those to whom the world has done its worst.

To those humanitarians who, too often, have to accept that they can only: *"bring a measure of humanity, always insufficient into situations that should not exist". (Rieff 2003, p. 19)*

To those humanitarians who, in the modern world, often risk injury and death to fulfil their task and in particular my colleagues who have given their lives while delivering humanitarian service to others.

<u>Acknowledgements</u>

__Abstract__

When specialists in the evaluation of humanitarian action meet they often complain that *"evaluations tell us nothing new"*. Can this complaint be justified and, if so, can any additional insights into the reasons be discovered?

An analysis and comparison is made of the recommendations arising from the evaluations of the humanitarian response to two major forced displacements and two natural disasters. The comparison is used to identify the extent to which recommendations made in the evaluation of the earlier of each pair were repeated in the subsequent evaluations.

An analysis of the subjective influences reported as impinging upon nine of the earliest evaluations of humanitarian actions is made. A series of 22 elite interviews with staff of humanitarian organisations and independent consultants directly involved in the evaluation of humanitarian action reveals the continuing influence of these 'subjective' factors and indicates of a lack of confidence in the evaluation process as a tool for institutional learning.

The roles of agency and structure in the subjective influences impinging on the evaluation of humanitarian action are analysed and recommendations made for the improvement of the evaluation/learning cycle.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

4Rs	Repatriation, Reintegration, Rehabilitation and Reconstruction
AFP	Agence France-Presse
ALNAP	Active Learning Network for Accountability and Performance
CAP	Consolidated Appeal Process
CAQDAS	Computer Assisted Analysis of Qualitative Data
CAR	Central African Republic
CEO	Chief Executive Officer
CERF	Central Emergency Response Fund
CERP	Commander's Emergency Response Program
CIMIC	Civil Military Cooperation
CNAR	Commission Nationale pour l'Accueil et Réinsertion des Réfugiés
CNN	Cable News Network
CRS	Catholic Relief Services
CSO	Civil Society Organisations
DAC	Development Assistance Committee (of the OEDC)
DFID	Department for International Development (of the UK Government)
DHA	Department of Humanitarian Affairs
DRC	Democratic Republic of the Congo
DRR	Disaster Risk Reduction
ECHO	Humanitarian Aid and Civil Protection Department of the European Commission
ELRHA	Enhancing Learning and Research for Humanitarian Assistance
ERC	Emergency Relief Coordinator
ERT	Emergency Response Team
EU	European Union
EWR	Early Warning and Response
FAO	Food and Agriculture Organisation
FDF	Fondation de France
FTS	Financial Tracking System
GD	Government Donor
GDP	Gross Domestic Product
GHD	Good Humanitarian Donorship
GIS	Geographic Information System
GWOT	Global War on Terror
HAP	Humanitarian Accountability Partnership
HR	Human Resource
HRI	Humanitarian Response Index
IASC	Interagency Standing Committee
IC	Independent Consultant
ICISS	International Commission on Intervention and State Sovereignty
ICRC	International Committee of the Red Cross
IDP	Internally Displaced Person
IFRC	International Federation of the Red Cross
IO	International Organisation
IOB	Operations Evaluation Department
IP	Implementing Partner

JEEAR	Joint Evaluation of Emergency Assistance to Rwanda
JSI	Joint Standards Initiative
LRRD	Linking Relief, Rehabilitation and Development
MINUSTAH	UN Stabilization Mission in Haiti
MOU	Memorandum of Understanding
MSF	Médecins Sans Frontières
NATO	North Atlantic Treaty Organisation
NGO	Non Governmental Organisation
OCHA	Office for the Coordination of Humanitarian Affairs
OEF	Operation Enduring Freedom
OECD	Organisation for Economic Cooperation and Development
OFDA	Office for Foreign Disaster Assistance
R2P	Responsibility to Protect
RPA	Rwandese Patriotic Army
RPF	Rwandese Patriotic Front
SCF	Save the Children Fund
SFP	Supplementary Feeding Programme
SGBV	Sexual and Gender Based Violence
Sida	Swedish International Development Agency
TEC	Tsunami Evaluation Coalition
TOR	Terms of Reference
TRCS	Tajikistan Red Crescent Society
UK	United Kingdom (of Great Britain and Northern Ireland)
UN	United Nations
UNDRO	United Nations Disaster Relief Coordinator
UNHCR	United Nations High Commissioner for Refugees (UN Refugee Agency)
UNICEF	United Nations Children's Fund
UNOCHA	See OCHA
UNREO	United Nations Rwanda Emergency Office
UNRRA	United nations Relief and Rehabilitation Administration
UNV	United Nations Volunteer
US or USA	United States of America
USSR	Union of Soviet Socialist Republics
WFP	World Food Programme

x

Chapter 1 – Introduction

Origins of the research

This research has its origins in the perception widely held amongst practitioners of humanitarian action, reported by Ramalingam, Scriven and Foley (2009, p. 2 & 9) *"that evaluations do not tell us anything new"*. This reflects a sense of frustration amongst humanitarian practitioners founded on a belief that, despite the proliferation of evaluations of humanitarian projects and programmes since the 1990s, recommendations made are not resulting in a learning process and the improvement of the quality of delivery of humanitarian action. In acute emergencies where the quality of delivery of protection and assistance may reasonably be assumed to translate directly into relief of human suffering and the saving of lives, it is a serious criticism that humanitarians continue to make the same mistakes despite these being clearly identified and recommendations made for improvement. If evidence can be found to show that this perception reflects the real situation, this indicates a major problem with the evaluation/learning process. A similar level of criticism and recommendations regarding a medical procedure or the design of an automobile which were ignored for a period of years would cause outrage and would demand political action.

Introducing the researcher[1]

This research is being undertaken following a 30 year career in the humanitarian sector working in NGOs, the United Nations and in academic teaching roles. Most of the researcher's career has been spent working in large refugee emergencies, as Head of UNHCR Offices and Director of Operations. The researcher also worked for seven years in Donor Relations and Resource Mobilisation. When the researcher retired from UNHCR in 2006 it was with many questions regarding the theory, policy and practice of humanitarian action in which he had been involved. The opportunity to teach post graduate courses in humanitarian subjects immediately following retirement has provided a fertile environment for analysing these questions and

---

[1] Throughout this work the author of this thesis will refer to himself as "the researcher" so as to avoid any confusion when referring to the author of a work being cited. The only exception to this is in the attribution of original tables and figures where "Source: Author" is used.

reading extensively what has been written by practitioners and academics. The varying points of view expressed, some of which did not coincide with the researchers experience have often provided challenges to understanding rather than answers. This research looks at one of these unanswered questions.

In this work the researcher utilises his experience to conduct insider research into some of the causes of the problem identified by Ramalingam, Scriven and Foley (2009) which have resulted in considerable frustration throughout the humanitarian profession. From the researcher's personal point of view an important starting point has been the question that, while much has been written on the objective policy and methodology of evaluation, are more subjective influences a significant part of the problem?

Introducing the research

This research establishes an evidence base for the repetition of similar recommendations in evaluations of humanitarian action and identifies reasons for this repetition. The research then analyses the subjective influences reported by evaluators who undertook some of the earliest evaluations of humanitarian action on the 1990s and, through interviews with current practitioners, establishes the extent to which the same influences continue to be experienced. The role of the structure, in which the evaluation of humanitarian action takes place, in perpetuating the subjective influences is analysed and recommendations for the improvement of the evaluation/learning cycle are made.

The researcher's initial reaction to the perception reported by Ramalingam, Scriven and Foley (2009)was to find himself holding two almost opposite responses. The researcher has read many evaluations in the course of his career with a sense of déjà vu, a feeling that the topics covered, the criticisms and recommendations made were familiar and felt unchanging. At the same time from the researcher's experience it was clear that the institutions and mechanisms for responding to humanitarian needs had made significant advances over the previous 30 years. These include the professionalization of organisations and their staff, improved technologies for logistics, financial accounting and communication, the proliferation of guidelines, handbooks and tool kits, the development of coordination mechanisms, the espousal of evaluation as a tool for accountability and learning.

Standard setting organisations such as The Sphere Project, People in Aid (PIA) and the Humanitarian Accountability Partnership (HAP) have all been established. The earliest evaluations of humanitarian action were undertaken in 1993 when there were doubts *"whether humanitarian assistance could be evaluated"* (O'Keefe et al. 2001, p. 19) and it seems impossible that in 20 years evaluations have not brought improvements of the humanitarian system or that evaluation has reached the end of its useful life and become merely repetitive.

 Ramalingam, Scriven and Foley (2009) used the perception *"that evaluations do not tell us anything new"* as a springboard to emphasise the need for a more innovative approach to humanitarian action rather than a reliance on the iterative, project, evaluation and learning cycle. The role of evaluation is not the main thrust of their article. However, the researcher found their reporting of the perception *"that evaluations do not tell us anything new"* to indicate a widely held subjective reaction to the important tool of evaluation which undermined its credibility. In this context a better understanding of the extent to which *"evaluations do not tell us anything new"* is important and became the subject of the first part of this research.

From the beginning of this research it appeared that when researching a subjective perception that was reported as widely held amongst *"critics of humanitarian aid, many from within the sector"* (Ramalingam, Scriven & Foley 2009, p. 2) and was *"one of the most frequent comments heard at ALNAP meetings"* (p. 9) it was unlikely that a clear "true" or "false" result would be obtained. Even if the perception was partially true then it would be important to identify some possible causes and the extent to which the problems were in the evaluation or the learning parts of the cycle.

Starting the research from a subjective perception regarding evaluation reported by Ramalingam, Scriven and Foley (2009) the researcher decided to identify what other subjective influences were perceived as existing in the evaluation of humanitarian action and how these might influence the evaluation learning cycle. In the relevant literature little attention appears to have been paid to the on-going effects of the subjectivity inherent in the stressful world of humanitarian action. Walkup (1997, pp. 37-60) identified that the stressful environment in which humanitarians operate leads to a high level of subjectivity, including delusion, defensiveness, myths and coping strategies regarding the quality of delivery of humanitarian action.

In order to begin this research it was important for the researcher to define some criteria for what constituted a subjective influence. Through the process of literature review, interviews and analysis the definition evolved, becoming more inclusive:

1. Influences not under the control of the policy or methodology of evaluation.

2. Influences perceived as limiting the objectivity of the evaluation learning process.

3. Influences on evaluation and learning perceived by participants as subjective.

Background to the research

In introducing this research it is important to clarify a number of issues that identify both the scope and rationale behind the work. These issues are:

1. How, for the purposes of this research, is humanitarian action defined and how is it clearly separated from other areas of aid?

2. What is the historical perspective of evaluation? How and when was it introduced to the sphere of humanitarian action?

3. Who are the important actors in humanitarian coordination, evaluation and learning?

1. <u>How, for the purposes of this research, is humanitarian action defined and how is it clearly separated from other areas of aid?</u>

Humanitarian action is one specialised form of aid, a broad topic that spans a wide spectrum from assistance given at the national level to governance, infrastructure, and economic development to emergency assistance given at the individual level to those who have lost the capacity for continued survival. It is in the latter terms that the scope of this research is based, with the term "action" being used to indicate the inclusion of both assistance and protection. It must, however, be recognised that humanitarian action has a larger, and growing, definition and the term is used for a variety of different approaches. The so called "new humanitarianism" puts humanitarian action into a much larger context. Fox (2002) summarising Mikael Barford's presentation to the UK International Development Select Committee on Conflict and Post Conflict Reconstruction in 1999 writes,

> *"what characterises new humanitarianism is: the integration of human rights and peace building into the humanitarian orbit; the ending of the distinction between development and humanitarian relief; and the rejection of the principle of neutrality."* Fox (2002, p. 276)

It must be concluded though, that however large the definition of humanitarianism becomes it cannot bypass or remove the most basic component, the need to step in and provide survival assistance following a disaster. The researcher will use the definition that he developed during his 30 years of experience in the field and continues to use in his teaching. Humanitarian actions are: *"Actions to ameliorate human suffering in a situation which follows a gross threat to human survival (Humanitarian Emergency)."* This definition focusses on a single objective, the amelioration of human suffering. While the immediate necessity is the survival of those affected it does, however, cover actions beyond and rapidly carries over into what has been termed "care and maintenance" i.e. the on-going provision of shelter, food, water, sanitation, hygiene as well as services such as medical care and education. As survivors, particularly those who have been displaced from their region of origin may remain in need of humanitarian assistance for considerable periods, aspects such as quality of life and livelihoods also need to be addressed. Whereas these activities may be similar activities undertaken from a development perspective, and indeed may overlap with such activities, humanitarian actions usually consider sustainability to be less important than immediate impact and do not consider the economic rate of return as a criterion. Funding for humanitarian action is generally considered to be expenditure rather than an investment.

2. <u>What is the historical perspective of evaluation? How and when was it introduced to the sphere of humanitarian action?</u>

Scriven (1996) writes that *"evaluation is a very young discipline – although it is a very old practice"* (pp. 394-395) and its long history from 1792 has been documented in some detail by Hogan (2007) in seven ages

> "*the first the period prior to 1900 the Age of Reform; second, from 1900 until 1930, the Age of Efficiency; third, from 1930 to 1945, called the Tylerian Age; fourth, from 1946 to about 1957, called the Age of Innocence; fifth, from 1958 to 1972, the Age of Development; sixth, from 1973 to 1983, the Age of*

Professionalization; and seventh, from 1983 to 2000 the Age of Expansion and Integration." (p. 6)

As will be seen from the discussion of the history of evaluation in humanitarian action below it is only the last of these that concern this research although the evaluation of other areas of public funded activity began earlier. It was during the Age of Expansion and Integration that *"Professional associations were developed along with evaluation standards. In addition, the Joint Committee on Standards for Educational Evaluation developed criteria for personnel evaluation."* (Hogan 2007, p. 6)

In their web page summary of the history of evaluation Evalsed (2009) identify that *"Evaluation emerged as a distinct area of professional practice in the post-war years in North America."* Three sources were important in "*this early period*", evaluation of educational innovations, linking evaluation with resource allocation and the evaluation of anti-poverty programmes. Evalsed (2009) points out that these roots initiated some major components of evaluation methodology at least two of which *"cost benefit and economic appraisal methods; and participatory and qualitative indicator methods involving the intended beneficiaries of programmes in the evaluation process"* feature strongly in current evaluation practices applied to humanitarian actions. These two approaches link closely to the factors of "accountability vs. lessons learned" discussed later in this research.

Stemming from this historical background Evalsed (2009) also identified four groups who all have an interest in evaluation but whose interests may be in competition; groups that are all reflected in the field of humanitarian action.

- Policy makers – particularly elected officials responsible for government donors interested in accountability for public money spent on humanitarian projects and programmes and *"justification for policy decisions".*

- Professional and specialist interests – humanitarian practitioners technical specialists interested "*in opportunities to improve the quality of their work"*

- Managers and administrators – *"often concerned with the delivery of policies and programmes how well they are managed and organised"*

- Beneficiaries – Interested in shaping programmes better to "*meet their needs".*

6

Coming from public service roots and with the similarity of interest groups it is hardly surprising that evaluation has now been applied systematically to humanitarian endeavour.

Evalsed (2009) identifies the spread of evaluation into *"Northern Europe and in those parts of Europe, in particular, that had close links with the United States and Canada"* and further into other parts of Europe in the 1970s.

The influence of politics on the wide adoption of evaluation is also noted citing the requirement for budgetary reform in France in 2000 and the change of government in Britain in 1997 as significant events, as well as the 1998 Structural Funds Regulation of the European Commission. The appointment of Mukesh Kapila as head of Conflict and Humanitarian Affairs in DFID by Clair Short in 1997, which followed closely on the publication of the Joint Evaluation of Emergency Assistance to Rwanda (Borton, Brusset & Hallam 1996) proved a major impetus to the widespread adoption of evaluation by British humanitarian organisations.

The evaluation of humanitarian action was based on the Principles for Evaluation of Development Assistance (Development Assistance Committee 1991). Although the application of evaluation to humanitarian action can be documented back to 1993 (Wood, Apthorpe & Borton 2001c, p. 10) guidelines for the evaluation of the complex emergencies of the type that evolved in the 1990s were not published until 1999 (Development Assistance Committee 1999).

It is interesting to note that both these documents (Development Assistance Committee 1991) and (Development Assistance Committee 1999) indicate a very similar two fold purpose for evaluation

- *"to improve future aid policy, programmes and projects through feedback of lessons learned;*

- *to provide a basis for accountability, including the provision of information to the public."* (Development Assistance Committee 1999, p. 5)

Perceptions of the problems caused when these two purposes are applied simultaneously in the evaluation of humanitarian action will be discussed later in this research.

Wood, Apthorpe and Borton (2001c) observe that despite the Development Assistance Committee (DAC) Expert Group on Aid Evaluation (subsequently to become the DAC Working Party on Aid Evaluation) being formed in 1982; *"the application of evaluation to the field of humanitarian aid (or emergency assistance or disaster relief as it was known) was comparatively slow"* and the cause for this is identified as *"a combination of attitudinal, technical, practical and methodological factors"* which included:

- Resentment of aid organisations based on the attitude that *"we did our best under extremely difficult circumstances and do not accept that someone who was not involved in the operation should come and criticise us"* (p11)

- Rapid change and lack of baseline data in emergency situations making them quite different to the development situations to which evaluation had already been applied. (p. 11)

- Pressure to respond rapidly leading to poor documentation of the response. (p. 11)

- The multidisciplinary nature of humanitarian response creating "*organisational barriers to initiating and undertaking evaluations*". (p. 11)

Wood, Apthorpe and Borton (2001c) suggest that these factors have been addressed and overcome (p. 11) but this research indicates that far from being overcome they have been largely ignored and continue to contribute to subjective factors that influence the evaluation process. The same document points to the major change agent being the *"pressure fuelled by the increasing level of resources used by the international humanitarian system and also wider changes in the approach to accountability within Western societies"* (p. 11) (particularly government funding agencies), that overrode the objections listed above and enforced evaluation practices already used in development projects onto humanitarian action.

Wood, Apthorpe and Borton (2001c, pp. 11-12) cite some US studies made in the 1970s into the response to the Sahelian droughts and a book, by Shawcross (1984), as documents that prompted a more robust approach to the evaluation of humanitarian action but trace the first application of "*evaluation procedures, as applied to development assistance"* to emergency assistance occurring in 1993.

The discussion on whether it was indeed possible to evaluate humanitarian action steadily lost ground to increasing donor government pressure for evaluations through the early 1990s and the *"first attempt to evaluate humanitarian assistance provided in response to a conflict or complex emergency"*` was the evaluation undertaken in 1993 by the Netherlands Government into assistance they had funded in Somalia. (Wood, Apthorpe & Borton 2001c, p. 12)

In their account of undertaking their trail blazing evaluation, O'Keefe et al. (2001) cite the 1990s as being a time of massive growth in humanitarian expenditure due to *"the collapse of the Soviet Union and the end of a bipolar world. Cambodia, Afghanistan, the Caucasus, Iraq, former Yugoslavia, the Horn countries and Nicaragua were trouble spots. Bosnia dominated and there especially, the issues of humanitarian intervention were tied to diplomatic and military considerations".* (p. 19) They also document the on-going discussion on whether humanitarian assistance could, or should, be evaluated. They cite discussions of:

- *"Would the focus of the evaluations be largely on the implementation of activities rather than the results?"* (p. 19)

- *"The ethical issue that the evaluators might hinder aid delivery, that evaluators would 'get in the way' of those providing the humanitarian assistance."* (p. 19)

They further report *"severe doubts"* (p. 21) among the staff of the commissioning agency that such an evaluation was feasible.

Kirkby et al. (2001) cite arguments used against the evaluation of humanitarian aid as:

- *"It was impossible to undertake the evaluation of an emergency since, by the time of the evaluation, the emergency was over"* (p.116)

- Factors such as timeframe of aid, types of intervention, information and funding sources, relationship to the target group and the planning framework. (p.116)

Acknowledging input from a 1993 source Kirkby et al. (2001) sum up some of the differences between humanitarian and development evaluations in Table 1.1.

Table 1.1 Differences between the evaluations of humanitarian and developmental projects

Requirements	Humanitarian	Development
Timeframe	Instant, short term	Sustained, long term
Nature of intervention	Responsive	Planned
Means of acquiring info.	Quick, estimates, surveys	Long term research
Funding required	Timely charity/donations	Long term commitment
Field requirements	Logistics (transport)	Infrastructure (institution building)
Nature of target groups	Passive (receiving)	Active (participating)
Relationship with target group	Top-down	Two-way
Planning frame	Collaboration/cooperation	Coordination/integration

Source: (Kirkby et al. 2001)

3. <u>Who are the important actors in humanitarian coordination, evaluation and learning?</u>

   a. Coordination

The formal coordination of humanitarian action is undertaken by the United Nations Office for the Coordination of Humanitarian Affairs (OCHA). The history of this organisation may be traced back to the adoption of resolution 46/128 by the United Nations General Assembly (1991). This resolution was

> *"designed to strengthen the United Nations response to complex emergencies and natural disasters, while improving the overall effectiveness of humanitarian operations in the field. The resolution also created the high-level position of Emergency Relief Coordinator (ERC)."* (United Nations Office for the Coordination of Humanitarian Affairs 2013b)

The role of the Emergency Relief Coordinator (ERC), Inter-Agency Standing Committee (IASC), the Consolidated Appeals Process and the Central Emergency

Revolving Fund were also created under resolution 46/128. (United Nations General Assembly 1991)

b. Evaluation

Evaluations are usually commissioned by donors or the humanitarian organisation responsible for the project. In a few cases, such as the aid delivered following the genocide in Rwanda, (Borton, Brusset & Hallam 1996) and the humanitarian response to the Indian ocean Tsunami, Borton, Brusset and Hallam (1996) Telford, Cosgrave and Houghton (2006) evaluations may be commissioned by consortia of interested stakeholders. The scope and duration of the evaluation are specified in Terms of Reference (TOR) written by the commissioner of the evaluation which should be agreed with the evaluators.

Evaluations may be carried out by the staff of the organisation responsible for the funding or implementation of the project (internal evaluators) or by independent consultants (external evaluators). Evaluations may be carried out at the end of the project or at a mid-point in the implementation. More recently, real time evaluations (Cosgrove, Ramalingam & Beck 2009) of humanitarian actions are being carried out early in the implementation so as to allow the feedback from the evaluations to be immediately incorporated into improved delivery of the humanitarian action.

Evaluations are documented in the form of an evaluation report which contains the observations of the evaluators as well as recommendations for improvement which may be framed as "lessons learned". It is considered to be best practice for humanitarian organisations to put evaluation reports into the public domain.

c. Learning

Learning, in this context, is the process by which the findings of evaluation and other recorded humanitarian experience is incorporated into best practice and utilised in the design and implementation of future humanitarian actions. This may be done "in house" by donors and humanitarian organisations through the production of guidelines, e.g. (United Nations Inter-Agency Standing Committee 2008), handbooks e.g. (United Nations Childrens Fund 2005) or toolkits e.g. (Norwegian Refugee Council 2004). These documents are usually designed

specifically to meet the needs of the publishing organisation but, in some cases, may gain a wider acceptance in the humanitarian community.

Several organisations have been formed by groups of interested humanitarian organisations specifically to further the learning process.

The Active Learning Network for Accountability and Performance (ALNAP) *"was established in 1997, following the multi-agency evaluation of the Rwanda genocide. ALNAP is a learning network that supports the humanitarian sector to improve humanitarian performance through learning, peer-to-peer sharing and research."* (Active Learning Network for Accountability and Performance 2013a) ALNAP is funded by contributions from its members which include governmental donor organisations, UN organisations and NGOs with the largest contributions coming from governmental donor organisations. See (Active Learning Network for Accountability and Performance 2013b, p. 24)

People in Aid was established in 1995 by a group of humanitarian and development organisations to "*improve organisational effectiveness within the humanitarian and development sector worldwide by advocating, supporting and recognising good practice in the management of people."* (People in Aid 2013a) People in Aid have developed a code of good practice to enhance the quality of human resource management in humanitarian and development organisations. People in Aid is funded by grants from governmental donor organisations (43% in 2013) with the balance coming from what People in Aid describe as "charitable activities" including members' subscriptions and contributions. See (People in Aid 2013c, p. 26)

The Sphere Project was established in 1997 and has a governing board representing 18 humanitarian organisations. The Sphere Handbook, Humanitarian Charter and Minimum Standards in Humanitarian Response, is a widely recognised set of common principles and universal minimum standards for the delivery of humanitarian assistance. This publication The Sphere Project (2011), is revised regularly as part of the on-going work of the Sphere Project. No Annual Report or audited accounts are available from the Sphere Project website. The Sphere Project does not have an independent legal identity but is currently a project hosted by the International Council of Voluntary Agencies (ICVA). The annual report of ICVA for 2012 does not indicate separately the contributions to the Sphere Project. ICVA is, however largely funded by governmental donor organisations See (International Council of Voluntary Agencies 2012, pp. 21-22)

The Humanitarian Accountability Partnership (HAP) was formed in 2003 as a derivative of the 'humanitarian ombudsman' concept recommended by Borton, Brusset and Hallam (1996, p. 167). HAP focusses on making *"humanitarian action accountable to its intended beneficiaries through self-regulation"* (Humanitarian Accountability Partnership 2013a) HAP has developed the Standard in Accountability and Quality Management which humanitarian organisations can adopt in order to become accredited by HAP. The HAP standard is regularly updated by a consensus process so as to remain relevant. The Humanitarian Accountability Partnership is funded largely by contributions from governmental donor organisations (71% in 2013) See (Humanitarian Accountability Partnership 2013c, p. 24)

This brief introduction to the research has outlined the origins of the research, the research itself and the environment in which the research operated as well as introducing the researcher. Many of the topics contained in this chapter will be discussed more fully in subsequent chapters. The next chapter reviews the literature to identify the influences that have acted upon humanitarian action both in its historical development and its present implementation.

Chapter 2 – Literature Review

Humanitarian organisations may be analysed in terms of the interaction between their agency (independence of action) and the structure (environment) in which they operate. It is this analysis that is used in this chapter.

Humanitarian organisations may also be analysed in terms of the principal agent problem: "The problem of motivating one party (the agent) to act on behalf of another (the principal)" (Financial Times 2013), where the principal is the donor and the agent is the humanitarian organisation. This is particularly the case where the humanitarian organisation is heavily dependent on a few influential donors who are able to use a mixture of incentives and sanctions to motivate the humanitarian organisation to act on their behalf. In the analysis that follows these organisations are identified as strongly Wilsonian[2] in nature. This analytical structure is applied in Chapter 8 to specific cases where a clear contractual arrangement exists between two or more parties. However, as, in the wider structure in which humanitarian action is carried out, there are multiple principals acting on a multiple agents, the "agency structure" analytical framework is used.

Giddens (1984) defines agency in terms of *"events of which an individual is the perpetrator, in the sense that the individual could, at any phase in a given sequence of conduct have acted differently" (p. 9)* Sibeon (1999) shows that organisations that have *"causal powers"* can be defined as *"social actors"* which have agency. Several examples are listed including *"voluntary organisations and executive agencies*" (p. 141)

For this analysis the agency of humanitarian organisations will be considered as having two distinct components, the agency that is derived from the humanitarian principles (neutrality, impartiality and independence), and the agency that is derived from their ability to attract resources and gain access to those in need of assistance. In this work these will be referred to as principle-agency and operational-agency and will be hyphenated so as to avoid any ambiguity of meaning when used in context.

---

[2] The classification of humanitarian organisations as Wilsonian is discussed on page 28

Structure includes other actors including donors, host governments, military, media, other humanitarian organisations etc. and circumstances including conflict, disaster, social norms, religious beliefs etc. Giddens (1984) points out that *"Structure is not to be equated with constraint but is always both constraining and enabling"* (p. 25) This point is emphasised by Horst (2005): *Thus the agency of actors is both enabled through and constrained by the structural properties of social systems.* (p. 11)

At its simplest analysis, humanitarian organisations claim a high degree of principle-agency through the humanitarian principles of neutrality, impartiality and independence. However, the social, political and military structure in which these organisations must operate often facilitate their operational-agency (e.g. through provision of resources and access) while at the same time constraining their operational-agency regarding how and where they operate as well the extent to which they can implement their principle-agency. As a result humanitarian actors are constantly trading principle-agency for operational-agency and as a result being instrumentalised[3]. The degree of instrumentalisation encountered is directly related to the extent to which principle-agency is traded for operational-agency and forms an important dividing line between Dunantist[4] and Wilsonian humanitarian actors as analysed below.

Consideration of the effects of the dual role of structure in the delivery of humanitarian action leads the researcher to question; firstly, the extent to which evaluation can be expected to bring about change in the delivery of humanitarian action; and secondly, the extent to which structure limits the agency of the evaluation process itself.

Barnett (2011, pp. 9-12) emphasises the complexity of humanitarianism and of its historical roots referring to these roots as *"the crooked timber of humanitarianism"*. The researcher will show that these crooked timbers which shaped the concept of humanitarianism raise questions regarding the relationship between 21st century humanitarianism and the Dunantist principles that are widely held to underpin it. In the complexities and apparent contradictions of modern humanitarianism, what is the role of the external, independent evaluation that forms the focus of this research?

---

[3] The term instrumentalised means the use of humanitarian resources or actions for non-humanitarian, i.e military or political, objectives or for personal gain..

[4] The distinction between Dunantist and Wilsonian organisations is discussed on page 28

While no records exist, it may reasonably be assumed that at some point in the evolution of humankind as social, migratory and communicating animals, the concept of helping members of the same species, that were not perceived as a threat, and the possibility of receiving similar help in return proved advantageous to the species as a whole. This chapter will show that this mixture of compassion and self-interest pervades the history of humanitarianism and persists to the present time, making the delivery and instrumentalisation of humanitarian action an inevitable symbiotic partnership. Compassion may be considered the motivation for the principle-agency of humanitarians, and self-interest (particularly of others) as the motivation of structures. This partnership has the effect of complicating the philosophy and practice of humanitarian action as well as its evaluation.

Barnett (2011) postulates that *"If we equate humanitarianism with compassion, then humanitarianism is as old as history"*.(p. 19) Nearly all humans exhibit compassion, which in some may be limited to a nuclear family and in others may be demonstrated on a much wider scale. Walker and Maxwell (2009, p. 13) cite the tomb inscription of *"Harkhuf, the governor of Upper Egypt in the 23rd century BCE"* and the *"collection of Chinese cultural and religious practices from the 8th to 5th century BCE"* written by Li Ki as the earliest documented sources.

Walker and Maxwell (2009, p. 14) cite the sending by Cyrene of grain to 41 communities in Greece to alleviate famine between 330 and 320 BCE and the response of the Roman emperors Augustus and Tiberius to famine in Rome between AD 6 and AD 32.

The influences of the monotheistic religions of Judaism, Christianity and Islam which all exhort their followers to charitable acts may be considered the precursors to what today we call humanitarianism. The Jewish law contains several exhortations to charity towards the poor, two notable examples

- *"And when you reap the harvest of your land, you shall not wholly reap the corners of your field, neither shall you gather the gleanings of your harvest. And you shall not glean your vineyard, neither shall you gather every grape of your vineyard; you shall leave them for the poor and stranger: I am the LORD your God."* (Leviticus 19:9-10, The Bible: New International Version 1978)

- *"share your food with the hungry and to provide the poor wanderer with shelter–when you see the naked, to clothe them* (Isaiah 58:7, The Bible: New International Version 1978)

In the New Testament (Luke 10:25-37, The Bible: New International Version 1978) it is the stranger that shows compassion to the man who has been robbed and injured and makes provision for his care in the parable of the good Samaritan.. Christian believers are exhorted to behave in the same humanitarian manner.

In the The Quran English Translation (1930) Zakat (charity) is one of the five fundamental duties of a Muslim. *"We made a covenant with the Children of Israel: You shall not worship except GOD. You shall honour your parents and regard the relatives, the orphans, and the poor. You shall treat the people amicably. You shall observe the Contact Prayers (Salat) and give the obligatory charity (Zakat)."* (Sura 2:83)

It is, however the Judaeo-Christian influence which moulded the type of modern humanitarianism which is the subject of this research. The Islamic concept of Zakat has moulded a parallel system of humanitarianism, that is of growing influence in the 21st century, but which operates outside the framework of the largely Western Organisation for Economic Cooperation and Development (OECD) and its principles of evaluation. See Development Assistance Committee (1991), Development Assistance Committee (1999), Beck (2006).

The two cases cited above, the provision of aid in the Roman Empire and the teachings of the Christian Church on charity came together in AD 325 when the Emperor Constantine adopted Christianity and made it the state religion. (Runciman 1947, p. 3) Walker and Maxwell (2009) make a play on the word adopted and write that *Constantine effectively co-opted Christianity as a state religion"* and *"shifted the burden* (of charity) *to the church, saving the Roman state tax money"* (p. 14) As a result, organised charity in Europe was dominated by the Catholic Church, the state church that became a church state, for over a thousand years.

In his history of humanitarianism Barnett (2011, p. 7) highlights the second half of the 18th century as the time when the next major developments in the concept of humanitarianism took place, in what he refers to as the "*humanitarian big bang".* (pp. 49:56). From this point in history he identifies "*three distinctive ages of humanitarianism":*

1. *"**Age of imperial humanitarianism** from the late 18th century to World War II"* Barnett (2011, p. 7) represents as the period of growth and sustainment of the great, largely European, empires. Barnett and Weiss (2011, pp. 35-46) refer to the period from 1864-1945 as one of *"Birth and maturation"*

2. *"**Age of neo-humanitarianism** from the end of World War II to the end of the Cold War"* Barnett (2002, p. 7) represents as a period when the great empires were dissolving into independent states most of which rapidly became surrogates and proxies of one side or the other in the Cold War, with humanitarian action largely funded by one side or the other exclusively. Barnett and Weiss (2011, pp. 47-69) refer to humanitarianism in this period as *"The traditional enterprise"* while Munslow and O'Dempsey (2010) refer to it as *"humanitarianism curtailed off"* (p. 1224)

3. *"**Age of liberal humanitarianism** from the end of the Cold War to the present"* Barnett (2002, p. 7) represents as the period of the peace dividend, free trade, failed states and localised wars in which humanitarian ethics were rethought and stated in three major documents; identified by Slim (2002) as the Red Cross NGO Code of Conduct, the Humanitarian Charter and the Sphere Standards. Barnett and Weiss (2011, pp. 70-87) refer to this as *"The turbulent post-Cold War era: the new humanitarianism"* and as *Rhetoric meets reality"* (pp. 88-104)

Walker and Maxwell (2009, pp. 60-78) in their consideration of the history of humanitarianism cut short this third period, in a similar way to Munslow and O'Dempsey (2010), and consider it from the end of the Cold War to the global war on terror. Munslow and O'Dempsey (2010) identify this with the title *"humanitarianism comes of age in an era of complex political emergencies (CPEs)"* (p. 1225). They add a fourth, the age of the Global War on Terror from 2001 to 2010, representing a period when two large and prolonged military campaigns have been undertaken by the United States and its allies in the Islamic countries of Afghanistan (with spill over into Pakistan) and Iraq. These conflicts have greatly compromised claims to neutrality and impartiality made by UN and Western aid agencies largely funded by combatant nations. Munslow and O'Dempsey (2010) cut short the period dominated by the Global War on Terror in 2010 and suggest that *"As the preoccupation with a war on terror fades, some of the big humanitarian issues of the 1990s may come back onto the agenda."* (p. 1228) They recognise that it *"inevitably would be a long transition period"* (p. 1228) away from the Global

War on Terror and with the focus on Syria and weapons of mass destruction little of this transition has yet taken place.

In their discussion of the origins of the international humanitarian system Walker and Maxwell (2009) identify three *"competing agendas of humanitarianism"*

1. **Compassion** which they define as "*private acts of charity and state acts to alleviate suffering"* (p. 21). This identifies a pure strand of humanitarianism, unconditional and without any other agenda, the "purity" of "uncompromised" neutrality, impartiality and independence.

2. **Change** which may be identified with those who campaign for reforms that address the root causes of human suffering. Historically this is related by the authors to famines in "*colonial India and the reform of the poor laws back in the British homeland"* (p. 22). Today the researcher would relate this agenda to the concept of "build back better", economic development, human rights, democracy and governance. Barnett (2011) identifies this agenda as that of *"emancipation"*, an agenda in which humanitarianism *"aspires to keep people alive, to expand their opportunities and to give them greater control over their fates"*(p. 11)

3. **Containment** which Walker and Maxwell (2009) define as where *"the worst manifestations of the calamity are addressed primarily to maintain the status quo, to prevent rebellion, and promote security and stability"*. (p. 21) Historically this agenda may be identified in actions to relieve famine in empires from Rome to British India and more recently in the massive amount of aid made available to ex-Yugoslavia and Haiti to prevent large scale migration to EU countries and the United States respectively.

Walker and Maxwell (2009) argue that these agendas are *"doing battle throughout the history of humanitarian action"* and *"competing to determine the role of the humanitarian system in the 21st century".* (p. 21) However, these agendas may not be clearly separated in any one humanitarian action, all three may well be present at the same time, sometimes in the same organisation, competing in a subtle way for opportunistic precedence, often in terms of spin and political justification or securing of resources. A fourth important influence exists: that of the commercial survival and growth of humanitarian organisations and those who work in the field of humanitarian action. This aspect will be discussed later in this chapter.

Age of imperial humanitarianism

Barnett (2011, pp. 45-79) and Walker and Maxwell (2009, pp. 13-21) chart a history of the development of humanitarian concepts through the writings of de Vattel (1757) in defining the role and duties of states and the Enlightenment including the USA the Declaration of Independence, Constitution and the Bill of Rights as well as the ideology (if not the practice) of the French Revolution when along with Liberté, Fraternité and Egalité there was a development of the concept of "*humanité implying a deeply felt concern for the welfare of one's fellow human beings.*" (McLoy 1957, p. 1)

In Britain the role of the Methodist movement is cited by Rack (2004) who identifies that "*Methodist experience of lay organization arguably helped to influence working-class movements in later time.*" (Barnett 2011) writing of the subsequent movement to abolish slavery concludes that *"it is difficult to imagine this rather robust social reform movement without religion"* (p.54)

Slavery and Colonial Humanitarian action

The movement against slavery took nearly 50 years to achieve its major goals from the adoption of its campaign slogan "Am I not a man and a brother" in the 1780s through the Abolition Acts of 1806-07 to the emancipation of slaves in 1834. In doing so it brought together religious roots, civil society (e.g. the anti-Slavery Society and The Aborigines' Protection Society), public advocacy, political action for Government legislation and government action for its funding and enforcement. In embryonic form, many of the influences on the current humanitarian system had begun to work together. (Barnett 2011, pp. 57-60)

With their base in Britain the supporters of the anti-slavery movement quickly began to assert the need for a "benevolent colonialism" using tools of Christian duty and secular self-interest which Porter (2001) observes *"the humanitarian coalition had already shown could become politically unstoppable"* (p. 209) In this context Porter (2001) quotes the speech made by Burke to the British parliament in 1783 during the debate on the East India Bill. Burke speaks of power being "*exercised for the benefit*" of those over whom power is given, that power is a "*trust*" and that *"it is*

in the very essence of every trust to be rendered accountable". (Porter 2001, p. 199) It is interesting to note that the power held by humanitarians over the beneficiaries of their action continues to be high on the agenda of current humanitarian thought, see Humanitarian Accountability Partnership (2010, p. 1)

Walker and Maxwell (2009, p. 18) as well as Barnett (2011, pp. 62-64) who cite the response of the British colonial administration to the famines in India during 1803 and again in 1837 as milestones in the historical development of humanitarian action. Presented as acts of benevolence to the British public at the time, they are viewed much more cynically by 21st century writers.

Sharma (2001) identifies that:

> *"A close reading of the narrative of the famine of 1803–4 reveals that official reactions were riddled with a conflict between a cautious and distrustful attitude and the desire to project an ideology of welfare for their new subjects… It was argued that the superiority of British rule needed to be established on the basis of 'the humane and enlightened policy of the British Government'…The famine was the right time to strengthen the legitimacy of British rule."* (p. 46)

The response to the famine of 1837 was quite different. The events resulting from the famine were seen as a risk to British rule. The motivation of containment took centre stage.

> *"During the 1837–8 famine, 'crime' and the 'breakdown of law and order' were used as indicators of the seriousness of famine for virtually the first time in the region. The famine shattered the complacency of the administration… The events of 1837–8 redefined the place of famine in the colonial agenda by transforming notions of benevolence and responsible rule."* (p. 76)

The concept that long term development projects may be accomplished by cash-for-work and food-for-work projects, aimed at assisting victims of the disaster, was proposed and carried out, although with the subtext that the overall result would be the recovery of the expenditure through increased revenue.

> *'Government should devise some means to relieve present distress and ensure an occupation for those, who would otherwise be forced to emigrate,*

on a work or series of works which would be of permanent advantage to the country, and assist in improving the revenue of future years.' (p. 138)

The results of these interventions were disastrous; in every case the scale of the projects could absorb only a fraction of those who travelled to seek employment. (Sharma 2001, p. 138) Wages were cut below subsistence level but the desperate population had no alternative but to seek the work in increasing numbers. (Barnett 2011, p. 63) Even so the "*works of public utility" were officially also described as works of charity".(Sharma 2001, p. 136)*

The rapid development of humanitarianism that started in the second half of the 18th century with largely a synthesis of "compassion" and "change" with movements for the ending of slavery had, under the hand of privatised colonialism in India, become one of "containment" perhaps tempered with a little "compassion". However cynical colonial governments' responses to humanitarian crises may have been, the Enlightenment had brought about a change in public awareness of a responsibility for the poor, demonstrated by the passing in 1834 of the New Poor Law *"which implicitly recognised that the state had an obligation to help the poor (and then move them back into the labour market)."* (Barnett 2011, p. 63) This approach continued in Britain up to the foundation of the British Welfare System in 1949. (Fuller 2013) The approach taken in Britain, and similarly in other countries, to the humanitarian relief of poverty was, work and wealth were virtues, idleness and poverty were vices, which were considered to be the responsibility of the individual. The seeds of this popular perception are depicted in the work of William Hogarth (1697-1764). (see Paulson 1988)

Humanitarianism in war

In the middle of the 19th century, public compassion was having some success in moderating state self-interest in normal circumstances. However, in war and armed conflict are found circumstances in which, in the heat of battle, self-interest became paramount and there was little or no room for compassion to play a role. At the strategic level, a case of "take the objective whatever the cost" and, at the individual level, "kill or be killed". Some changes had begun to take place and from 1854 to 1857 Florence Nightingale famously worked treating the sick and wounded in the Crimean War. (see Bostridge 2009) It was, however, two years after Florence

Nightingale had left the Crimea that Henry Dunant witnessed the aftermath of the Battle of Solferino in Italy and wrote an account of what he saw. (Dunant 1862)

> Dunant *"described the battlefield as "a disaster from the point of view of humanity." On the battleground lay corpses amid pools of blood and over 23,000 wounded… In response to the atrocious conditions of the injured soldiers at Solferino, Dunant began a drive to assist the wounded in war via private societies."* (Forsythe & Rieffer-Flanagan 2007, p. 6)

Dunant was not alone in his conclusions. Earlier in the 18th Century Jean-Jacques Rousseau, had written that

> "*States declare war against one another, not against individual soldiers. Therefore, when soldiers stopped fighting because of injury or surrender, they cease to be legitimate targets and should be treated with respect."* (Rousseau 1968, pp. 56-57)

Dunant, together with a few like minded Swiss formed "a private society" called International Committee for the Relief to the Wounded in Situations of War (later to become the International Committee of the Red Cross) in 1864 to further Dunant's ideas. (Forsythe & Rieffer-Flanagan 2007, p. 7) Commenting on the realisation of Dunant's vision, Barnett (2011) contrasts Dunant and his compatriots' inspiration of *"God and compassion"* (p. 79) with that of states who *"answered to a higher authority – themselves"* (p. 79). Once again we see the interaction of private compassion and state self-interest. Here we see a situation where the structure in which Dunant had to work limited his agency to realise his vision.

Barnett (2011) identifies:

- the increased violence of military technology,
- more regular use of conscripted armies and
- the development of war reporting and the resulting stirring of pacifist sentiments amongst the public

as motives that convinced states that it was in their own interest to show they were ameliorating the suffering caused by war.(pp.79.80)   Forsythe and Rieffer-Flanagan (2007) relate that

"the result of the convergence of national self-interest and humanitarian spirit in 1863 and 1864 was the first Geneva Convention for Victims of War (1864) which was signed by 12 Western states including Prussia and France (the United States and Great Britain would require more convincing). Its primary contribution was to neutralize the war wounded and the medical personnel who tended them." (p. 8)

Here we see a situation where the self interest of the structure in which Dunant worked facilitated his agency.

The full title of this convention was 'The Geneva Convention for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field' (Geneva Conventions 2013). Thurer (2007) writes that at a time *"when sovereignty was becoming so important the Geneva Convention for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field… constituted a radical break with the past."* (p. 50) He concludes that *"The original Geneva Convention of 1864 constituted a first step towards a fundamental change in the structure of international law, gradually opening it up to embrace individuals and civil society."* (pp. 50-51)

Over the next 85 years, as public compassion and state self-interest were brought together by graphic examples of wartime suffering, three more conventions were facilitated.

- In 1906 the Geneva Convention for the Amelioration of the Condition of the Wounded and Sick in Armies at Sea was signed (Forsythe & Rieffer-Flanagan 2007, p. 43) in the immediate aftermath of the naval Battle of Tsushima which had been fought a year earlier. (Encyclopædia Britannica 2013)

- In 1929 the Convention Relating to the Treatment of Prisoners of War was signed in response to the conditions of World War I *"which saw various abuses committed against captured military personnel." (Forsythe & Rieffer-Flanagan 2007, p. 44)*

- In 1949 the Convention Relative to the Protection of Civilian Persons in Time of War was added following a 20 year delay which had included the mass bombing of cities as well as the development and use of nuclear weapons.

> *"The ICRC after World War I attempted to get legal protections for civilians caught up in armed conflict. However, because of governmental policies, it would have to wait another 20 years before this would come to fruition"* (Forsythe & Rieffer-Flanagan 2007, p. 45)

At the same time, in 1949, revisions were made to the first three Conventions taking account of the significant changes that had taken place since they were drafted. These became known as the four Geneva Conventions and which now have wide acceptance amongst states. (Forsythe & Rieffer-Flanagan 2007, p. 49)

Dunantist Principles

Growing out of the philosophy of the founders of the Red Cross movement and refined by its practitioners came a series of principles widely adopted by the humanitarian community and referred to as "Dunantist". These principles *"define humanitarianism as the neutral, independent, and impartial provision of relief to victims of conflict and believe that humanitarianism and politics must be segregated."* (Barnett 2005, p. 728) These principles have, however, been adopted more widely to cover the delivery of humanitarian action in all circumstances, not just in combat situations, and are enshrined in a resolution of the United Nations General Assembly (1991). *"Humanitarian assistance must be provided in accordance with the principles of humanity, neutrality and impartiality."* (Section I Guiding Principles Paragraph 2).   The centrality of these principles to all humanitarian action is enlarged upon and disseminated by the United Nations Office for the Coordination of Humanitarian Affairs (2011)

> *"Humanitarian principles are central to establishing and maintaining access to affected populations whether in the context of a natural disaster, an armed conflict or a complex emergency."* (p. 1)

Stoddard (2003) sums up the implications for organisations that seek to base their humanitarian action on Dunantist principles: *"Dunantist organisations seek to position themselves outside of state interests."* (p. 2) The history of the development of the four Geneva Conventions referred to above shows, however, that neutral, impartial and independent action can only be realised when states, (or other powerful interest groups) perceive it to be in their interests.

Dunantist philosophy regards their principle-agency, based on neutrality, impartiality and independence, to be the fundamental component of agency for humanitarian organisations by which they achieve access to, and security in, areas that would otherwise be inaccessible. For the ICRC this applies specifically to conflict situations. Dunantist organisations, however, need to attract resources (operational-agency) within a given structural context. To achieve this they claim that it is their principle-agency, and the access it gives them, that makes them deserving of resources. While they are reluctant to give up elements of their principle-agency in order to increase their operational-agency, most would recognise that some compromise is necessary.

<u>Wilsonian Principles</u>

Dunantist principles, and organisations that follow them, are contrasted with Wilsonian Principles which:

> *"Follow in the footsteps of Woodrow Wilson's (1856-1924) belief that it was possible and desirable to transform political, economic, and cultural structures so that they liberated individuals and produced peace and progress* (and) *desire to attack the root causes that leave populations at risk".*(Barnett 2005, p. 728)

However, as one of the first actions of Dunant and his collaborators was, by 1864, to negotiate a convention which transformed political and military structures with the very Wilsonian objective of attacking the "root causes" that left the "wounded and sick in armed forces in the field" and by this definition appear to be acting in a Wilsonian manner.

Not all authors, however, see Wilsonianism in such a positive light as Barnett. Stoddard (2003) defines Wilsonianism as characterising *"most US NGOs"* (p. 2) and quotes Rieff (2003) as writing that Wilsonianism, "*sees a basic compatibility with humanitarian aims and US foreign policy objectives."* Stoddard (2003) observes another indication of the closeness between Wilsonian organisation and government policy: *"Wilsonians have a practical, operational bent, and practitioners have crossed back and forth into government positions."* (p. 2)

The researcher concludes that Wilsonian philosophy regards operational-agency as the fundamental component of the agency of humanitarian organisations by which they are able to achieve their objective of relieving human suffering. Principle-agency is considered a useful tool that may facilitate access and make the organisation politically more deserving of resources. The primary interaction of Wilsonian organisations with structure is aimed at maximising operational-agency.

While Wilsonian thinking had its origins in the 1920s its main impact on humanitarian action took place in the 21st century encapsulated in Colin Powell's referring to "*NGOs who are such a force multiplier for us, such an important part of our combat team". (Powell 2001)* The effects of this approach will be commented on later, at a more appropriate point in the time line. (See page 49)

Humanitarian Space

The creation of the concept of humanitarian space is attributed to Brauman (1996) when he *wrote "Je parle d'un espace symbolique, hors duquel l'action humanitaire se détache du fondement éthique"* (p. 43)*.* Collinson and Elhawary (2012) indicate a traditional relation between the Dunantist principles of Neutrality, Impartiality and Independence and the definition of humanitarian space, citing United Nations Office for the Coordination of Humanitarian Affairs (2003, p. 14):

> *"Humanitarian space is equivalent to a conducive 'humanitarian operating environment' in which agencies can adhere to the principles of neutrality and impartiality and maintain a clear distinction between their roles and functions (saving lives and alleviating suffering) and those of military and political actors"* (Quoted in Collinson & Elhawary 2012, p. 1)

However, Collinson and Elhawary (2012) propose a definition of humanitarian space *"as a complex political, military and legal arena"* highlighting

> *"the highly political nature of the task humanitarian agencies seek to achieve and that humanitarian needs (and their relief) are a product of the dynamic and complex interplay of political, military and legal actors, interests, institutions and processes".*(p. 1.)

This definition reflects very closely the experience of the ICRC, over a period of 85 years (1864 to 1949), in negotiating the four Geneva Conventions in a lengthy

interplay of political, military and legal actors, interests, institutions and processes. Barnett (2011) describes some elements of the process of negotiating the first Convention in 1863-64.

> *"While God and compassion might have inspired Dunant and his compatriots, states answered to a higher authority – themselves – and without their blessing nothing would happen. Their initial, fairly predictable, reaction was negative… They changed their minds, though after concluding that Dunant's proposal might help them legitimate and save war." (Barnett 2011, p. 79)*

It would be naive to believe that similar interplays did not feature in the negotiation of the other three conventions and have continued into the negotiation of more recent treaties banning landmines and various other weapons. Forsythe and Rieffer-Flanagan (2007) sum up the unlikely mix of influences that brought about the first Geneva Convention and which, in the researchers opinion pervade all aspects of humanitarianism in the 21st century *"And therein lies an enduring aspect of the Red Cross Movement: state pragmatism and self-interest alongside humanitarian goals".* (p. 7)

Humanitarian space may thus be equated to the sum of a humanitarian organisation's principle-agency and operational-agency in the environment in which the humanitarian action is taking place.

Private Humanitarian Organisations

 Forsythe and Rieffer-Flanagan (2007) indicate how Dunant's vision was "*to assist the wounded in war via private societies."* (p. 6) The concept of private societies being involved in lobbying for political and social change was not new. As indicated above, organisations such as the anti-Slavery Society and the Aborigines' Protection Society had played an important role in bringing about change in the 18th century.   Dunant's vision, however, differed in two respects from these organisations. Dunant's private societies would be directly involved in delivering humanitarian assistance and would do so in areas of combat and conflict. Barnett (2011) writes *that "Dunant had imagined European volunteers wandering into war to care for the wounded"* (p. 80) Forsythe and Rieffer-Flanagan (2007) report *that*

"Dunant began a campaign to develop a network of private aid societies. In doing so, his goal was to organize volunteers to care for those injured in war" (p. 1)

Walker and Maxwell (2009) indicate that the next important development in private humanitarian organisations took place in the aftermath of the First World War when *"Eglantyne Jebb established the Save the Children Fund to raise funds to send relief to the children behind the blockade"* of Germany and Austria-Hungary. (p. 25)

The same authors state:

> *"SCF is important to the history of the humanitarian system because it represents the first true NGO… It was the first NGO to fundraise, direct its own relief actions, and lobby for international legislation to protect victims of abuse and crisis. In doing so, SCF laid down the model of the independent, activist and operational NGO."* (p. 25)

SCF set a precedent which has been followed by many of the thousands of NGOs which have been founded since and have worked in the delivery of humanitarian action.

International Humanitarian Organisations

Between the two World Wars and under the auspices of the League of Nations the first steps were taken to create International Humanitarian Organisations most notably the High Commissioner for Refugees.

In their biographical notes on Fridtjof Nansen The Nobel Foundation (1922) write: "*In June, 1921, the Council of the League, spurred by the International Red Cross and other organizations, instituted its High Commission for Refugees and asked Nansen to administer it."* The same source credits Nansen with

- The invention of *the "Nansen Passport, a document of identification which was eventually recognized by fifty-two governments."*
- Ministering *"to hundreds of thousands of refugees – Russian, Turkish, Armenian, Assyrian, Assyro-Chaldean"*

- *"Utilizing the methods that were to become classic: custodial care, repatriation, rehabilitation, resettlement, emigration, integration."*

By the end of the 1930s the League of Nations was crumbling and the world was led into a Second World War. United Nations Publications (2000) outline the disintegration

"Several Big Powers failed to support the League: the United States crucially never joined; Germany was a member for only seven years from 1926 and the USSR for only five years from 1934; Japan and Italy both withdrew in the 30s."

By the beginning of the Second World War many of the types of structures which today typify humanitarian organisations were, albeit in embryonic form, already in existence.

Age of neo-humanitarianism

The world emerged from the Second World War almost directly into the Cold War. It was on 5th March 1946, in Westminster College, Fulton, Missouri in the United States of America that Winston Churchill made his "Iron Curtain Speech" that is taken as marking the beginning of the Cold War. In this speech Churchill declared that:

"From Stettin in the Baltic to Trieste in the Adriatic an iron curtain has descended across the Continent… The safety of the world, ladies and gentlemen, requires a unity in Europe, from which no nation should be permanently outcast… I do not believe that Soviet Russia desires war. What they desire is the fruits of war and the indefinite expansion of their power and doctrines". (Churchill 1946)

The division of, first Europe, and later the world, into two power blocs left humanitarian organisations largely isolated on one side. Tight state regulation in the Eastern Bloc stifled the development and even the existence of private humanitarian organisations. The International Humanitarian Organisations, being established by the United Nations but largely dependent on voluntary contributions, were left reliant on the nations of the Western Bloc, for the majority of their funding. This isolation resulted in what Walker and Maxwell (2009) refer to as a period of *"Mercy and Manipulation in the Cold War"* (pp. 46-59)

United Nations Humanitarian Agencies

After its foundation in 1945 the United Nations began to establish agencies to meet humanitarian needs as these needs were perceived to be of international importance by the member states. Over the intervening years four of these organisations have emerged as major players on the humanitarian scene; the High Commissioner for Refugees (UNHCR), the International Children's' Emergency Fund (UNICEF) the World Food Program (WFP) and the Office for the Coordination of Humanitarian Affairs (UNOCHA). The formation of these organisations did not take place rapidly and spans the period from 1946 to 1991.

UNICEF

The first of these organisations to be formed was UNICEF. It was formed in 1946 by the United Nations General Assembly resolution 57(I) (United Nations General Assembly 1946). Its mandate was to assist *"Children and adolescents of countries which were victims of aggression"* (Paragraph 1.1(a)) and for the same group in *"countries at present receiving assistance from the United Nations Relief and Rehabilitation Administration (UNRRA)"*(Paragraph 1.1(b)) The dependence of UNICEF on voluntary contributions is emphasised with the wording *"to the extent of its available resources"* (Paragraph 1.1) and *"The General Assembly expresses the earnest hope that Governments, voluntary agencies and private individuals will give the Fund their generous support"* (Paragraph 10)*.*

UNHCR

As the successor to several previous refugee agencies, including the High Commissioner for Refugees under Nansen, UNHCR was formed in 1950 by United Nations General Assembly resolution 428(V). (United Nations General Assembly 1950) The High Commissioner was given a mandate to

> *"Provide international protection, under the auspices of the United Nations, to refugees… and of seeking permanent solutions for the problems of assisting refugees by assisting… voluntary repatriation… or assimilation within new national communities."* (Annex, Chapter 1, Paragraph 1)

In the case of the High Commissioner for Refugees the dependency on voluntary contributions is defined as, only administrative costs can be funded from the UN budget, "*all other costs shall be financed by voluntary contributions*". (Annex, Paragraph 20) United Nations High Commissioner for Refugees (2013) states *"UNHCR is funded almost entirely by voluntary contributions"* and receives only "2% *from the UN regular budget"*

WFP

WFP was established, on an initial three year experimental basis, in 1961 by the United Nations General Assembly resolution 1714 (XVI) (United Nations General Assembly 1961) with a mandate to *"Establish adequate and orderly procedures on a world basis for meeting emergency food needs and emergencies inherent in chronic malnutrition".* (Annex, Paragraph 13(i)) Its dependence on voluntary contributions was established in the wording *"An initial experimental programme for three years of approximately $100 million with contributions on a voluntary basis." (Annex, Paragraph 1)*

OCHA

OCHA traces its history (United Nations Office for the Coordination of Humanitarian Affairs 2013b) to the general assembly resolution 46/182 (United Nations General Assembly 1991) which was aimed at *"Strengthening of the coordination of humanitarian emergency assistance of the United Nations"*. This resolution

"*created the high-level position of Emergency Relief Coordinator (ERC)… the Inter-Agency Standing Committee (IASC), the Consolidated Appeals Process and the Central Emergency Revolving Fund as key coordination mechanisms and tools of the ERC."* (United Nations Office for the Coordination of Humanitarian Affairs 2013b)

OCHA was not created as a new agency but resulted from a development of an existing UN Agency *"The high-level official should be supported by a secretariat based on a strengthened Office of the United Nations Disaster Relief Coordinator".(UNDRO) (Paragraph 36)* UNDRO had been established in 1971 by the United Nations General Assembly resolution 2861 (XXVI) (United Nations General Assembly 1971). This resolution indicates that the funding *of "disaster relief*

assistance to be carried out by the United Nations its agencies and programmes" (Paragraph 1(d)) would be by voluntary contributions. The strengthened UNDRO was renamed the Department of Humanitarian Affairs (DHA) and located in the UN Secretariat (giving it apparent equal status to the Department of Political Affairs). United Nations Office for the Coordination of Humanitarian Affairs (2013b) reports that

> *"In 1998, as part of the Secretary-General's programme of reform, DHA was reorganized into the Office for the Coordination of Humanitarian Affairs (OCHA). Its mandate was expanded to include the coordination of humanitarian response, policy development and humanitarian advocacy."*

In the case of the United Nations humanitarian agencies the concept of principle-agency may be equated to their "mandate" which is given to them by the General Assembly resolutions by which they were established. In the cases of UNHCR and UNICEF their principle-agency is enhanced by their role as the guardians of specific international conventions. The United Nations humanitarian agencies must, however, obtain their operational-agency largely from the governments of the member states of the United Nations by voluntary contributions. A trading off of principle-agency for operational-agency is inevitable which results in limitations on how, where and to what extent the mandates are realised.

The United Nations international humanitarian organisations have all adopted the humanitarian principles of neutrality, impartiality and independence which are disseminated on behalf of all of them by United Nations Office for the Coordination of Humanitarian Affairs (2011) in order to increase their principle-agency and distance themselves from the United Nations political organs, the General Assembly, and particularly, the Security Council.

Private Humanitarian Organisations

The Cold War years were a time when private humanitarian organisations began to multiply and flourish. They were part of the manifestation of the freedom of association and the development of civil society that was increasingly possible and encouraged by the democratic doctrines of the Western Bloc. As such they were a phenomenon of one side of the Cold War. Increasingly these organisations received funding from Western Governments. Walker and Maxwell (2009) cite three

examples of Private Humanitarian Organisations which, following much of the pattern established by Save the Children Fund, were either established or grew in the Cold War period.

- Oxfam which was formed in 1942 (as the Oxford Committee for Famine Relief) to assist civilians during the Second World War but which broadened its mandate to assist Palestinian refugees in 1951 and then continued to grow. (pp. 42-44)
- CARE formed at the end of the Second World War initially to provide relief to the population of a devastated Europe but which expanded into Latin America in the 1950s and into Africa in the 1960s (pp. 44-45)
- World Vision began assisting children in orphanages in Korea in 1953, expanding into other Asian countries in the 1960s and becoming a global organisation in 1974. (p. 45)

Barnett (2011, pp. 107-131) charts a similar pattern using almost the same organisations as his prime examples.

One organisation which was established as a result of the experiences of some ICRC doctors in Biafra, Médecins Sans Frontières (MSF) is also described by both Barnett (2011, pp. 143-147) and Walker and Maxwell (2009, pp. 46-49) in their accounts of the histories of humanitarianism. ICRC had traditionally interpreted its principle of neutrality in terms of maintaining a public silence regarding events it witnessed while carrying out its mandate to relieve suffering during conflicts. Bernard Kouchner was amongst a group of doctors who having witnessed the acute human suffering in Biafra believed that neutrality could be demonstrated in a more effective way by bearing witness to all atrocities without distinction as to which party to a conflict was responsible. *"Returning to France, Kouchner violated his vow of silence in the most spectacular manner, organising marches and media events to raise awareness and lobbying states." (Barnett 2011, p. 144)* The subsequent formation of MSF resulted in an organisation

> *"that was committed to the humanitarian imperative, but which dispensed with the niceties of national sovereignty and which would not presume that silence in the face of atrocity was the price humanitarians have to pay for neutrality" (Walker & Maxwell 2009, p. 48)*

In order to maintain the high degree of independence from governments that such a position requires MSF has focussed on raising funds largely from individual private donors who sympathise with the moral stand that the organisation takes. Médecins Sans Frontières (2013) claim *that "More than 4.5 million individual donors around the world provide some 90 per cent of our funding. This helps to ensure operational independence and flexibility."* MSF, relying strongly on Dunantist principles, protects its principle-agency by depending heavily, for the resources that it needs for its operational-agency, on parts of structure, individual private donors, that support the organisation without limiting its principle-agency.

Development of humanitarianism in the Cold War

In selecting examples to illustrate the way in which humanitarianism developed during the Cold War and the challenges it faced various writers have chosen different but overlapping cases.

Biafra

As indicated above both Barnett (2011) and Walker and Maxwell (2009) refer to Biafra, not only as the situation that led to the birth of the Sans Frontières movement but also for the way the crisis hit on the international media.

> *"For months the international community ignored Biafra in the same way it ignored other conflicts in the decolonizing world, and then suddenly in early 1968 the famine became worldwide news, transforming Biafra into a cause."* (Barnett 2011, p. 134)

This sudden publicising of human suffering had not only humanitarian implications but political implications as well.

> *Those in the West who might not have cared about the political agenda of the Biafran leadership suddenly became supporters because of the famine, assuming that a people suffering such hardship must have a worthy cause.* (Barnett 2011, p. 134)

The researcher would observe that a very similar linkage between suffering and a perceived rightness of a political cause was repeated in the case of the war in Bosnia and in particular the siege of Sarajevo, where a far more intensive media

coverage of events, ultimately created significant pressure for the intervention that ended the war.

Biafra, Bosnia and, as will be shown later Pakistan and Kosovo, all indicate the importance of the role of the media as part of the structure within which humanitarian organisations must react in order to maintain both principle-agency and operational-agency. This has become known as the CNN effect.

Vietnam

Barnett (2011) cites Vietnam as *"shattering the age of innocence amongst some of the largest US based aid agencies"* (p. 147) He cites CARE and Catholic Relief Services (CRS) as being organisations which *"the United States generously funded their activities, and these agencies in return acted in ways that furthered U.S. foreign policy"* and in doing so *"followed the United States into the quagmire of Vietnam"* (p. 148) Subsequently the negative consequences of their actions were exposed by a journalist which *"caused the agency* (CRS) *such embarrassment that it was forced to close some of its more controversial programs."* (p. 148)

Here is an example where an organisation over relied on operational-agency while ignoring principle-agency with an outcome that was detrimental to the organisation itself. The researcher would suggest that elements of this case study may be repeating themselves in the current Global War Against Terror.

Cambodia

The plight of Cambodians under the Khmer Rouge regime and of the Cambodian refugees in Thailand has been recognised as an important milestone in humanitarian history by several authors. Shawcross (1984) writes critically, in detail, about the situation from the point of view of a journalist looking at events from the outside. Walker and Maxwell (2009) refer to the events in Cambodia and in the refugee camps in Thailand as *"humanitarian action in a Cold War quagmire"* (p.49). Terry (2002, pp. 114-154) and Barnett (2011, pp. 149-158), in discussing the actions of humanitarian organisations in more detail, reveal some large organisations compromising many of their principles before the host government and donors and undercutting each other's negotiating positions as they competed for access and resources.

Terry (2002) focusses on the humanitarian actions undertaken in the Cambodian refugee camps in Thailand where *"Cambodian refugees were an integral part of the shifting political alliances and conflict that had enveloped the Indochinese region since the 1950s"* (p. 115) Terry identifies:

- The *"compromising of moral standards in the pursuit of political expediency"* by major donor governments and some states in the region as they *"pledged financial backing for China's attempts to forge a united front among all* (Cambodian) *opposition movements in Thailand"* (p. 117)*,*

- The role played by Thailand as *"a classic military sanctuary for Cambodian resistance"* (pp. 118-119)

- The impossibility of distinguishing refugees from fighters in the population of camps along the border where *"the proximity of the military camps to the UN assisted camps put the latter squarely under the influence of the former."* (p.122)

- At the same time aid agencies were aware that in the military *camps "many camp inhabitants were Cambodian peasants, confined there against their will and forced to perform military functions"* (p. 122)

WFP, with little access to the camps themselves, delivered food to Thai army warehouses for onward delivery to camps in the south of Thailand (p.123): a situation in which Shawcross (1984, p. 229) claimed to have evidence that WFP knew that the food was being used to feed combatants. Terry reports that in 1979 UNICEF and ICRC worked together to organise food distribution to camps in Thailand along the northern border of Cambodia. With little access to monitor the final distribution these efforts fell victim to inflated population figures and diversion of half the food delivered to feed combatants. In 1980 both UNICEF and ICRC tried to withdraw from this operation but came under heavy pressure from both the Thai and U. S. governments. ICRC withdrew but UNICEF remained having negotiated a significant reduction in the amount of food being delivered and the involvement of two NGOs, Catholic Relief Services (CRS) and World Relief in the delivery. (pp. 129-136) Terry is severely critical in her conclusion

> *Geopolitical strategic and ideological interests dominated the humanitarian concerns, restricting the room for aid organisations to manoeuvre vis-à-vis the host governments and the conditions they imposed. The Thai*

government and Western donors dictated the terms of the border aid program by imposing strict regulations and controlling the financial arrangements that underwrote the relief system." (p. 141)

Barnett (2011) focusses on the emergency unfolding inside Cambodia where *"a terrorised, brutalised and emaciated Cambodian society needed all kinds of aid"* (p. 150) The difficulty was that *"the Cambodian government wanted to use the aid for its own political and military purposes"* and made two demands on humanitarian organisations who wanted to work in Cambodia. The Cambodian government would distribute the aid and organisations which worked in Cambodia could not work in the refugee camps in Thailand. (Barnett 2011, p. 150) Barnett graphically explains the dilemma the humanitarian organisations faced.

> *"Aid agencies were suddenly in a terrible bind. They were desperate to work in Cambodia as reports now portrayed it as the new Holocaust… Yet relief agencies were also reluctant to violate their principles of neutrality and impartiality and capitulate to the Cambodian government's demands."* (p. 150)

ICRC and UNICEF opened negotiations with the Cambodian government in which they were willing to allow the Cambodian government to undertake distribution inside Cambodia but would not compromise their work in Thailand. Oxfam opened parallel negotiations in which they were willing not to work in the camps in Thailand in exchange for the Cambodian government agreeing to allow Oxfam some role in the distribution of aid inside Cambodia. (Barnett 2011, pp. 150-151)   Barnett concludes that James Howard, the Oxfam negotiator

> *"broke with Oxfam's policy of neutrality and impartiality, parted ways with the rest of the NGOs and undercut the positions of the ICRC and UNICEF for several reasons"… a genuine desire to respond… pressure from the British public to respond... and the possibility that Oxfam would become the leader of a consortium of NGOs in this high-profile event."* (p. 151)

Shawcross (1984) relates the story rather differently indicating that *"Brian Walker, Oxfam's director general flew into Phnom Penh on Oxfam's third relief plane."* (p. 150), that *"Walker accepted the government's conditions."* (p. 151) *and "proposed a consortium of voluntary agencies led by Oxfam to begin a huge relief program"* (p.

151) He concludes *that "The aid business is probably no more competitive than any other, but it is sometimes a shock to outsiders that it seems no less."* (p. 149)

Here is an example where inter-organisational competition combined with personal and/or organisational ambition for increased operational-agency resulted in the abandonment of principle-agency. Oxfam was largely successful in achieving some of the operational-agency it wanted but at considerable cost to the organisation's reputation (part of principle-agency).

Brauman (1998) sums up the different positions as *"The choice was… not between a political position and a neutral position but between two political positions: one active and the other by default".* (p. 181) Terry (2002) concludes *"the price of humanitarian access to Cambodia was compromise and silence"* (p. 144) and

> *"Humanitarian principles were compromised along the border and inside Cambodia and room for aid organisations to manoeuvre to claim humanitarian space was extremely limited… Most agencies judged that it was better to remain silent in order to continue to participate in the relief program."* (p. 145)

Along with silence, organisations decided an almost complete abandonment of principle-agency was acceptable if it preserved at least a minimum of operational-agency. The researcher questions to what extent principle-agency can be abandoned and the operational-agency continue to be correctly defined as humanitarian?

Afghan Refugee Camps in Pakistan

Neither Barnett (2011) nor Walker and Maxwell (2009) include the humanitarian operations in the Afghan refugee camps in Pakistan in their histories of humanitarianism. This was, however, the last large humanitarian operation of the Cold War. The Soviet withdrawal was completed in February 1989 and the fall of the Berlin Wall, which is often taken to mark the end of the Cold War, took place in November the same year. Terry (2002, pp. 55-82) discusses the role played by humanitarian organisations in the refugee camps. This also remains an operation of particular personal importance to the researcher as his first posting with UNHCR

was as Head of Office in Quetta (1985-1988) and the second was as Head of Office in Peshawar (1988-1990).

Terry (2002) begins her account recognising that *"Pakistan was one of the most generous and compliant asylum states of the 1980s… Pakistan hosted over 3 million Afghan refugees and resistance fighters"* (p. 55) Centlivres and Centlivres-Demont (1988, pp. 141-152) identify that the Afghan refuges culturally defined themselves in a way that made no distinction between being mohajer (refugee) and mujahed (fighter of Jihad or holy war). The two identities are intimately related in that the mohajer, displaced from his home by infidels, finds refuge with other Muslims. He has a duty to return to his homeland as a mujahed and fight a holy war. He returns into exile from time to time to support his family. Centlivres and Centlivres-Demont (1988) recognise that *"Mohajer does not correspond to the UN definition of refugees and the UNHCR in particular has to distinguish clearly fighters from expatriated victims."* (p. 150) The researcher's experience over the whole of his five years in Pakistan was that this was a hopeless (and thankless) task.

Terry reports that the United States Government, concerned that the Soviet Union was in the process of expanding its influence throughout southwest Asia, provided massive covert aid to the mujahedin (plural of mujahed) fighters which by 1987 grew to $630 million: a sum closely matched by Saudi Arabia. (pp. 58-59) The exodus of millions of Afghans into Pakistan, given a high level of coverage in the Western media, had the same effect as the reports of the plight of Biafrans cited above. Zolberg, Suhrke and Aguayo (1989) argue that "refugees constitute a legitimising population for the warriors", (p. 277) which Terry (2002) concludes resulted in *"a greater justification in the West for continuing active support for the guerrilla war."* (p. 63) The result was a situation in which Terry (2002) states *"the aid community, journalists and academics alike exhibited broad tolerance for the ambiguous (refugee) camp functions.* (p. 55)

The dual identity of the Afghan refugees, the dual nature of the camps in which they lived and the massive amount of military aid being channelled to the mujahedin ensured a situation in which *"the Pakistani Government exercised authority in the camps at every level" (Terry 2002, p. 67)* The humanitarian community, largely funded by the governments and public of the U. S. and its allies, accepted the status quo in the camps. Some NGOs, including MSF, chose to work clandestinely cross border inside Afghanistan supporting the "victims against the oppressors". (p. 73)

Brauman (1997) is clear *"in Afghanistan, MSF never sought to take a neutral stance… we had implicitly picked our side."* (p. xxii)

In this case it was a combination of several factors related to structure that limited the principle-agency of the humanitarian organisations.

- Social concepts that were derived from religious doctrine contradicted established humanitarian definitions and principles,
- The size of the displacement and the media presentation justified the cause of the Afghan mujahedin.
- A significant number of the humanitarian organisations present had been set up to support the Afghan cause.
- The geopolitical position of the donor governments directed humanitarian and military resources to one side of the conflict.

There appears to have been considerable confusion regarding how to apply principle-agency when the beneficiaries operated in a different principle framework, one with which humanitarians came to at least partially agree. Massive funding for operational-agency and media generated public support for the Afghan refugees allowed humanitarian organisations to work with minimal regard for principle-agency without the risks that had been experienced in the Vietnam and Cambodia situations.

Conclusions

The age of neo-humanitarianism defined by Barnett (2011) almost completely coincided with the Cold War and the development of humanitarianism was greatly shaped by its close association with one side of this conflict. Western organisations, staffed at senior levels by western staff and largely funded by western government money, were willing to compromise their humanitarian principles in order to assist the largest number of beneficiaries and in doing so grow as important players in the delivery of humanitarian action. A general belief in the rightness of western freedom to fight against Soviet oppression justified compromise and compliance. Brauman's comment regarding MSF in Afghanistan could have been applied to a greater or lesser degree to many aspects of humanitarianism in the closing decades of the Cold War.

Academic criticism had, however, begun to focus on some of the deficiencies of humanitarianism in the 1980s. Harrell-Bond (1986) criticised the humanitarian operation for the Ugandan refugees in Sudan. She took this operation as an example of humanitarianism in general in a criticism that focussed on the paternalistic attitude of humanitarians and the disempowerment of the beneficiaries. "Imposing Aid" was a landmark in the academic criticism of humanitarian action and revealed attitudes and method that humanitarians have failed to rectify completely in the intervening decades. These early critical analyses of humanitarian action were resisted by some humanitarians and Waldron (1987) opens his paper almost apologetically stating that "*Criticising refugee relief is an effort likely to produce much the same response as, for example, sending mother's apple pie to the FDA for chemical analysis*" (p. 1)

As the Soviets withdrew from Afghanistan, the Berlin Wall fell and the Cold War ended, the humanitarian environment was about to change radically.

"***Age of liberal humanitarianism*** *from the end of the Cold War to the beginning of the Global War on Terror (GWOT)*

Middleton and O'Keefe (1998) observe that:

> "*It is unusual to come across commentary on contemporary humanitarian emergencies which does not, somewhere, point to the end of the Cold War, and the consequent change in geo-politics, as part of the complexity of their causes.*" (p. 16)

In this work the researcher will maintain the tradition. As Middleton and O'Keefe (1998) point out, this was a time when states which had been supported for reasons of colonial history, valuable raw materials and/or superpower competition, were allowed to fail. An early attempt at a humanitarian intervention in Somalia had a disastrous outcome and deterred interventions in subsequent conflicts. (See Livingston and Eachus (1995) and Western (2002)) One outcome was that governments funded humanitarian relief efforts rather than engage in political action to solve the underlying conflict. Weir (2006) summed up the result "*As Sadako Ogata (former High Commissioner for Refugees) has said, "there are no humanitarian solutions to humanitarian crises," mass displacement, hunger, disease, etc. – are the humanitarian fallout of political failures.*" (p. 16)

Two of these conflicts were particularly traumatic for humanitarians as well as for the affected population. In the breakup of the former Yugoslavia, a brutal war in the middle of Europe was allowed to continue for three years (2002-2005). Humanitarian suffering continued while only the most ineffective political action was taken. Finally the war was ended in a few weeks when ultimately there was the political will, see (Glenny (1996) and Holbrooke (2011)) At the same time (2004), in Rwanda, genocide was allowed to take place only to be followed by political manipulation, repression and disease in the refugee camps in the Democratic Republic of Congo, (see Dallaire and Beardsley (2004), Barnett (2002) and Terry (2002)). It was a situation that Rieff (2000) refers to as *"the worst version of the humanitarian idea – aid as a fig leaf."* (p. 23)

In this case the principle-agency of humanitarian organisations was upheld both by the organisations themselves and the donors who generously provided resources that enhanced operational-agency. The results demonstrated the futility of humanitarian action without political action to address the root causes of conflict and suffering. One result was to encourage the conceptualisation of the New Humanitarianism discussed by Fox (2002)

Evaluation of humanitarian action

Evalsed (2009) identifies the origins of evaluation:

> *"Evaluation emerged as a distinct area of professional practice in the post-war years in North America. Three strands that were most important in that early period were the evaluation of educational innovations (e.g., the effectiveness of new curricula in schools); linking evaluation with resource allocation (e.g., through a Planning, Programming and Budgeting system); and the evaluation of anti-poverty programmes (e.g., the Great Society experiments of the 1960s)".*

Evalsed (2009) also identifies *"four main groups whose interests sometimes compete with each other in defining evaluation priorities:*

1. *Policy makers, e.g., elected officials and politicians… mainly interested in accountability.*
2. *Managers and administrators, e.g., civil servants and managers of local public agencies… concerned with the delivery of policies and programmes how well they are managed and organised.*

3. *Citizens and those affected by public action, e.g., the presumed beneficiaries of planned interventions… regard evaluation as a tool for democratic accountability and an opportunity to shape public interventions to their needs.*

4. *Professional and specialist interests, e.g. teachers in education or scientists in research… mainly interested in evaluation as an opportunity to improve the quality of their work or even the autonomy of their own professional group."*

The first three of these groups all focus on accountability for quality and performance in one form or another, the fourth, however, has more of a focus on evaluation as a social learning tool.

The evaluation of humanitarian actions began to be undertaken in the 1990s largely commissioned by government funding organisations and with an accountability focus. The closer involvement of humanitarian organisations in the commissioning and management of evaluations (see Fig 8.1 p259) has tended to move the emphasis of the evaluation of humanitarian action towards a greater focus on social learning.

The history of the genesis and development of this branch of evaluation between the end of the Cold War and the beginning of the Global War on Terror is discussed more fully in Chapter 6, through an analysis of the experiences of undertaking nine evaluations reported by Wood, Apthorpe and Borton (2001b) The criteria used for the evaluation of humanitarian actions were initially established for the evaluation of development projects (Development Assistance Committee 1991). They consisted of five criteria: efficiency, effectiveness, impact, sustainability and relevance. These were subsequently adapted for the evaluation of complex emergencies, (Development Assistance Committee 1999, pp. 30-32):

- sustainability was replaced with connectedness after Minear (1994)
- two new criteria were added, coherence and coverage

This resulted in the seven criteria in common use today for the evaluation of humanitarian action. However, humanitarian action is often carried out in situations that are not defined as complex emergencies for which Development Assistance Committee (1999) was written and the more recent guidelines of Beck (2006) are now more appropriately used.

In 1999 Kosovo marked a political turning point that had a major effect on some future humanitarian responses. *"NATO decided to make war on Yugoslavia rather than, as in Bosnia, to stand on the side-lines and allow Milocevic to have his way"* (Rieff 2000, p. 25) Perhaps unsurprisingly after being instrumentalised by inaction *"a number of NGOs, while not overtly calling for outside military intervention on humanitarian grounds, certainly fuelled these calls with their reporting"* (Rieff 2000, p. 27) Rieff's criticism of the humanitarian organisations is scathing. Caught up in "the media-driven humanitarian crisis par excellence" (p. 26), *"NGOs effectively became sub-contractors to one side of the war"* (p. 28)*,* finding it *"virtually impossible… to abstain from participating in a major operation… not only do they risk losing market share… but alienating their principle governmental and institutional funders."* (p. 29)

Rieff (2000) makes the case that in Kosovo media attention and political pressure from major donors to increase their operational-agency in order to facilitate the donor's political and military objectives overwhelmed humanitarian organisations' ability to maintain their principle-agency. As in Pakistan it proved to be a relatively risk free option. These were circumstances where it was safer to be Wilsonian than Dunantist.

The 1990s were a period when either as a result of political/military action or inaction humanitarian organisations' principle-agency was constrained by the structure which facilitated their operational-agency with access and resources while instrumentalising their operations. Rieff (2000) sums up the humanitarians' "agency – structure" trade off from an outside, journalistic view. *"Relief groups may be misused by governments but for the most part they connive in this misuse"* (p. 32)

Humanitarian reaction

The decade of the 90's had a significant effect on humanitarian thought. A number of key elements can be identified:

- A rapid increase in humanitarian budgets and a resulting demand for evaluation mechanisms (discussed in Chapter 6)

- A growth in the number of humanitarian actors and a call for better coordination which led to humanitarian reform with the creation and strengthening of coordination mechanisms through OCHA, referred to above.
- A movement of humanitarian action into conflict areas.
- The restatement, development and consolidation of a broad based and widely accepted humanitarian philosophy.

Slim (2002) identifies a decade starting in 1992 (the beginning of the war in Bosnia) when the unprecedented media coverage of extreme violence of civil war brought together highly interdisciplinary groups studying the questions of war and humanitarian action. These groups included "*NGO activists from a variety of different traditions"* and "*in this process the fundamental values of the humanitarian ethic and the principles of its practical application have been 'rediscovered', held up to the light, scrutinised, dismissed or reaffirmed"* (pp. 113-114) Slim emphasises that this was not an exercise carried out by academics or politicians but by *"the community of international NGOs concerned with humanitarian action."* (p114)

Three important documents are identified as emerging from this process

- The Red Cross/NGO Code of Conduct
- The Humanitarian Charter
- The SPHERE Minimum Standards in Disaster Response.

At the time of writing this thesis, all three of these documents have been incorporated into a single publication, The Sphere Project (2011). Slim (2002) *concludes, "Although initially seen as a rigorous attempt to 'put their house in order', this process of humanitarian writing has resulted in a deeper process of explicit recommitment to humanitarian values"* (p. 114) These values may be identified as principle-agency.

The Code of Conduct begins by stating the importance of the humanitarian imperative, defining and demanding it. It is a clear statement of rights and duties, a statement of intent and a solid position from which to negotiate.

> *The humanitarian imperative comes first – The right to receive humanitarian assistance, and to offer it, is a fundamental humanitarian principle which should be enjoyed by all citizens of all countries. As members of the international community, we recognise our obligation to provide*

humanitarian assistance wherever it is needed. Hence the need for unimpeded access to affected populations is of fundamental importance in exercising that responsibility. (The Sphere Project 2011, p. 370)

If the Code of Conduct is a statement of humanitarian strategic intent the Sphere Standards are a practical tactical guide as to exactly how humanitarian responsibility should be carried out. To take a military metaphor, the researcher would refer to them as the humanitarian 'rules of engagement'. Slim (2002) is even more graphic *"Sphere standards are not just the stuff of general moral obligation but present a very precise latrine-based ethics."* (p.116)

The third document, the Humanitarian Charter, is best left to introduce itself as the legal and ethical basis on which humanitarianism is based

The Humanitarian Charter:

"is in part a statement of established legal rights and obligations; in part a statement of shared belief.

In terms of legal rights and obligations, it summarises the core legal principles that have most bearing on the welfare of those affected by disaster or conflict. With regard to shared belief, it attempts to capture a consensus among humanitarian agencies as to the principles which should govern the response to disaster or conflict, including the roles and responsibilities of the various actors involved." (The Sphere Project 2011, p. 20)

The closing years of the 20th century were a time of highly productive thinking, writing and organising in the humanitarian arena.

- People in Aid was formed in 1995 to improve *"organisational effectiveness within the humanitarian and development sector worldwide by advocating, supporting and recognising good practice in the management of people." (People in Aid 2013a)*
- ALNAP was formed in 1997 as a network to promote learning and improvement amongst humanitarian organisations.
- Walkup (1997) identified the causes and effects of stress on humanitarian workers and on the organisations they made up.

- Anderson (1999) identified the ways in which aid delivered in conflict areas could do more harm than good and proposed ways in which unintended consequences could be minimised.
- The Development Assistance Committee (1999) set down guidelines for evaluating humanitarian assistance in complex emergencies.
- It was at the turn of the century that the United Nations Secretary General asked the important humanitarian question to the nations of the world

> *"if humanitarian intervention is, indeed, an unacceptable assault on sovereignty, how should we respond to a Rwanda, to a Srebrenica— to gross and systematic violations of human rights that offend every precept of our common humanity?" (Annan 2000)*

It was during this decade that the evaluation of humanitarian action was born and grew up. Kirkby et al. (2001) chart the doubts and concerns that surrounded the first evaluations *"with some agencies arguing that it was impossible to undertake an evaluation of an emergency since, by the time of the evaluation, the emergency was over."* (p. 116) One of the authors of this paper had undertaken the first evaluation when this question was raised. See (O'Keefe et al. 2001) However,

> *"the argument about the feasibility of evaluating humanitarian assistance was never resolved. It disappeared as bilateral donors commissioned a series of evaluations, essentially avoiding the argument."* (Kirkby et al. 2001, p. 118)

A number of relatively experimental evaluations of humanitarian projects were undertaken between 1993 and 1996 but it was the horrors of the Rwanda genocide that gave the impetus to the commissioning of the first major evaluation of a humanitarian crisis, the response to which was the subject of the evaluation conducted by Borton, Brusset and Hallam (1996). The experiences of the evaluators undertaking nine evaluations between 1993 and 2000 are analysed in Chapter 6 and the recommendations made by Borton, Brusset and Hallam (1996) are analysed in Chapter 4

By the time "Claiming the Humanitarian Imperative" (Slim 2002) had been published the events of 11 September 2001 had taken place and humanitarianism was beginning to face working in a world dominated by the Global War on Terror.

Humanitarianism in the Global War on Terror

Not only was Slim's triumphal proclamation of the claiming of the humanitarian imperative overtaken by the destruction of the World Trade Centre but it was also overtaken by Colin Powell's address to a group of American NGOs, reported by Slim (2004) as *"remarks that have now become notorious in NGO circles" (p. 43)* Powell is quoted as saying, *"I am serious about making sure that we have the best relationship with the NGOs who are such a force multiplier for us, such an important part of our combat team.* (Powell 2001) For the Dunantist, and even for many of the more Wilsonian NGOs, the idea that they were part of a "combat team" was simply unacceptable.

Amongst military personnel, however it was not just the NGOs who were to be conscripted as *"part of our combat team"*; it was the very fabric of humanitarianism. Although the researcher has verbally heard similar statements from various military personnel, two documented sources have been identified for inclusion in this literature review. At an official strategic level the United States Army "Commanders Guide to Money as a Weapons System" (United States Army 2009) states

> *"The Commander's Emergency Response Program (CERP) enables local commanders in Afghanistan and Iraq to respond with a nonlethal weapon to urgent, small-scale, humanitarian relief, and reconstruction projects" (p. 13)*

A horrific tactical realisation of the conscription of humanitarian relief into military strategy is reported in an interview with Lieutenant General Victor E. Renuart, Jr., of the United States Air Force in which he states:

> *"Another non-kinetic means that can be very effective is humanitarian assistance. During* (Operation Enduring Freedom) *OEF on the very first night of lethal operations, we dropped 75,000 pounds of bombs on targets in Afghanistan and began dropping 75,000 humanitarian daily rations out of C-17s. The people understood we weren't threatening them, that we were feeding them and killing bad guys.* (Hollis 2004, p. 7)

The bright yellow wrappers of the humanitarian daily rations are a similar colour to some unexploded cluster munitions.

At a more systematic level the military has begun to adopt the Sphere Standards for military humanitarian interventions which they perceive as legitimising the use of military assets in a humanitarian context.

> *"Applying Sphere project standards* (which) *are well-proven means to measure humanitarian assistance effectiveness and would place the* (Department of Defence) *DoD on a firm ground with regards to international humanitarian law and on an equal footing with other providers of health care in humanitarian relief"* (Drifmeyer & Llewellyn 2004, p. 167)

While this approach may have some validity in major natural disasters, it meets with justifiable criticism when applied in combat situations where the line between combatant and non-combatant is blurred when military personnel directly engage in operations they refer to as humanitarian.

> *"In such a context, it is likely to see the Sphere Minimum Standards being used by non-humanitarian actors to legitimise their actions. An officer from the Coalition in Afghanistan expressed his surprise at NGOs reactions' against the engagement of the military in humanitarian operations: 'Why are they against us? We also use the Sphere standards.'".* (Dufour et al. 2004, p. 139)

Where does this leave humanitarianism in the second decade of the 21st century and what challenges does it face?

On one hand, never before have humanitarians been so numerous or had access to so many resources. In the first few months following the earthquake in Haiti "*over one thousand international organisations had provided humanitarian assistance*" (Grünewald, Binder & Georges 2010, p. 7) The United Nations Office for the Coordination of Humanitarian Affairs (2013a) report that Global Humanitarian Contributions in 2012; Including Consolidated Appeals, natural disasters response, bilateral aid, and all other reported humanitarian funding, was US$ 12.7 billion. Beneath this prosperity, however, lie some disturbing factors. The analysis of the figures behind the US$ 12.7 billion bottom line is indicative of one of the key elements:

- 30% of the funding comes from one donor, the United States of America.
- 50% comes from the five top donors USA, European Commission, United Kingdom, Sweden and Japan

- 68% comes from NATO member states and the European Commission

About two thirds of the funding available for humanitarian action comes from the USA and its allies in NATO (the European Commission is included as the majority of the member states of the European Union are NATO members) the countries most involved in the Global War on Terror.

In the majority of the international humanitarian organisations, 68% of their funding and many of their senior international staff all originate from nations actively involved in the Global War on Terror. These humanitarians find themselves claiming neutrality, impartiality and independence, with the protection that international humanitarian law affords them as non-combatants, while the military of the same nations are engaging in combat operations, as well as actively carrying out similar "humanitarian" operations. It is hardly surprising that the security of humanitarian workers is one of the biggest challenges faced by many humanitarian organisations.

In the Global War on Terror, humanitarianism has not only been instrumentalised by politicians but also by the military acting in the name of national security

Humanitarian Challenges in the 2010s

This review looks at two recent works that give similar overlapping lists of the major challenges faced by humanitarianism in the second decade of the 21st century. While these lists may not be considered exhaustive they do cover a very broad spectrum of the issues of concern to humanitarians at this time.

Table 2.1 Comparison of the challenges to humanitarianism in the 21st century as identified by Barnett and Weiss (2011) and Walker and Maxwell (2009)

General topic	Barnett and Weiss (2011)	Walker and Maxwell (2009)
Philosophical and ethical issues		
Scope & limits of humanitarianism	What should humanitarian do? (pp. 105-108) What can humanitarians accomplish? (pp. 108-109)	Dilemma of addressing effects and causes? (pp. 140-142)
Humanitarian principles	Which principles, when? (pp. 109-110)	Future of humanitarian principles in today's politicised world (pp. 136-140)
Humanitarian ethics	What kind of ethics? (pp.112-114)	
Humanitarian Military Cooperation	Do we want a militarised humanitarianism (pp. 110-112)	
Institutional Issues		
Growth	Is bigger better (pp. 114-115)	The challenge of growth and deregulation (pp. 142-143)
Professionalization	Is it possible to be too professional? (pp. 115-118)	Organisational and system learning and change (pp. 147-148)
Donor relations		Financial Independence (pp. 143-145)
Victims or survivors	Can we mobilise action without creating a world of victims? (pp. 118-121)	
Host society issues		
Host country and population	Do local views matter? (pp. 121-123)	The challenge of relationship (pp. 145-147)
Universality	Is humanitarianism universal? (pp. 124-127)	

The researcher will consider the challenges faced by humanitarianism under the headings in the first column of table 2.1 "General Topic".

Philosophical and Ethical Issues

Scope & limits of humanitarianism

The key question regarding the scope and limits of humanitarian action concerns how far down the problem tree (See Overseas Development Institute 2009), which moves from causes to effects, should a humanitarian organisation operate? Weir (2006) identified the difficulty in the quotation used above *"As Sadako Ogata (former High Commissioner for Refugees) has said, "there are no humanitarian solutions to humanitarian crises, mass displacement, hunger, disease, etc. – are the humanitarian fallout of political failures."* (p. 16) This being the case should humanitarians be content in addressing the effects or follow the causes and become involved in their political solution?

Barnett and Weiss (2011) point out that limiting humanitarian action to meeting immediate needs with humanitarian aid is an achievable goal that remains politicised to the minimum extent and allows access to the population in need. This approach does not solve the problem which may persist or recur on a regular basis. However, it may not be obvious how to change the political causes, and the political involvement of humanitarians may be resisted and limit or block access to the population in need. *"Those in power hardly welcome aid agencies that want to change the world."* (p. 106) Walker and Maxwell (2009, p. 141) cite Darfur as being a recent example where advocacy has cost humanitarian agencies their access, an example of where the implementation of principle-agency heavily reduced operational-agency.

Barnett and Weiss (2011) point out, however, that

"*A too limited version of humanitarianism may well down-grade what is possible. It is not as if we are as ignorant today as we were several decades ago regarding what may be done to improve the lives of vulnerable populations"* (p. 108)

Humanitarian Principles

Barnett and Weiss (2011, p. 109) point out that the principles of neutrality, impartiality and independence are not of long standing dogmatic origins, but were codified by the ICRC in the 1960s. They had been found to be *"functional for engaging in rescue and relief in highly charged political circumstances"* (p. 109)

Barnett and Weiss (2011, p. 109) identify the principles of neutrality and independence as "more instrumental than intrinsic". This may be contrasted with the observation of Walker and Maxwell (2009, p. 137) that *"neutrality had almost become a dirty word in the aftermath of the Balkans wars and the Rwanda genocide"* and cite this as one of the motivations for the concept of the New Humanitarianism (p. 138) proposed by the United Kingdom Department for International Development (United Kingdom Parliament 1999) and discussed by Fox (2002). One effect of the New Humanitarianism was to encourage humanitarian organisations to take an active view of neutrality, "*holding all sides in a conflict accountable to the same standards"* (Walker & Maxwell 2009, p. 139) rather than a more passive approach of keeping silence. The same authors point out that the environment of the Global War on Terrorism has restricted the ability of humanitarians to implement this active approach to neutrality. (p. 139)

<u>Humanitarian Ethics</u>

Barnett and Weiss (2011, pp. 112-113) identify two approaches to the ethics of humanitarian action.

- Deontological approach based on duty. This is further defined by Slim (2002) when he refers to;

    "the notion of duty sharing so prominent in the Code, the Charter and Sphere. If a person, a government or an organisation cannot or will not abide by a humanitarian duty then that duty automatically falls to others. As a categorical imperative, humanitarian duty is boundless" (p. 118)

- A consequential approach in which *"we are morally obliged, therefore, to act in a way that produces the best consequences given the alternatives".* (Barnett & Weiss 2011, p. 113) Consequences and unintended consequences are at the foundation of the concept of 'do no harm' when applies to humanitarian action. (See Anderson 1999). This approach raises a number of problems associated with deciding whether the best consequences have been achieved, including how visible all the consequences are, the timeframe in which they are observed and *"Whose consequences matter"?* (Barnett & Weiss 2011, p. 113)

Humanitarian-Military Cooperation

Barnett and Weiss (2011) introduce this topic with the rather simplistic question *"Do we want a militarised humanitarianism?"* (p. 110). To a large extent this question must be considered in the context of an on-going Global War on Terror which carries the implication, based on the historical facts of the invasions of Afghanistan and Iraq, that the United States and its allies are willing and able to attack any country from which they believe terrorists threaten their security. As indicated above, the direct delivery of aid by the military in Afghanistan, using the term humanitarian, based on their adoption of the Sphere Standards has shocked humanitarian organisations.

There is, however, a clear humanitarian doctrine for working relations between the humanitarians and the military in natural disasters, see United Nations Office for the Coordination of Humanitarian Affairs (2007). O'Keefe and Rose (2008) summarised the key principles as:

- *"Complementarity, which implies that the military will not be used if civilian assets are available;*
- *Control of the military in support of humanitarian action must be the responsibility of civil authority;*
- *No costs associated with the military can be charged to the affected population*
- *The military must withdraw at the earliest possible moment."* (p. 461)

In complex emergencies United Nations Inter-Agency Standing Committee (2008) sets down guidelines, a summary of an earlier version of this document is given by Bessler and Seki (2006) The application and compliance with guidelines that are widely accepted by the humanitarian community should provide a basis for a working relationship between the military and humanitarian organisations. Such a relationship facilitates the delivery of humanitarian aid with the minimum of devaluation of the principles by which it is given. However, it is also one in which humanitarian organisations will have to trade some principle-agency to gain the operational-agency of access to, and through, conflict areas. Depending on their position on the Dunantist-Wilsonian spectrum, humanitarian organisations will differ on their willingness to accept the compromises necessary

The turbulent history of the concept of Responsibility to Protect (R2P), covered in some detail in Chapter 4 of this thesis, has added layers of suspicion to humanitarian/military cooperation particularly following the use of R2P language to bring about regime change in Libya, (see United Nations Security Council (2011), Paragraph 4)

Institutional Issues

<u>Growth</u>

Both Barnett and Weiss (2011, pp. 115-115) and Walker and Maxwell (2009, pp. 142-143) address the issue of growth of individual humanitarian organisations. Walker and Maxwell (2009, p. 120) list the expenditures of the five largest humanitarian NGOs for 2006 which, if totalled, amount to US$ 5.46 billion. The largest of the group is cited as being World Vision International with an expenditure in 2006 of US$ 2.1 billion. Karajkov (2007) using similar figures for 2006, comments: *"In terms of size and financial strength, some of the biggest N.G.O.'s are to be found in the realm of humanitarian work."* He compares the size and growth of the largest of the NGOs with that of a small country, writing,

> *"Montenegro, one of the latest countries to gain independence, in 2006 had a G.D.P. of $2.2 billion. It is considered one of the fastest growing economies in the world, by states' standards. N.G.O.'s, however, grow faster."*

Barnett and Weiss (2011, p. 115) focus on the way in which, simultaneously, growth allows organisations to apply the economies of scale, while at the same time making it harder for them to innovate and change the way they operate. Walker and Maxwell (2009, p. 143) highlight that growth has brought with it more demands for accountability, standards and professionalism.

Growth into large international organisations, and the influence of powerful patrons, has led to large NGOs appointing Chief Executive Officers (CEOs) from the world of business and government who command high salaries. Hope (2013) reveals that 14 of Britain's leading foreign aid charities have executives earning over UK£ 100,000 per annum, with the chief executive of the British Red Cross earning UK£184,000. This may be compared to the national average wage of about UK£ 23,000 (Office of National Stastics 2013). These revelations caused more than a little embarrassment to organisations that advertise for small contributions to aid the poorest on the

planet. CEOs drawn from commercial backgrounds are also more likely to see continuous growth as a sign of a healthy organisation and consider factors such as market share and competition in the market as important policy objectives (See Rieff 2000 cited above)

Professionalization

This is another topic on which both Barnett and Weiss (2011, pp. 115-118) and Walker and Maxwell (2009, pp. 147-148) comment with different emphasis. Barnett (2002) emphasises the risks of moving from a volunteer cadre to a professionally qualified one, including the unequal access to training for qualification and the increase in technocracy, that typifies professions and which "*places rules above people".* (p. 118)

Walker and Maxwell (2009) also recognise the problems of humanitarian technocracy but focus on the need to improve the *"limited capacity for institutional learning among humanitarian agencies".* (p. 148) They identify humanitarian organisations as *"unwilling to take responsibility to learn when things go wrong"* (p. 148) and donors who *"don't want to hear about failures"* (p. 148) They attribute the lack of change in *"agency strategy and behaviour"* coming from evaluations as being due to organisations being *"locked into standard procedures and responses".* (p. 148) Walker and Maxwell (2009) recommend the classic features of professions, links between academia and practice with the humanitarian equivalent of teaching hospitals and medical journals, as well as more opportunities for graduate and post graduate education in humanitarian issues. (p. 149)

Donor Relations

Walker and Maxwell (2009, pp. 143-145) consider this topic as one of "financial independence" and the difficulty in maintaining this when, as indicated above, 50% of humanitarian funding comes from five donors. Barnett (2005) refers to the few donors who *"now comprise an oligopoly"* (p. 727) as a major factor of politicisation of humanitarianism through *"the political economy of funding… as many states either expect something in return or evidence that their money was being well spent."* (p. 727) A few Dunantist organisations such as the ICRC and MSF manage to maintain a high degree of independence from donors, but by differing methodologies. The ICRC is funded by voluntary contributions from the states party

to the Geneva Conventions (governments) amongst others. It maintains its independence (operational flexibility):

"The ICRC's operational flexibility was preserved as a number of governments continued either not to earmark their contributions or to do so in a relatively broad fashion (mostly by geographical region)" (International Committee of the Red Cross 2012, p. 75)

MSF maintains a funding base that is 90% dependent on private contributions as indicated above (Médecins Sans Frontières 2013)

<u>Victims or Survivors</u>

Barnett and Weiss (2011, pp. 118-121) address this issue with a question *"Can we mobilise without creating a world of "victims"?"* (p. 118) and suggesting that mobilising resources for humanitarian action *"may require some degree of exploitation of the very people we want to help… the more graphic the image and the more it screams 'innocent victim' the more effective it will be in mobilising compassion, action and money."* (p. 119)   Horst (2005, pp. 93-103) points out a related problem, that humanitarians tend to believe their own propaganda. They expect the beneficiaries to act as vulnerable victims rather than resourceful survivors. When the beneficiaries do not behave as vulnerable victims they are often regarded as cunning crooks.

Host Society Issues

<u>Host Country and Population</u>

Walker and Maxwell (2009, pp. 145-147) consider the changes taking place inside and between organisations in the field. Large international organisations are converting field offices into "local autonomous organisations" (p. 145) creating a structure which, on the outside at least, looks like the Red Cross model of national societies. The question is whether they will become true international federations giving up power to their national organisations and, if they do, how will accountability to Northern donors be maintained? (p. 146) Walker and Maxwell (2009) also consider the growing problem of coordination as organisations, all espousing independence as a principle, proliferate. They observe that *"Everyone wants coordination; no one wants to be coordinated"* (p. 146)

Barnett and Weiss (2011, pp. 121-123) approach the subject again with a question *"Do local views matter?"* They point out that

> *"There is a worrying gap between what most humanitarians would say about the crucial importance of input from the beneficiaries, on the one hand, and what they tend to do in an emergency on the other"* (p. 121)

The Humanitarian Accountability Partnership (HAP) was established in 2003 to *"make humanitarian action accountable to the beneficiaries"* as its webpage header proclaims. A comparison of the organisations which are members of HAP *("and are Entitled to vote at HAP General Assembly meetings… eligible for election to the HAP Board.")* with the organisations which have been HAP Certified (assessed for compliance with the HAP Standard in Accountability and Quality Management by an independent audit) is revealing. (Humanitarian Accountability Partnership 2013b)

HAP has 69 organisations as full members of which only 16 (23%) are certified as compliant with the HAP standards. If the six largest humanitarian NGOs as identified by (Walker & Maxwell 2009, p. 120) are considered; five are members of HAP but none are certified. This suggests that these large international NGOs are more interested in being associated with the concept of accountability to beneficiaries than they are with adopting this accountability in a transparent manner.

In the interviews conducted for this research, participants were questioned regarding the level of participation of beneficiaries in the process of evaluating humanitarian actions. These findings are discussed in Chapter 7.

Host Governments

There appears to be little consideration in the literature of the effect on humanitarian action of host nations taking the control implied by the United Nations General Assembly (1991)

> *"Each State has the responsibility first and foremost to take care of the victims of natural disasters and other emergencies occurring on its territory. Hence, the affected State has the primary role in the initiation, organization, coordination, and implementation of humanitarian assistance within its territory."* (Paragraph 4)

This concept is repeated in NGO publications such as Norwegian Refugee Council (2008)

> *"States, in line with the obligations and responsibilities of sovereign bodies, are responsible for providing protection and humanitarian assistance to internally displaced nationals and refugees within their territories – including those living in camps or camp-like settings"* (p. 22)

Where states exercise this responsibility, particularly the initiation, organisation and coordination of activities undertaken by NGOs, bringing them under the political control of the host government, resentment and frustration are expressed by the NGOs that their independence is being compromised. It may be expected that host governments will take firmer control of relief operations in their territories in the future as the resources available to these operations become larger and the number of international organisations involved becomes greater.

Universality

Barnett and Weiss (2011, pp. 124-127)   question *"whether humanitarianism is universal"* They state that *"Humanitarianism assumes its universality and claims to operate under the principles of humanity, impartiality, neutrality and independence"* (p. 124). This attitude they contrast with the Feinstein Center (2004)

> *"Many in the South do not recognise what the international community calls the universality of humanitarian values… Humanitarian action is viewed as the latest in a series of impositions of alien values, practices and lifestyles."* (p. 55)

Conclusion

Slim (2003) is emphatic, that the politicisation of humanitarianism is self-evident.

> *"Humanitarianism is always politicized somehow. It is a political project in a political world. Its mission is a political one – to restrain and ameliorate the use of organised violence in human relations and to engage with power in order to do so. Powers that are either sympathetic or unsympathetic to humanitarian action in war always have an interest in shaping it their way."* (p.1)

It is this "shaping" by structure that allows humanitarianism to have a level of operational-agency and through the compromises necessary, that humanitarianism is instrumentalised. The researcher would only question why Slim limits his comment to *"humanitarian action in war"?*   The agency/structure relationship

pervades all humanitarian action; war may simply make the humanitarian position weaker and the necessary compromises larger.

O'Dempsey and Munslow (2008) highlight the extent to which humanitarianism has been instrumentalised

> *"Objective humanitarian need, on the basis of the extent of human suffering, is never the criteria used to determine whether or not the international community will act. Power and money determine the workings of international politics, tempered to a variable degree, by the humanitarian impulse, which is influenced by the media and electoral sentiment."* (p. 466)

Humanitarianism has been instrumentalised in an interlocking web of interests; donor government politics and military diplomacy, the commercial, professional and personal interests of humanitarian organisations and the political manipulations of host states. As these structures have solidified they have established a pattern of humanitarian operation that discourages and limits change. Can the evaluation of humanitarian be expected to bring about significant change in such a constrained environment? To what extent can the evaluation of humanitarian action claim to be immune from instrumentalisation itself?

This research looks at some of the influences which act on such evaluations. The methodology by which this has been undertaken is the subject of the following chapter.

Chapter 3 – Design and Methods

As indicated in Chapter 1 the starting points of this research were an observation made by Ramalingam, Scriven and Foley (2009) as they sought to focus the attention of the humanitarian community on innovation rather than incremental learning

> *Critics of humanitarian aid, many from within the sector, complain that humanitarian evaluations and other learning exercises repeatedly highlight the same problems and shortcomings, and 'tell us nothing new'.*(p. 2)

> *"One of the most frequent comments heard at ALNAP meetings and discussions is that 'evaluations do not tell us anything new'. The implication seems to be that, although there is an expectation that evaluations will contribute to improving humanitarian action, they seldom deliver novel and interesting perspectives or solutions to old problems."* (p. 9)

From this starting point these statements raise several questions:

- Can the statements be justified from a review of the recommendations made by evaluations of similar humanitarian emergencies which were separated by a significant period of time?
- If the statements are justified then can reasons why *"evaluations and other learning exercises repeatedly highlight the same problems and shortcomings and tell us nothing new"* be demonstrated?
- Can the sources of the problem be identified: the evaluation process, the learning process or the problems encountered in the delivery of humanitarian action?
- If the sources are identified can recommendations be made for mechanisms that may facilitate the evaluation-learning cycle?

This research aims to provides some answers to these questions.

Research objectives

The methodology devised for this research design was shaped to meet the six objectives defined in the research proposal. These are:

1. To compare the extent to which similar recommendations and "lessons learned" are repeated in evaluations separated in time and to review whether

published evaluation results of work are used in practice to improve the learning capacity of humanitarian organisations through the evaluation process.

2. To document the perceptions of the evaluation process by key participants in the humanitarian system (donors, practitioners, evaluators and researchers) as well as the barriers and facilitation mechanisms that are involved in implementing lessons learned.

3. To explore perceptions held by participants regarding the impact that institutional changes in humanitarian organisations have made on the effectiveness of humanitarian actions.

4. To describe participants' recommendations for mechanisms that may be implemented in the future to improve the institutional learning processes of humanitarian organisations.

5. To analyse the information gathered in interviews in order to understand better the common themes and the relationship between them, as well as to identify both incremental and innovative improvements to the evaluation and learning processes in humanitarian organisations.

6. To recommend actions that may be taken to improve the evaluation/learning process and to apply the institutional knowledge base to improved planning and implementation of humanitarian action

The research seeks to accomplish two different but interrelated tasks. The first objective seeks to establish the validity of the serious criticism, identified by Ramalingam, Scriven and Foley (2009, p. 2 & 9), of the evaluation and lesson learning systems currently utilised in the field of humanitarian action.

The second task encompasses objectives 2 to 6 inclusive which seek to identify currently held perceptions regarding the evaluation and learning process and document the suggestions and recommendation that those involved in the process have for its improvement.

Two points may be noted in the realisation of these two tasks

- Elements of the second task are dependent on the outcome of the first.

- Quite different methodologies were used for each of the two parts of the research.

The implementation of the evaluation of humanitarian action and the identification of subjective influences

Guidelines, specific, to the methodology to be used in evaluation of humanitarian action have existed for a considerable period of time. See: Development Assistance Committee (1991) Hallam (1998) and Development Assistance Committee (1999) These have been supplemented by a large number of agency specific guidelines, such as those for International Federation of Red Cross and Red Crescent Societies (2011a) and United Nations High Commissioner for Refugees (2010). Rather than seeking to focus on these methodological systems for evaluation, or on the, often organisation specific, evaluation policies, this research looks at the experiences of the use of these methodologies and policies. This specifically included the perceptions that the use of these methodologies had left with those who used them, or whose work had been the subjects of evaluation. The researcher also sought to identify both the effect these perceptions had on the evaluation process and on the utilisation of the recommendations that emerged from evaluations to improve the quality of humanitarian action. The researcher defined the policy and methodology as the objective influences in the evaluation and influences not controlled by these as subjective influences.

Epistemology

Part 1 of the research is undertaken by the comparison of texts without questioning the accuracy of the texts themselves or the thoroughness of the work undertaken by their authors. The objective is to compare the recommendations made by evaluators to similar humanitarian actions.

Part 2 of the research is phenomenological study of the perceptions of practitioners regarding the influences that limit the objectivity of the evaluation of humanitarian action. As defined by Smith, Larkin and Flowers (2009) the participants are *"persons who are able and willing to offer us a view of the phenomena under investigation" (p. 40)* and the object of the research is *"Exploring, describing, interpreting and situating the means by which our participants make sense of their experiences." (p. 40)*

Methodology for Part 1 of the research

The methodology used for the first objective was:

- to compare the findings and recommendations from some of the larger evaluations undertaken into the initial response to major humanitarian emergencies spanning a period from 1994 to 2010

- to examine the extent to which similar problems were repeatedly encountered and similar recommendations made.

It should be noted, however, that these objectives were aimed solely at gaining an understanding of the extent to which "*evaluations do not tell us anything new*" (Ramalingam, Scriven & Foley 2009, p. 2 & 9) and that a systematic and exhaustive study of all recommendations in all evaluations was neither intended nor undertaken.

Two pairs of similar humanitarian responses were identified. The first pair of responses were made to two forced migrations, the second pair or responses were made to two natural disasters. These pairs of humanitarian responses were chosen using similar criteria:

- The earlier humanitarian response in each pair had been the subject of a large multiagency evaluation which included a well-structured and comprehensive set of recommendations.
- The later humanitarian response in each pair was chosen to have circumstances as similar as possible to the earlier one.
- The time between the earlier and later responses was at least five years; a period it was considered would have allowed recommendations made in the earlier response to have been incorporated into best practice for the design and implementation of the later response.

To meet these criteria the pair of responses to forced displacement chosen were the mass displacements

- as a result of the genocide in Rwanda 1994
- as a result of the 'genocidal conflict' in Darfur 2003

These two emergencies, despite having many differences, shared an important number of common aspects. Both were large displacements resulting from genocide or "genocidal conflicts" in Africa. Both caused displacement into areas in the centre of Africa where there were few resources available to provide the necessary assistance. Both displacements were to areas far from ports and to which logistic supply routes were difficult. Both receiving countries Zaire (DRC) and Chad had weak governmental structures in the reception areas and neither local

66

government officials nor law enforcement bodies were able to cope with the demands placed upon them. The humanitarian response to the displacement from Rwanda was the subject of a major interagency evaluation. (Borton, Brusset & Hallam 1996) The inceptions of the two humanitarian responses were separated by approximately nine years.

The two responses to natural disasters chosen were:

- Indian Ocean tsunami 2004
- Haitian earthquake 2010

These two disasters similarly had important differences and shared an important number of similarities. Although the source of the disaster in both cases was an earthquake the damage was caused in two different ways, a tsunami in Asia as compared to direct earthquake damage in Haiti. As a result, the humanitarian crisis in Asia was spread widely across a relatively large region and involved several countries whereas in the second the crisis was contained in a limited area of Haiti. Aspects that the crises shared in common were rapid massive loss of life and the destruction of infrastructure, upon which those who survived the initial impact and those who would provide assistance, would have to depend. Both disasters rapidly developed a high public profile and attracted massive funding although, once again for different reasons. In South Asia, it was the death of hundreds of nationals of large donor countries who were vacationing in the affected area that produced the high media profile and prompted both public and private donor response. In Haiti, it was the proximity to the United States and the fear of both the public and the government, that unless a massive and effective aid effort was put into place very rapidly massive migration towards the United States would take place. The humanitarian response to the Indian Ocean tsunami was the subject of a large interagency evaluation of five volumes and synthesis report. (Telford, Cosgrave & Houghton 2006) These two disasters were separated by about five years.

Methodological limitations

The use of such a comparative methodology has to take into account one of the most basic dichotomies encountered in the delivery of humanitarian action, that between standardisation and contextualisation. In one respect the basic needs for humans to survive the aftermath of a disaster are universal and can relatively easily be quantified as has been done in The Sphere Project (2011) and in numerous

other guidelines and handbooks, for example, United Nations High Commissioner for Refugees (2007). The Sphere Project (2011) defines standards as:

"evidence-based and represent sector-wide consensus on best practice in humanitarian response… The minimum standards describe conditions that must be achieved in any humanitarian response in order for disaster-affected populations to survive and recover in stable conditions and with dignity. (p. 4)

The universal application of such standards, were it possible, would render comparison a simple matter of measuring compliance with appropriate standards. However, the world of humanitarian action is not that simple. Although physiologically, especially when treated as the sum of relatively large groups, the basic survival and emergency needs of a human population may be standardised in this way, the delivery of humanitarian action takes place in a wide variety of geographical, cultural, economic, social and political contexts. As a result the standards must be contextualised, as is recognised in the The Sphere Project (2011)

"Conforming with Sphere does not mean meeting all the standards and indicators. The degree to which agencies can meet standards will depend on a range of factors, some of which are outside their control. Sometimes difficulties of access to the affected population, lack of cooperation from the authorities or severe insecurity make standards impossible to meet".   (p. 8)

Recognising these constraints to the methodology, the researcher has been careful to select situations which are as comparable as possible and to contextualise the recommendations made in the evaluations, both in regard to the humanitarian situation to which they apply and the changes in humanitarian emphasis that have taken place over time.

The comparison process

Comparison of the recommendations made in the major evaluations of the first pair of emergency responses was fairly straightforward. The evaluations chosen were those undertaken by Borton, Brusset and Hallam (1996) on Rwanda and by Bartsch and Belgacem (2004) on Eastern Chad. The former evaluation on humanitarian aid to victims of the conflict in Rwanda, was seen as one of the early landmark evaluations. The choice of the latter evaluation may be criticised in that it was undertaken by a single agency (UNHCR), utilising one of its own staff and is being

compared with a joint evaluation undertaken by a team of independent consultants. Bartsch and Belgacem (2004) is, however, highly critical and constructive covering wide aspects of the humanitarian response and despite being undertaken by the organisation which was coordinating the response to the refugee crisis. The transparency of the evaluation process was demonstrated by the report being rapidly placed in the public domain.

For the comparison of responses to the 2004 Indian Ocean tsunami the evaluation undertaken by the Tsunami Evaluation Coalition (TEC) were used. Of the multiple volumes of this evaluation the synthesis report (Telford, Cosgrave & Houghton 2006) and the consolidation of recommendations from all the reports (Tsunami Evaluaton Coalition 2007) were used as the main source of material for the comparison study.

The selection of material for the response to the earthquake in Haiti proved to be much more difficult. When this research began in April 2010, just four months after the earthquake had taken place, no evaluations were available. The initial strategy was to wait for an evaluation or evaluations comparable to those of the TEC and to use these as the primary data source. By September 2012, however, a review of the ALNAP resources on Haiti revealed 103 entries, the majority of which were evaluations of one, or at best a few organisations, or evaluations commissioned by a specific donor regarding projects it had funded. While in the longer term a larger joint evaluation of the whole emergency response may well appear, at the time of writing it was necessary to base the comparison on a selection of the evaluations available. An initial selection of six evaluations was made with the objective of having the best overview of the whole response rather than of individual projects or funding sources. The six selected were:

1. Binder and Grünewald (2010) *"assesses the operational effectiveness and the main outcomes of the cluster approach, as well as its interactions with other pillars of humanitarian reform."* (p. 7)

2. Grünewald, Binder and Georges (2010) *"analyses the humanitarian response following the Haiti earthquake between January and April 2010. In order to identify what worked, what did not and why? It presents key findings, recommendations and lessons learned for immediate implementation in Haiti and for future large-scale disasters."* (p. 8)

3. Interagency Standing Committee (2010) *"describes the response of the humanitarian community to the earthquake, outlining the main achievements*

and challenges encountered, proposes lessons which can be learned from the initial phase of the humanitarian response, and summarizes some aspects of the way forward." (p. 1)

4. DARA[5] (2010) views the initial response within the context of other emergencies

5. DARA (2011) extends the view of the 2010 report into the longer term.

6. Patrick (2011), commissioned by the Haiti Evaluation Task Force, this brief paper *"highlights emerging lessons"* and *"will feed into a more detailed synthesis work to take place at a later stage".* (p. 1)

The Joint Evaluation of Emergency Assistance to Rwanda (JEEAR) (Borton, Brusset & Hallam 1996) provided a framework of 11 categories (specifically tailored to the situation being evaluated) which were used for an initial classification of the findings and recommendations of evaluations into the later humanitarian responses with new categories being developed where necessary. Direct comparison of the 11 categories across all the four evaluations was not possible, however the recommendations of the TEC reports into the Indian Ocean Tsunami do to some extent follow those of the JEEAR. The categories used in this research are shown in table 3.1

---

[5] DARA is an independent non-profit organisation committed to improving the quality and effectiveness of humanitarian action for vulnerable populations affected by armed conflict and natural disasters. DARA is the full title of the organisation and not an acronym.

Table 3.1 Categories used for comparison of recommendations from evaluations

JEEAR Used for comparison of humanitarian response in Eastern DRC with that in Eastern Chad		TEC Used for comparison of humanitarian response to the Indian Ocean Tsunami with that to the Haiti earthquake	
1	Lack of policy coherence	1	Policy Coherence
2	Donor funding and preparedness issues	2	Funding
3	Humanitarian early warning and contingency planning		
4	Coordination	3	Coordination
5	Stand-by capacity and the role of the military forces in humanitarian operations	4	Other actors (other than humanitarian – includes both military and media)
6	NGO performance		
7	Improving accountability	5	Accountability
8	Improving camp security		
9	Food issues and registration		
10	The role of the media		
11	Mitigating the impact on host communities	6	Linking Relief, Rehabilitation and Development (LRRD), Disaster Risk Reduction and Early warning
		7	Ownership

Source: Author after JEERA and TEC

Methodology for Part 2 of the research

The methodology developed for the second objective combined examples of subjectivity reported in the literature with information obtained from interviews with humanitarian workers who had participated in evaluations, i.e. commissioners of evaluations, evaluators and those whose humanitarian work had been evaluated.

In order to investigate the subjective influences the basic methodology utilised was that of qualitative research which, as Silverman (2005) indicates, gives "*access to the nitty-gritty reality of everyday life viewed through a new analytic lens*". (p. 171)

While the evaluation of humanitarian projects is certainly not the "everyday life" of a large segment of the general population, it does feature widely amongst practitioners of humanitarian action. In accord with this observation, this research was pursued through a review of published accounts of undertaking evaluations and interviews with practitioners from a number of different organisations which specialised in humanitarian action.

Published accounts of participation in the evaluation of humanitarian action are rare but the accounts presented by Wood, Apthorpe, et al. (2001b) provided a remarkable insight into the experiences of evaluators working on the evaluation of nine humanitarian responses over the period 1993 to 2000. This data was used to *"kick-start"* the data analysis, as proposed by Silverman (2005, p. 150). These accounts have been treated as non-technical literature and have been used, as recommended by Corbin and Strauss (2008), *as "primary data to supplement interviews"* and to *"make comparisons, enhance sensitivity, and provide questions for interviews"* as well as to *"stimulate questions during the analysis."* (pp. 39 & 37)

From the accounts of undertaking evaluations of humanitarian actions between 1993 and 2000 published by Wood, Apthorpe and Borton (2001b), the researcher identified 14 subjective influences. (See Chapter 6) While these 14 subjective influences are derived from the documented experiences, they have not previously been identified and categorised in the manner undertaken in this research. In the interviews subsequently undertaken with participants in the research these influences were tested against the participants' own current experiences along with additional subjective influences identified as the interviews progressed.

The part of the research involving interviews with participants was in many ways similar to that carried out by Schwartz et al. (2010) where the perception and effect of ethical dilemmas on humanitarian workers was investigated. As a result the methodology used by Schwartz et al. (2010) has been taken as a starting point for this research and adapted as necessary and appropriate. The similarities are in some ways obvious, both sets of participants were involved in humanitarian action and both faced frustrating situations they felt powerless to change. Some of the differences are similarly straightforward. Schwartz et al. (2010) focussed on participants from a single professional group, health care professionals, and documented the participants' experiences and the effect these experiences had on them. This research sought to look at the experiences of a wider group of humanitarian workers, all of whom had been involved in the evaluation-learning

process as applied to humanitarian action, with the objective of identifying from them not only the problems they faced but their suggestions for improvement.

Recruitment

Research ethics committee approval was obtained through the Liverpool School of Tropical Medicine Research Ethics Committee before recruitment was started.

The research was to be carried out in conjunction with several humanitarian agencies, which were willing for the researcher to interview a number of their staff who had been involved in one way or another with evaluations. A Summary Research Proposal was prepared and sent to prospective partner organisations from each of the categories: donors, international agencies, large NGOs, small NGOs, as well as a number of independent evaluators. Organisations that responded favourably were asked to circulate the Summary Research Proposal amongst their staff requesting volunteers to participate in the research project. The objective was to recruit several participants from each of about five organisations.

Using the estimate given by Britten (1999) that "*each hour of interview can take six or seven hours to transcribe, depending on the quality*" (p. 17) of the recording and his observation that "*Large qualitative studies do not often interview more than 50 or 60 people*" (p.18) and heeding the warning of Gillham (2005) *that "It is easy to construct unrealistic expectations of what you can hope to achieve*" (p. 4) an initial target sample size of 20 to 25 interviews was decided upon.

Volunteers were asked to express their interest by sending an email to the researcher including some basic information regarding their experience in participating in the evaluation procedure, from the commissioning of evaluations to the dissemination and implementation of their recommendations.

All volunteers received a response thanking them for their interest and those selected for interview were sent a consent form (See Annex 2) as well as a self-definition questionnaire (See Annex 3) which collected more detailed data on their organisational experience and current role. The objective was to recruit participants from a wide range of roles within their organisation with the common factor that they had experience of participation in the evaluation/institutional learning functions. Gillham (2005) sums up this approach

> *"To use the language of survey sampling, one may seek informants who come from different 'strata' within the group – in terms of status, occupational category, or degree of experience, for example. But this is more to do with trawling for a range of information than trying to establish a representative sample. People placed differently within a given setting can provide different information, and can suggest different ways of understanding it."* (p. 43)

Volunteers who returned the consent form and completed the questionnaire were enrolled as participants. Where too few volunteers were obtained from a partner organisation a snowball approach was used, volunteers being asked to identify others in the organisation who would be willing to participate.

The criteria for selection of participants consisted of humanitarian workers who had experience of working with the evaluation/institutional learning processes. While all participants may be considered as 'specialists', preference was given to those with the most direct experience with these processes and also to the more senior members of the organisation. The latter criterion was justified as there are far fewer senior staff than the more junior and the opportunity to obtain an "elite" interview with at least one senior staff member in each organisation would make an important contribution to the research.

Data Collection

The basic interview technique used was semi structured using open ended questions which were refined by trialling with colleagues and then piloted with a small group of volunteers who met the selection criteria but who were not later recruited as participants.

The interviews were designed to meet the objectives listed by Britten (1999)

- *"Be interactive and sensitive to the language and concepts used by the interviewee*
- *Keep the agenda flexible*
- *Go below the surface of the topic being discussed*
- *Explore what people say in as much detail as possible*
- *Uncover new areas or ideas that were not anticipated at the outset of the research"* (p. 13)

The questions themselves were written to meet the criteria specified by Pope *"good questions in qualitative interviews should be open-ended, neutral, sensitive and clear to the interviewee"* (see Britten 1999, p. 14)

Before the interview participants were asked to think of instances when they had been involved in the evaluation/institutional learning processes and felt that subjective influences were impinging on the objectivity of the process and to, if possible, identify the nature of the subjective influence and the effect it had on the process. The definition of subjective was left to the participant so as not to limit the range of responses. Where participants at the beginning of the interview were ready to talk about these experiences of evaluation/institutional learning, a more unstructured approach was taken in order to allow the participant the maximum freedom of expression. In these cases the researcher used the questions and prompts devised for the semi structured interviews as a checklist (See Annex 4) to ensure that relevant points were covered. Less forthcoming participants were encouraged, as necessary, by the minimum use of the open ended, pre-planned, questions. Participants were prompted to discuss their experience and/or perception of the 14 subjective influences (as interviews progressed one additional influence was identified) with questions introduced with phrases like, 'It has been reported that' or 'Some people have said that'. These questions encouraged participants to either agree with or argue against the suggested influence, their responses being used as a means of testing how widely each of the influences was perceived. Participants were also asked how improvements could be made and for their reflections on possible improvements.

The structure of the interviews was dynamic, with later interviews used to validate some of the previous information received: "*Validating here refers to a checking out of interpretations with participants as the research moves along."* (Corbin & Strauss 2008, p. 48)

Where possible, interviews were conducted face to face (5 Interviews) but most were conducted over the telephone or by Skype (17 Interviews). While face to face interviews do have a greater level of intimacy which builds confidence, telephone has allowed interviews with participants from Switzerland, Denmark, France and the Netherlands, as well as those widely dispersed within the UK. The researcher observed that telephone interviews with participants who did not directly know the researcher tended to take longer to "warm up" than interviews undertaken face to face. However, once this initial period had been overcome participants appeared to

be more relaxed and spontaneous than was the case when face to face with the researcher and the recorder on the table in front of them.

The maintenance of recorded sound quality to enable easy and accurate transcription was initially a problem. Whereas the elimination of video from Skype calls did result in a significant improvement, it was generally found that Skype could only be really satisfactorily used when the use of open computer or desk based microphones was eliminated. Best results were obtained when the researcher used a computer connected telephone type handset and Skype was used to call the participant on a landline based telephone. All interviews were professionally transcribed into Microsoft Word documents.

Analysis of data

This research has been carried out using the principles of interactionism (Silverman 1993, pp. 94-95) and *"has preferred open ended interviews"* for the reasons given by Denzin (1970) :

1. *"It allows respondents to use their 'unique ways of defining the world'*
2. *It assumes that no fixed sequence of questions is suitable to all respondents.*
3. *It allows respondents to 'raise important issues not contained in the schedule"* (p. 125)

It is recognised that

> "*Where the researcher maintains a minimal presence, asking few questions, this can create an interpretive problem for the interviewee about what is relevant. Moreover the passivity of the interviewer can create an extremely powerful constraint on the interviewee to talk"* (Hammersley & Atkinson 1983, pp. 110-111)

This technique was deliberately used to provoke the widest possible response from the participants.

The Computer Assisted Analysis of Qualitative Data (CAQDAS) program Nvivo was used in the manner described by Lee and Fielding (1995) where they state that *"Most researchers who use CAQDAS confine their use to coding and retrieval of text segments, using the computer as an electronic filing cabinet"*. (p. 29) Silverman

(2005) points out that while *"CAQDAS does support theorizing…Theory building is generally done in the mind, or with the aid of paper, if at all"* (p. 202) and in justification continues

> *"Although in one sense all research studies, indeed all observations, are 'theory driven', not all research studied need to be explicitly 'theorised'. Much qualitative research is, in common sense parlance, 'descriptive' and does not require the explicit elaboration of conceptual thought generally referred to as theory".* (p.202)

The objective of this research is not to develop theory but to use the information gathered from participants to categorise experiences and feelings and from these to develop recommendations for improvements to systems currently in use. Throughout the analysis of information obtained in this research, the researcher used his personal experience in the manner presented by Corbin and Strauss (2008)

> *"We can use our experience not as data per se, but as a comparative case to stimulate thinking about the various properties and dimensions of concepts … to "bring up other possibilities of meaning… to offer a negative case or something new to think about that will make us confront our assumptions about specific data."* (p.80)

The transcribed interviews were loaded into the CAQDAS software "Nvivo" and simultaneously checked against the original sound file while being coded against pre-established and emerging nodes. Before this process began an initial 14 nodes had been created in Nvivo corresponding to the 14 subjective influences previously identified. Following the methodology suggested by Corbin and Strauss (2008) to *"begin with a list of codes derived from the literature and then revise the codes as the codes are compared against actual data"* (p. 67) Additional nodes were created as necessary as the interview coding process was undertaken. As the definition of subjective influences had initially been left to the participant and the researcher had deliberately limited his interruptions to the flow of information being given by the participant, the interviews contained material relating to issues outside the scope of the questions asked, hence the coding was expanded to keep the maximum amount of information readily available for further analysis. The discussion of suggested solutions frequently involved discussion of very objective processes that may be used to improve the evaluation/learning processes and these were also carefully coded.

When all the interviews had been coded in the manner indicated above, each node was exported as a Microsoft Word document, printed out and marked up linking similar ideas expressed by different participants. These were grouped, using a cut and paste technique, into those confirming or opposing the relevance of the node topic. Selections from each of these documents were used to establish and illustrate the results of the research. This approach was adopted as an effective compromise between the physically complex and labour intensive direct marking up of print outs of interview transcriptions and the large investment in time required to become fluent in the features of a complex CAQDAS computer program.

In this research the word 'participant' is used to indicate subjects who were interviewed for the purposes of this research. To give a measure of the frequency by which participants discussed the node topics each main heading used in Chapter 7 contains a reference to the number of interviews in which quotations for the corresponding node topic were coded. These are in the form $x/22 - y\%$ where x is the number of participants coded against the topic and y is this fraction expressed as a percentage. As the interviews were relatively unstructured and some had to be terminated due to other commitments of the participant, not all interviews covered all the node topics. The use of these fractions and percentage figures may be indicative of the participants' interests but should not be regarded as in anyway statistically significant.

In the discussion and analysis of the results of the research (Chapter 8), the analytical structure developed in the literature review of the instrumentalisation of humanitarian action (Chapter 2) is used to demonstrate the manner in which each of the subjective influences identified serves to instrumentalise the evaluation of humanitarian action. The analytical structure is based on a refinement of the agency structure model, identifying two components of the agency of humanitarian organisations, that derived from their principles (principle-agency) and that derived from resources and access made available to them (operational-agency). In order to maximise their effectiveness organisations compromise their principle-agency in order to maximise their operational-agency and are in the process instrumentalised by the providers of resources and access.

Aspects of the evaluation process that involve contractual obligations and financial rewards are also analysed in the context of the principal agent problem as stated in Financial Times (2013).

Problems encountered in the implementation of the methodology and solutions adopted.

Recruitment

The recruitment process detailed above unfortunately did not provide the sample required for the research. Organisations were contacted from each of the groups proposed and agreement obtained from these organisations for the researcher to contact key members of staff who would act as focal points for the research and assist in the identification of participants. In most cases these were staff whose primary role was the management or conducting of evaluations. These contacts, in each case, participated and gave interviews. What proved far more difficult, and in some cases impossible, was to build a snowball of participants in the same organisation. Partly this may have been a result of the researcher working by email and telephone and partly due to the role of the primary contact person. Where other participants did volunteer, interviews often uncovered levels of frustration but little analytical thinking regarding what lay behind that frustration. The result of this initial group of interviews being from participants largely working in roles that directly involved them in implementing evaluation policy, together with a few of those who felt themselves somehow victims of the evaluation process, was that saturation was approached at an unsatisfactorily small sample.

To increase the sample size and diversify the recruitment, the model was broadened to include a larger than planned group of individuals who had extensive experience of the evaluation of humanitarian action with a wide range of organisations, i.e. consultants who specialised in undertaking evaluations. This part of the sample included some participants who had published their accounts of undertaking evaluations between 1993 and 2000 (Wood, Apthorpe & Borton 2001b) which are used as "baseline data" in this research. Some of these participants were still carrying out evaluations when interviewed in 2012. In these interviews no reference to their earlier publication was made by the researcher, although several of the participants did make reference to the earlier work.

The necessity for this change in recruitment strategy, and indeed the nature of the change itself, may be the subject for some criticism of the methodology. On one hand, the initial approach may appear naïve in its expectation of a high degree of openness of organisations to researcher access to staff, as well as the willingness

and ability of a range of staff to participate. Only two organisations invited the researcher to meet staff at their offices. The headquarters of humanitarian organisations are, however, busy places with staff under both workload and organisational political pressures and these factors may well have contributed to the way the situation played out. Ideally, following perceptions held in a single organisation regarding that organisation's own evaluation and lesson learning process would have been an excellent approach to this research topic, but unfortunately it was not possible to realise. Had this not been the first approach tried, despite the difficulties, then there would have been justifiable reason for criticism on the grounds that there was a better way that should have been attempted. As Gillham (2005) points out, "*The 'real-world' researcher is constantly having to adapt or compromise on methods because of the constraints encountered.* (p. 4) The final sample contained 13 staff from eight organisations and nine consultants, 22 in total.

Participant Observation

Participant observation is defined by Macionis and Plummer (2005) as *"a method by which researchers systematically observe people while joining their routine activities"*. (p. 56) While conducting the research the researcher must balance *"Playing the participant"* which *"gains for the researcher acceptance and access to people's lives; yet playing the observer"* which *"affords distance and perspective needed for thoughtful analysis"*. (p. 57)

Plummer (2005) identifies *"two main types of participant observation; overt and covert"* which may be summarised as*:*

- In covert observation the researcher is accepted by the group as a member but not recognised as a researcher.
- In overt observation the researcher is accepted by the group as a researcher but less as a member of the group.

Insider research provides a third option in which the researcher is accepted as a full member of the group and at the same time can overtly conduct research without jeopardising acceptance as a member of the group.

<u>Data Collection</u>

The researcher recognised that from his three decades of experience in the field of humanitarian action there was a strong likelihood that he would be known, either directly or by hearsay, to at least some of his informants. The researcher has considered the implications of undertaking such insider research which has been described by Rabbitt (2003) who observes that

> "*Being familiar with the local culture and customs and having already established a relationship provides the opportunity for the researcher to gain participants easily and to be privy to 'insider' information that would not be trusted to a stranger. Yet being known has its shortcomings. Prior knowledge, underlying personal bias and preconceived ideas can render disadvantages to this intimate type of 'insider research'.*" (p. 1)

On balance the researcher considered that the advantages, for this particular research topic, would outweigh the disadvantages. Detailed knowledge of humanitarian action and experience of being both an evaluator and the subject of evaluation in this particular field of work has been a definite advantage in understanding both the substance and context of participants' experience as well as being sensitive to many of the nuances expressed during interviews. Whether access has been facilitated by the researcher being recognised as "an insider" is much more open to question. Organisationally the researcher's access may have actually been limited by his close identification for over two decades with one large UN organisation closely involved in humanitarian action. In the highly competitive and politicized world of humanitarian organisations such large organisations are often grudgingly respected but rarely liked or completely trusted. While relatively senior staff with considerable experience, felt confident enough to participate in academic research conducted by a retired insider, such confidence may not easily be felt by younger more junior staff. Consultants, however, several of whom had a similar length of experience as the researcher, were more accessible and ready for a 'senior professional to senior professional' interaction.

Participants have, in the researcher's opinion, been more willing to discuss their experiences of the evaluation process frankly than they may have been with an "outsider". Experience of situations similar to those being narrated has been an advantage and care has been taken to avoid inclusion of any data referring to situations in which the researcher has actually been involved where it was recognised that prior knowledge, underlying personal bias and preconceived ideas

may have hindered objective analysis. Overall it is concluded that while insider research may have reduced the range of the participant sample, it probably improved the quality of the information gained from those who did participate.

<u>Analysis of the data</u>

Experience in the case of this research has been something of a two edged sword. On one hand experience facilitates a familiarity with the systems with which, and the context in which, the participants are working and so gives an understanding and empathy with the experiences being described. On the other hand experience also brings awareness of the resistance to change and the political complexity of the humanitarian environment. This knowledge has a tendency to inhibit the formulation of recommendations that are 'outside the box' of current trends. As Kirk and Miller (1986) put it *"There is a world of empirical reality out there. The way we perceive it and understand that world is largely up to us, but the world does not tolerate all understandings of it equally".* (p. 11)

Chapter 4 - A comparison of the recommendations of evaluations of the humanitarian response to the forced displacements from Rwanda (1994) and from Darfur (2003)

In this chapter the recommendations of the third volume of the Joint Evaluation of Emergency Assistance to Rwanda (JEEAR), "Humanitarian Aid and Effects" (Borton, Brusset & Hallam 1996), will be taken as a baseline against which the recommendations of an evaluation of the humanitarian action taken nine years later in the displacement of Sudanese from Darfur to Chad are compared.

Situation overview

Forced displacement from Rwanda

According to Sellstrom and Wohlgemuth (1996) the ethnic polarisation of Rwanda between the majority Hutu population and the minority Tutsis can be traced back to pre-colonial times. The polarisation of ethnicity and the power of the Tutsi minority was strengthened under German and Belgian colonial rule and cemented in 1933 with the introduction of Identity cards which specified ethnicity. However in 1950 there was a rapid change when the Belgian colonial regime, in preparation for decolonialisation (independence was achieved in 1960), shifted support from the Tutsi to the Hutu majority. This was followed by the Hutu revolution of 1959-61 and the change to an independent Hutu led Republic, all in less than 3 years. One result was the displacement of tens of thousands of Tutsis into neighbouring countries, particularly Uganda, where the close relationship between Rwandan Tutsi elements and the Ugandan army of President Yoweri Museveni, led to the formation of the Rwandese Patriotic Front (RPF), and the formation of its military wing, the Rwandese Patriotic Army (RPA) in 1988.

Following the death of Rwandan President Habyarimana, when the plane in which he was travelling was shot down approaching Kigali airport on 6 April 1994, powerful elements of the Hutu population were mobilised to perpetrate the genocide of the remaining Tutsi population and moderate Hutus. Simultaneously a major offensive from the Tutsi led RPA took place and on the 18 July 1994, with Kigali taken, the RPA declared the war to be over.

"Within a period of three months in 1994, an estimated five to eight hundred thousand people were killed as a result of civil war and genocide in Rwanda. Large numbers were physically and psychologically afflicted for life through maiming, rape and other trauma; over two million fled to neighbouring countries and maybe half as many became internally displaced within Rwanda." (Sellstrom & Wohlgemuth 1996, p. 5)

As a result of the genocide and the war a massive displacement of refugees took place. The first refugee movement was to Burundi where 270,000 fled in April 1994. Larger movements took place into Tanzania where 580,000 arrived in April-May and into Zaire where the largest group of 1,200,000 arrived in July. As the war ended with a Tutsi led victory the group arriving in July 1994, in the North Kivu province of Zaire, contained many of the elements who, a few weeks earlier had perpetrated the genocide in Rwanda. These elements took political control of the refugee community, refusing to allow them to return to Rwanda and greatly complicating the protection of refugees and the delivery of humanitarian assistance

Forced Displacement from Darfur (Sudan) into Chad

While the humanitarian response to the influx of refugees into Chad from Darfur in 2003 has similarities to the response to Rwandan refugees in Zaire (DRC) almost a decade earlier, it was the survivors of genocide rather than a group that included the perpetrators who constituted the refugee population. The British Broadcasting Corporation (2010) reported that *"Darfur, which means land of the Fur, has faced many years of tension over land and grazing rights between the mostly nomadic Arabs, and farmers from the Fur, Massaleet and Zaghawa communities."* And that *"The Sudan Liberation Army (SLA) and Justice and Equality Movement (Jem) began attacking government targets in early 2003, accusing Khartoum of oppressing black Africans in favour of Arabs."* The (British Broadcasting Corporation 2010) further reported that refugees said that government bombing raids *"would be followed by attacks from the Janjaweed, who would ride into villages on horses and camels, slaughtering men, raping women and stealing whatever they could find".*

Weisman (2004) reports United States Secretary of State Colin Powell as saying, *"When we reviewed the evidence compiled by our team, we concluded — I concluded — that genocide has been committed in Darfur and that the government of Sudan and the Janjaweed bear responsibility, and that genocide may still be*

occurring." However, *"a UN investigation team in 2005 concluded that war crimes had been committed but there had been no intent to commit genocide."* (British Broadcasting Corporation 2010) The use of the term "genocide" in connection with the events in Darfur remains highly controversial. (See Prunier 2005)

In its evaluation (Bartsch & Belgacem 2004), UNHCR reports a time line for the emergency beginning in April 2003 when, with the outbreak of fighting in Northern Darfur, the first arrival of refugees from Sudan into Chad began. UNHCR further reports that more than 180,000 Sudanese refugees entered Chad between May 2003 and June 2004. They entered along a 600Km stretch of the Chad/Sudan border and many remained close to this border. |By August 2004, 150,000 refugees had been relocated into eight newly created camp sites. However, UNHCR places the beginning of its emergency operation at the end of December 2003 due to the arrival of 30,000 refugees in that month bringing the total to 95,000, when at the same time incursions of Janjaweed took place into Chad highlighting the need to relocate refugees away from the border. Also at this time *"the first assessment missions into Sudan's Darfur Region painted a very grim picture of a protection crisis inside Darfur, thus raising the spectre of further influxes".* (p. 5)

The evaluations

The Joint Evaluation of Emergency Assistance to Rwanda (JEEAR)

The humanitarian situation in Rwanda, and in the countries to which displacement took place, became so politically far reaching, complex, controversial and overwhelmingly expensive that a unique evaluation, the Joint Evaluation of Emergency Assistance to Rwanda (JEEAR), was commissioned in 1996 and the preface to the first volume (Sellstrom & Wohlgemuth 1996, p. 5) describes the circumstances and mandate as follows:

> *"Recognizing both the magnitude of the Rwanda emergency and the implications of complex disasters for constricted aid budgets, the Danish Ministry of Foreign Affairs, through its development cooperation wing, Danida, proposed a Joint Evaluation of Emergency Assistance to Rwanda.* (p. 5)

> *This initiative resulted in the launching of an unprecedented multinational, multi-donor evaluation effort, with the formation of a Steering Committee at a*

consultative meeting of international agencies and NGOs held in Copenhagen in November 1994. This Committee is composed of representatives from 19 OECD-member bilateral donor agencies, plus the European Union and the Development Assistance Committee (DAC) of the OECD; nine multilateral agencies and UN units; the two components of the International Red Cross and Red Crescent Movement (ICRC and IFRC); and five international NGO organizations" (pp. 5-6)

The objective of the evaluation is presented as follows:

"The main objective of the evaluation is to draw lessons from the Rwanda experience relevant for future complex emergencies as well as for current operations in Rwanda and the region, such as early warning and conflict management, preparation for and provision of emergency assistance, and the transition from relief to rehabilitation and development.

In view of the diversity of the issues to be evaluated, four separate evaluation studies were contracted" (Sellstrom & Wohlgemuth 1996, p. 6)

These four studies were:

1. Historical Perspective(Sellstrom & Wohlgemuth 1996),
2. Early Warning and Conflict Management (Adelman & Suhrke 1996),
3. Humanitarian Aid and Effects (Borton, Brusset & Hallam 1996) and
4. Reconstruction of Rwanda (Kumar et al. 1996).

A Synthesis Report (Eriksson 1996) was also prepared. While the historical background above is condensed from the first study it is the third study which forms the main source for the remainder of this chapter of the research.

The JEEAR was produced at a time when the evaluation of humanitarian projects was still not universally accepted and represents something of a landmark in the field not only for its size and scope but also for establishing evaluation of humanitarian actions as a common practice. Volume 3 Chapter 9 of the JEEAR (Borton, Brusset & Hallam 1996) makes 26 recommendations broken down over 11 topics.

The author of Volume 3, John Borton, has reviewed progress made by the humanitarian community on the recommendations made in the JEEAR on two occasions. On the 10th anniversary of the genocide, in 2004 *"ALNAP and Danida*

commissioned two of the evaluation's original authors, John Borton, leader of Study 3, and John Eriksson, leader of the Synthesis, to review the impact of the evaluation." (Borton & Eriksson 2004, Preface) On the 20th anniversary of the genocide Sphere published an interview with John Borton. (Borton 2014)

Borton and Eriksson (2004, p. 47) observe that 64 books, reports and articles published between 1995 and 2004 deal with issues relevant to the JEEAR and that 37 of these publications directly cite the JEEAR. However, commenting on the limits of the utilisation of the JEEAR and quoting a personal communication from Peter Walker. Borton and Eriksson (2004) recognise:

> *"the JEEAR [Joint Evaluation] quickly entered into the folk law of the humanitarian world, and within it achieved wide recognition. Outside of it, even in closely related fields such as security and development, it got less star treatment. Also, I find now that many people who entered the business post-1998 have not heard of it. This I think has to do with the continued lack of formal institutional memory arrangements. We use the JEEAR in our teaching, but it is now an academic historical document for students, not a living one."*

Borton (2014) recognises the catalytic role played by the JEEAR in the formation of "some of the key humanitarian quality and accountability initiatives" citing the Sphere Project, the HAP and ALNAP as examples. He also acknowledges the role played by the JEEAR in initiating the ongoing debate regarding the establishment of an NGO accreditation system. Commenting on whether the 'accountability revolution' in humanitarian action had realised its potential Borton (2014) said:

> *"It has achieved a lot. But whilst I admire all that has been done, I wonder whether we haven't actually walked around some of the issues that the Rwanda evaluation tried to confront. Compromises have been made and the structures that have been created are less than perfect - that's why I think many of the challenges are still there."*

Borton (2014) cites the humanitarian response to the earthquake in Haiti as an example of where recommendations made in the JEEAR were ignored.

Reference to Borton and Eriksson (2004) is made under a number of headings that follow in this chapter.

Forced Displacement from Darfur (Sudan) into Chad

The UNHCR evaluation (Bartsch & Belgacem 2004) was undertaken "*to determine the operational effectiveness of UNHCR's response to the Chad emergency and to identify lessons learnt of possible relevance to other operations.*" (p. 29) The evaluation consists of a single volume and findings are made against five benchmarks established by UNHCR for the evaluation of its emergency programmes. In this study however, they are regrouped to enable comparison with the findings and recommendations of the JEEAR report.

Structure used for the comparison of recommendations made in the evaluations.

The structure used for the comparison of the recommendations follows the headings used by Borton, Brusset and Hallam (1996) (See table 3.1 page 68).

Borton, Brusset and Hallam (1996) evaluated the humanitarian response delivered to beneficiaries inside Rwanda as well as that delivered to those displaced to Zaire (now the Democratic Republic of the Congo (DRC)) and to Tanzania. As the most complex and dangerous situations developed in Zaire, the analysis of the recommendations focusses more in Zaire than in the other locations.

   a.   **International Political Action (Lack of policy coherence referred to as a "Policy vacuum")**

Forced displacement from Rwanda

Two principle factors are identified by Borton, Brusset and Hallam (1996)

> "*conflicting interests between members of the Security Council and lack of resolve to overcome these differences, probably stemming from little interest in a small African country of marginal strategic importance to the main powers; and a lack of understanding of Rwanda's complex situation and misread signals prior to and immediately after the shooting down of the President's aircraft on 6 April.*" (p. 161)

The evaluation identifies important points:

- that humanitarian action cannot substitute for political, diplomatic or military action,
- that the responsibility for addressing complex emergencies rests primarily with those who control the necessary means
- the role of the Security Council in ensuring that this is done.

The role of the humanitarian community is identified as:

- highlighting the humanitarian consequences of action taken by the Security Council and the member states
- the need for the creation of better communication channels between the humanitarian and other actors.

The two recommendations made under this topic by the evaluation are the establishment of a Humanitarian Sub-Committee under the Security Council and the appointment of a team of senior advisors to address all complex emergencies and to bring policy options to the Security Council and the Secretary General.

Borton and Eriksson (2004, pp. 51-67) comment on these issues in the context of Volume 2 of the evaluation (See Adelman & Suhrke 1996) which fell outside the scope of this research. There is, however, some overlap with Borton, Brusset and Hallam (1996). Borton and Eriksson (2004, pp. 51-67) focus on the developments in UN peace keeping operations and in particular the "Brahimi Report". See (Brahimi 2000)

Borton and Eriksson (2004) recognise the link between the JEEAR and the developing concept of the 'report of the ICISS (See Evans & Sahnoun 2001) they state that:

> "Though it is nearly three years since the International Commission on Intervention and State Sovereignty produced its 'Responsibility to Protect' report, it appears to have had little if any discernible impact in the case of Darfur." (p. 106)

While action at this level is clearly outside the mandate or authority of the humanitarian community at large, the observations and recommendations laid some important guiding principles for the process of the humanitarian reform which was formulated some five years later in General Assembly Resolution 46/182 (United Nations General Assembly 1991) This resolution:

- Establishes the role of the Emergency Relief Coordinator (ERC), (Annex paragraph 34) at the Under Secretary General level, reporting directly to the Secretary General and with access to *"all relevant organisations",* which, practice has shown, includes the Security Council. (Annex paragraph 37)
- Strengthens the position of the ERC by the establishment, under the chairmanship of the ERC, of the Inter-Agency Standing Committee to coordinate policy with regard to the response to humanitarian emergencies (Annex paragraph 38)
- Establishes an office to support the ERC in the tasks assigned under the resolution. (Annex paragraph 36)

Borton and Eriksson (2004, p. 70) indicate the role of UN reform (See United Nations General Assembly 1997) in further developing the effectiveness of these institutions in line with the recommendations made in the JEEAR.

Responsibility to Protect (R2P)

In 2000 the then Secretary General of the United Nations, Kofi Annan, (Annan 2000) asked the *question "If humanitarian intervention is, indeed, an unacceptable assault on sovereignty, how should we respond to a Rwanda, to a Srebrenica – to gross and systematic violations of human rights that affect every precept of our common humanity?".* (p. 48) In response to this question and to criticisms and recommendations following humanitarian disasters, action began to be taken towards the concept of Responsibility to Protect (R2P).

The International Commission on Intervention and State Sovereignty (ICISS) was established under the co-chairmanship of Gareth Evans (ex-foreign Minister of Australia) and Mohamed Sahnoun (senior Algerian diplomat) which produced a report (Evans & Sahnoun 2001) in December 2001. The forward of the report begins with a succinct definition of its purpose

"This report is about the so-called "right of humanitarian intervention": the question of when, if ever, it is appropriate for states to take coercive – and in particular military – action, against another state for the purpose of protecting people at risk in that other state." (p. vii)

The forward continues to cite the Rwanda genocide as a classic failure of the international system and demonstrates the direct link between the events in

Rwanda and the concept of R2P. The ICISS reports goes further than was proposed in the recommendations of the JEEAR insomuch as it questions the absolute nature of sovereignty, a concept of international relations that was recognised from the Peace of Westphalia signed in 1648 to the Charter of the United Nations signed in 1945. As envisaged in the report of the ICISS (Evans & Sahnoun 2001) the most fundamental responsibility of a sovereign state is to protect its population and that

> *"Where a population is suffering serious harm, as a result of internal war, insurgency, repression or state failure, and the state in question is unwilling or unable to halt or avert it, the principle of non-intervention yields to the international responsibility to protect."* (p. xi)

The option of military intervention is, in the ICISS report (Evans & Sahnoun 2001, p. xii) defined within four precautionary principles of, right intention, last resort, proportional means and reasonable prospects. While the "right authority" for intervention is focussed on the Security Council, it is with the restriction that the five permanent members should refrain from the use of the veto and that if the Security Council is unwilling to act then action should remain open to regional bodies or "concerned states". (p. xiii)

The report of the ICISS was considered by the World Summit in 2005 and the deliberations of this body on the subject of R2P are documented in the General Assembly Resolution 60/1 (United Nations General Assembly 2005 Paragraphs 138-140) The World Summit weakened many of the recommendations of the ICISS and defaulted towards the status quo of the existing provisions of the UN Charter.

The changes made to the concept of R2P between the ICISS report and the World Summit Outcomes have been summarised by Labonte (2011)

Table 4.1 – Changes made to the concept of R2P between the ICISS report and the World Summit Outcomes

	ICISS	World Summit Outcomes
What triggers R2P (Just cause)	Large---scale loss of life, widespread ethnic cleansing and related human rights violations	Genocide, war crimes, crimes against humanity, ethnic cleansing
Response trigger	Host state "unwilling or unable"	Host state "manifestly failing"
Legitimate responders	Security Council General Assembly Coalitions of the willing	Security Council
Criteria	Ad jus bellum principles	None – Case by case consideration
Veto	Avoid veto – be prepared to explain	No restrictions on veto use
Obligation to react	Shared by the Security Council and the international community	No such obligation exists. UN member states "stand ready" but no duty is created
Response	Hierarchy from no forcible to forcible measures including unilateral measures	UN Charter Chapters VI,VII and VIII

Source: Labonte (2011)

R2P was brought again to the attention of the General Assembly in January 2009 in the report of the Secretary General on "Implementing the Responsibility to Protect" (United Nations General Assembly 2009 (1)). The Chairman of the General Assembly convened four days of discussion on the topic in July 2009 and the General Assembly adopted a remarkably short resolution (United Nations General Assembly 2009 (2)), which has less than 50 words in the main text, in which the General Assembly merely decides to *"continue its consideration of the responsibility to protect"*.(Paragraph 2)

Recent publications indicate that little enthusiasm also exists for the further development of any early warning and response mechanism. Wulf and Debiel (2009) highlight two of the major problems

> *"The most explicit response mechanism exists with regard to humanitarian emergencies at the UN level; however the early warning and response*

(EWR) mechanism is far from being efficient since the UN is a bureaucratic organisation with a 'silo' mentality among the different agencies and departments, and the UN Security Council is a highly politicised body." (p. 1)

Zenco, Fellow and Friedman (2011) give a very pessimistic analysis of the situation.

"every significant internal review of the UN's role in the maintenance of peace and security has called for the urgent need to develop a comprehensive and coordinated conflict early warning system. Despite the repeated efforts by the Secretariat to promote and implement the findings of these reports, member-state obstructionism has precluded development of a UN-wide early warning system of socio-political crises that could lead to political instability or armed conflict. Indeed, there is presently no UN-wide coordinating mechanism to collect, assess, prioritize and integrate all of the early warning reporting… nor are there plans for a debate on the creation of one." (p. 21)

Forced Displacement from Darfur (Sudan) into Chad

The UNHCR evaluation (Bartsch & Belgacem 2004) makes no reference to the first topic raised in the JEEAR report, that of lack of policy coherence, as it was clearly seen as out of the scope of an evaluation conducted by an individual UN agency. It is obvious, however, that neither the Security Council nor the Africa Union were able to prevent the mass killing in Darfur which caused the displacement into Chad. Two successive Special Representatives of the Secretary General to Sudan were expelled by the Government of Sudan for statements they made on the situation as it developed in Darfur: Mukesh Kapila in April 2004 and Jan Pronk in October 2006. The UN was as ineffective in 2003 in Darfur as it had been in 1994 in Rwanda in protecting a population against extreme violence by the Government or its agents.

b. <u>Donor funding issues (Donor funding and preparedness measures.)</u>

Forced displacement from Rwanda

The three recommendations made in JEEAR under this heading (Borton, Brusset & Hallam 1996, p. 163) are for better donor coordination regarding levels of

investment in contingency planning and preparedness, a greater readiness to provide flexible funding when events are "fast moving" including a larger Central Emergency Revolving (now Relief) Fund, and that, when funds are comparatively readily available they are used wisely, including the maximum use of locally available goods and services and the appropriate minimal use of airlift.

The issue of donor coordination in general began to be addressed as part of the Montreux Process (Montreux Process 2000) and was further developed as Good Humanitarian Donorship in Stockholm in 2003 when 17 major donors agreed 23 principles of good donorship. (Good Humanitarian Donorship 2003) Two of these principles are to enhance preparedness in disaster prone countries and within humanitarian organisations. While these recommendations have clearly met with less resistance than R2P, consideration of evaluations of the more recent major humanitarian crises reviewed in this research show them to be ineffective.

Forced Displacement from Darfur (Sudan) into Chad

The reliance of UNHCR on the donor community for almost the totality of the funding it receives allows only muted criticism of donors couched in diplomatic language. However, despite this limitation criticism is certainly implied. The report states that *"financial issues have dominated the early stages of the operation, with implementing partners (IPs) insisting on rapid disbursements of instalments whereas UNHCR, not having received the required cash, encouraged partners to pre-finance operations with their own funds".*(Bartsch & Belgacem 2004, p. 14) Later in the evaluation (p. 20) it is made clear that the operation did receive funds from the UNHCR Operational Reserve but that this was only able to be a "stop gap" measure and could not be sufficient to fund operations adequately. The critical factor was the unavailability of donor funding at an early stage in the operation which the evaluation attributes to the context of the refugee situation in the larger context of the Darfur crisis as a whole. If access to Darfur were to have become available then larger sums would have been required inside Sudan than in Chad. The challenging by donors of UNHCR's population figures is seen as an intentional downplaying by some donors of the magnitude of the crisis which may have negatively affected other donors. At the time the evaluation was written it is clear

that the situation was continuing with a conclusion that *"the operation has been constrained by the availability of both adequate and timely contributions".* (p. 20)

In summary, while the first two areas of recommendation have looked to the international community (United Nations and the donors) for their implementation the implementation of the remaining recommendations is more directly under the control of the humanitarian actors and so will be a better measure of whether evaluations are indeed showing nothing new.

c.   Humanitarian early warning and contingency planning.

Forced displacement from Rwanda

A single rather long and detailed recommendation is made (Borton, Brusset & Hallam 1996, p. 163), the main points of which may be summarised as:

- The establishment of an early warning cell within the DHA (OCHA) field office to which all agencies working in the humanitarian sector should submit information.
- Where there are gaps in this information the cell should either deploy observers or obtain information from aerial (now satellite) images.
- Contingency plans should be made and regularly updated and information should be widely distributed, if necessary daily, with due care regarding confidentiality.

In its exact formulation this recommendation has to some extent been overtaken by technological developments, particularly Geographic Information System (GIS) mapping and the ease of information sharing through the internet.   These developments should, however, have made the general objective of the recommendation, early warning of developments in an on-going humanitarian situation, easier to achieve. While the wording of the recommendation appears to focus on specific on-going humanitarian situations, the concept of early warning and contingency planning can readily be extended to a global watch on humanitarian risks and in reviewing evaluations that follow the JEEAR.

Borton and Eriksson (2004, p. 75) refers to the failure of UNHCR's preparedness and response capacity in relation to the 1999 movements from Kosovo into Albania and Macedonia and reports on the improvements that UNHCR subsequently made.

As is shown in the following paragraphs, however, UNHCR responded to the movement from Darfur to Chad with mixed success.

Forced Displacement from Darfur (Sudan) into Chad

In numerous places in the evaluation reference is made to the obvious lack of preparedness of UNHCR to respond adequately with field administration, logistics, telecommunications and appropriate staff to open and staff the required operation in a remote location such as Eastern Chad. The maintenance of material and human resources for a large scale response demands financial resources even when no crisis is apparent. Clearly such resources were not available to UNHCR at this time. The researcher's experience in the funding department of UNHCR showed that financial resources only allowed the maintenance of a limited emergency response capacity which was demonstrated as a weakness in 1999 with the displacement of refugees from Kosovo. It is difficult to tell from the evaluation (Bartsch & Belgacem 2004) of the emergency response to the influx of refugees into Chad how much of the deficiencies reported were a result of lack of preparedness resulting from a lack of funds and how much was due to a lack of early warning and contingency planning,

Clearly early warning and contingency planning did play an important part in UNHCR being "behind the curve" in its response to the displacement into Chad. In December 2001, as part of a global prioritisation exercise, undertaken in response to funding shortfalls, UNHCR closed its offices in Chad and left coverage of Chad to be undertaken by the office in Bangui in neighbouring Central African Republic (CAR). (p. 25) In October 2002, refugees from the Central African Republic (CAR) began arriving in Southern Chad and in April 2003 the large influx from Darfur started necessitating rapid re-establishment of the UNHCR presence in Chad. The evaluation reports that this delayed the response significantly. In retrospect it is not possible to identify clearly why the decision to close the UNHCR presence in Chad, rather than some other African location, was taken or whether a better early warning and contingency planning system would have provided information that would have resulted in the office being maintained. The evaluation (p. 28) does, however, clearly highlight that there was a considerable time delay between the outbreak of fighting in Northern Darfur and the first arrival of refugees in Chad in April 2003 and the joint assessment mission of UNHCR, WFP and the Chadian Government in September 2003. The report is critical of the speed with which UNHCR headquarters recognised the emergency nature of the influx into Chad and took

action to respond. In its first "Systemic recommendation" the evaluation (p. 23) identifies *"the need to develop an appropriate mechanism to declare an emergency, based on predefined triggering events, that would then set off a chain response."* The same recommendation cites the then on-going "Emergency and Security Management Initiative" as a mechanism to redress the problem identified with UNHCR's early warning and contingency planning mechanisms. This Initiative has subsequently published a "catalogue" (United Nations High Commissioner for Refugees 2006) which describes the intent of the "Synthesis of Global Analysis", a tool developed by UNHCR.

> *"Through a monthly digest of sources selected and analysed by a pool of country/regional experts the Synthesis of Global Analysis, commissioned by UNHCR through WriteNet practical management, functions as a support tool for UNHCR/Donor Governments in early warning, contingency planning and crisis management."* (p. 91)

d.   Coordination

Forced displacement from Rwanda

The subtitle used in the JEEAR "filling the hollow core" gives a clear indication of the evaluators' analysis of both the necessity for and the lack of coordination in the response to the humanitarian crisis. The report points out that the coordination of UN and Red Cross agencies along with over 250 NGOs all funded by over 20 major donors was a difficult task. (Borton, Brusset & Hallam 1996, p. 164) UNHCR is singled out as having done well in the coordination of the response to the refugee situation and the report attributes this to its clear mandate, technical competence and control of a significant proportion of the available resources. This formulation suggests an interesting form of coordination, one that can be enforced by an organisation where the mandate, supported by a level of technical competence, gives a credibility that attracts a critical mass of funding which necessitates compliance rather than coordination from partner organisations.   The author's experience as Chief of Operations for UNHCR in Kosovo in 1999 was of a similar situation of coordination through compliance where the Office of Foreign Disaster Assistance (OFDA), the Humanitarian Aid and Civil Protection Department of the European Commission (ECHO) and UNHCR controlled a sufficient proportion of the available funding to coordinate jointly the winter shelter programme through these

means. Such models of coordination are rare and more voluntary, cooperative structures prevail in normal practice.

The recommendation of JEEAR (Borton, Brusset & Hallam 1996) on the topic of coordination is in the form of three options:

The first option is to *"Strengthen and extend existing inter-agency coordinating arrangements and mechanisms through: the use of inter-agency Memoranda of Understanding such as that between UNHCR and WFP" (United Nations High Commissioner for Refugees & World Food Programme 1997)*; strengthening the Department of Humanitarian Affairs (DHA) (now Office for the Coordination of Humanitarian Affairs) through funding and provision of common services through DHA and the development of inclusive coordination structures led by DHA through a small group of senior officials with clearly defined roles. (See Borton, Brusset & Hallam 1996, pp. 164-165)

This is the option which has been adopted and developed by OCHA and was certainly the option of least political resistance. It is recognised in JEEAR that this *"would be the least costly and disruptive, but the findings of Study III suggest that these efforts would not be enough to eliminate the confusion and competition and considerable difficulties faced by DHA and the United Nations Rwanda Emergency Office (UNREO) during the Rwanda emergency".* (p. 165)

The second option is for a much stronger role for DHA as the recipient of all UN humanitarian funding with the authority to *"decide on priorities and determine the amount of funds each agency would receive."* (p. 165) This coordination by compliance model was strongly supported by some donors, notably the UK Department for International Development, at a pledging meeting in 1997 at which the researcher was present. JEEAR recognises that this option would be resisted by other UN agencies, (p. 165) which indeed it was. It is also recommended by Borton, Brusset and Hallam (1996, p. 165) that DHA (OCHA) may be strengthened by secondment of technical staff from other UN agencies, a practice that continues today, both on an ad hoc basis and in a more structured manner through the selection system for Humanitarian Coordinator.

The third option, by far the most radical, is for the formation of a single UN emergency response agency by the consolidation of *"the emergency response functions of DHA and the principal UN humanitarian agencies (UNHCR, WFP and UNICEF)"* (p. 165). Although the most radical it was the option recommended by

JEEAR Study III although JEEAR also recognises that such a development may further widen the gap between relief and development.

Forced Displacement from Darfur (Sudan) into Chad

Coordination difficulties feature strongly in the UNHCR evaluation of the first six months of its emergency operation in Chad. However, it is the coordination within UNHCR that features more strongly than coordination between organisations. This is not to say that coordination issues had been solved in the decade since the writing of the JEEAR evaluation, but rather, as the report also indicates, that very few organisations were present in the theatre of operations and that it was the mobilisation of NGOs that presented a difficulty. At the time of the evaluation, UNHCR had 16 implementing partners (IPs). Even this relatively small number of IPs caused concerns regarding coordination and the ability to maintain consistency of standards across the geographic area of operations.

Some of the internal coordination problems highlighted in the evaluation are attributed to the need to re-establish the management structures in N'Djamena and in field offices closer to the refugee influx. This, and perhaps the rapidity with which various technical missions arrived in Chad, may have contributed to a situation whereby, *"many of the recommendations tabled by the numerous technical support missions fielded in areas such as registration, public health or water to assist UNHCR Chad, do not seem to have been followed through consistently."* (p. 10) As the deployment of suitable staff is also identified as a serious constraint in the Chad Operation (p. 13) there was little utility in sending technical missions to make recommendations for which there were no staff available.

An area that meets with considerable criticism in the evaluation is the effectiveness of the handover between the last Emergency Response Team (ERT) and the newly established country offices. The researcher's own experience can identify with the reported *"certain degree of friction"* (p. 21) between the ERT and the regular country team. The evaluation highlights the necessity to integrate fully the operational accomplishments of the ERTs into the country management of the operation which failed in a number of areas. (p. 21) The evaluation also highlights *"the unwillingness to honour budgetary commitments entered into by the ERT"* and that *"key achievements of the ERT, such as the setting up of a pre-registration system, fell into disrepair".* (p. 21) The evaluation concludes that there was a *"malfunction of a key mechanism which was originally designed to firmly anchor UNHCR's emergency response within the respective country operation".* (p. 21)

e. **The role of military forces (Stand-by capacity and the role of military forces in humanitarian operations)**

Forced displacement from Rwanda

Borton, Brusset and Hallam (1996) introduce the term "stand-by capacity" when discussing *"the use of military capacity as an emergency standby to complement capacity within the humanitarian system".* (p. 63)

UNHCR requested and obtained considerable assistance from military contingents following requests made to donor governments for specific service packages to augment response capacity. This was a natural follow up to the close cooperation between UNHCR and military forces both in the Sarajevo airlift and in the delivery of aid to Bosnia as a whole. (See Cutts 1999) While these arrangements did provide a rapid response, questions remained concerning their effectiveness and efficiency and the three recommendations made reflect these concerns as well as proposing a way forward. (p. 64)

The first recommendation is for a *"systematic study"* of both the *"performance and costs of military contingents in humanitarian relief operations compared with those of official agencies, NGOs and the private sector performing the same functions"* as well as *"the most effective and cost-effective ways to maintain stand-by capacity between emergencies"* (p. 166) and the establishment of well-resourced and coordinated stand-by arrangements.

The second recommendation in this section is for the development of *"clearer frameworks"* for *"civil-military cooperation in relief operations".* It is recognised that *"this may require joint training courses and exercises for agency and military personnel."* (p. 166) While NATO military have developed their framework in the form of the doctrine of Civil Military Cooperation (CIMIC), it is based on the position that humanitarian actions are only justified in so much as they contribute to the achievement of the Force Commander's objectives. (See North Atlantic Treaty Organisation 2003). The humanitarian community, through the leadership of OCHA, has developed its own guidelines for cooperation with the military in conflict situations (United Nations Inter-Agency Standing Committee 2004) which have been summarised well by Bessler and Seki (2006).

A third recommendation is that donors *"should develop schemes enabling their principal national relief NGOs to train and retain competent personnel between periods of deployment".* (p. 166)

Forced Displacement from Darfur (Sudan) into Chad

In the UNHCR evaluation of Chad, the use of stand-by mechanisms is evidenced by the observation (p. 13) that

> *"The high ratio of secondments, United Nations Volunteers (UNVs) and other categories of additional workforce over UNHCR core staff raises some concerns about UNHCR's ability to effectively staff an emergency operation of this magnitude with its own resources."* (p. 13)

The evaluation reports that

> *"By June, out of 39 approved professional functions, only nine were filled through staff being formally assigned to the operation while another 14 were deployed on missions of varying durations. The remaining 17 staff comprise a variety of non-UNHCR staff, including secondments, UN Volunteers or SURGE deployments. This amounts to a rather high ratio of supplementary staff over core staff with all its negative repercussions including a high turnover and lengthy induction periods to brief new colleagues not always familiar with basic UNHCR procedures."* (p. 13)

This indicates both that various stand-by arrangements were available and that, in this operation, UNHCR was too reliant on these measures to be able to mount an effective response. Stand-by arrangements can be used to support an operation but become its Achilles heel if over relied upon.

f. <u>Performance issues (NGO performance)</u>

Forced displacement from Rwanda

With over 250 NGOs working in the Rwandan Emergency it is not surprising that concerns are expressed in JEEAR about the performance of some of these. *"A number performed in an unprofessional and irresponsible manner that resulted not only in duplication and wasted resources but may also have contributed to an unnecessary loss of life".* (Borton, Brusset & Hallam 1996, p. 166) These concerns

highlight the ineffectiveness of coordination (see above), the unwillingness of some NGOs to participate in coordination mechanisms and the lack of technical competence of some NGOs and their staff. JEEAR recognises the work being undertaken at the time of publication (1996) on what have now come to be known as the Sphere Standards (The Sphere Project 2011) as well as the pre-existing Red Cross/NGO Code of Conduct (International Federation of Red Cross and Red Crescent Societies 1994) but expressed doubts regarding the effectiveness of voluntary compliance. (p. 166)

Two recommendations are made; the first concerning regulation or enforcement, is given in the form of two options. The first option is one of self-management through NGO networks. The second, and that favoured by JEEAR, is for an international accreditation system. The subject of the accreditation of humanitarian organisations and humanitarian workers has been revitalised (See Enhancing Learning and Research for Humanitarian Assistance (2010)) following the perception of similar failures in Haiti to those reported in JEEAR almost 15 years earlier. The Humanitarian Accountability Partnership (2010) also operates a system of accreditation.

The second recommendation is *that "Donor organizations should give greater support to NGO emergency training and lesson learning activities."* (Borton, Brusset & Hallam 1996, p. 167) The researcher's experience is that almost all major international donors have put an increased emphasis on identifying lessons learned as a part of obligatory reporting formats. The extent to which lessons learned and reported to donors are incorporated into institutional practice is examined in the second part of this research.

Forced Displacement from Darfur (Sudan) into Chad

While Bartsch and Belgacem (2004) report that *"several camps have attained a level of development that allows for the effective delivery of basic assistance… gaps were observed, not only with respect to individual sectors but also in geographical terms."* (p. 8) In this context the plight of refugees who had not yet been registered in official camps is highlighted. Concerns regarding the nutritional status of refugees and the provision of water below the accepted standard of 15 litres per person per day are expressed. (p. 8) The evaluation does not indicate to what extent these deficiencies are due to poor performance by UNHCR and NGOs and/or

the constraints of responding to an on-going influx of refugees into an area where access is difficult.

g. <u>Accountability (Improving accountability)</u>

Forced displacement from Rwanda

In many ways accountability is closely related to performance and indeed they may, to some extent, be viewed as two sides of the same coin. Borton, Brusset and Hallam (1996) report that *"the availability of quality performance data"*, (p. 167) the currency in which accountability was measured, was better in Goma than in other parts of the response to the Rwanda emergency. The evaluation team reported that the information that *was "available did not provide a sufficient basis for assessing impact or performance, or – just as important – for adjusting programme activities to improve performance".* (p. 167) Part of the problem they report was caused by both UN and NGOs emphasising their achievements and under reporting problems encountered while at the same time *"basic data on staff, finances and activities were difficult or impossible to obtain from a number of NGOs".* (p. 167)

It is not surprising then that, in the first of four recommendations, the evaluation recommends that, *"systems for improving accountability need to be strengthened"* (p. 167), and proposes a number of options, some of which repeat, or develop, previous recommendations.

The first option (p. 167) relies upon improved coordination mechanisms, the development of standards and compliance with the Red Cross/NGO Code of Conduct. (See International Federation of Red Cross and Red Crescent Societies 1994) The Code of Conduct commits signatories *to "hold ourselves responsible to both those we seek to assist and those from whom we accept resources"* (Art. 9), and in doing so, perhaps inadvertently, highlights the dilemma of dual accountability that pervades humanitarian organisations today. Since the publication of JEEAR, organisations such as ALNAP have stressed accountability and performance with a focus on internal accountability and accountability to donors. The Humanitarian Accountability Project (HAP) has produced Standards in Accountability and Quality Management (See Humanitarian Accountability Partnership 2010), which emphasise accountability to beneficiaries, and which have been adopted by a number of humanitarian organisations.

The second option is for a UN body, possibly DHA (OCHA) to form a specialised unit only responsible for accountability issues, reviewing compliance with standards, serving as a humanitarian ombudsman, maintaining a database of humanitarian operations and reporting in the public domain. To some extent OCHA has fulfilled the last two of these roles through the Consolidated Appeal Process (CAP) and its public websites.

The third option, that recommended by Borton, Brusset and Hallam (1996) is *to "Identify a respected, independent organization or network of organizations to act on behalf of beneficiaries of humanitarian assistance to perform the functions described"* (p. 168) in the second option. While the networks ALNAP and HAP have worked strongly in the area of accountability, and accountability to beneficiaries, they have not undertaken the functions specified in the third option.

The second recommendation under this topic is that the Financial Tracking System of DHA (OCHA) should be improved and that public reporting of donor contributions should be undertaken. (p. 168) Since the publication of JEEAR the tracking system has been further developed and provides regularly updated information publically through the Internet. JEEAR also recommends that military and non-military support to humanitarian operations be listed separately and costs should be compared. Currently such a system is not in operation.

The third recommendation is for the Development Assistance Committee of the OECD to develop guidelines for performance data and reporting on humanitarian assistance activities. (p. 168) DAC has published Quality Standards for Development Evaluation (Development Assistance Committee 2010) covering a much wider topic, but the development of reporting standards for humanitarian assistance is widely fragmented with donors each defining their own reporting formats based on their own public accountability requirements. Organisations operating multi-donor projects often have to report the same activities in a number of different reporting formats which vary in content and style.

The fourth recommendation is that needs assessments are regularly carried out in the areas of epidemiology, nutrition and food security using methodology and presentation that *"allow comparability between agencies"*. (p. 168) Progress in this area has been through technical standardisation and through joint assessments involving multiple providers.

Forced Displacement from Darfur (Sudan) into Chad

Bartsch and Belgacem (2004) state that: *"In general, there is only very limited data available to afford a comprehensive analysis of the extent to which minimum standards in humanitarian assistance have been met".* (p. 8) Basic record keeping and reporting requirements had not been met.   A lack of accountability to the beneficiary population is indicated in the recommendation that understanding the coping mechanisms of refugees who have survived in border areas without receiving assistance: *"may offer a key to providing targeted assistance to complement indigenous capacities and will mitigate against the creation of undue dependencies in the long run".* (p. 8) These coping mechanisms were well identified in the television documentary "Living with Refugees" (von Planta & Atkins 2004) which was filmed at the time the evaluation was being carried out. Bartsch and Belgacem (2004) report this lack of accountability in the strongest terms*, "It is still not possible to verify whether the welfare of the refugee population has improved or indeed worsened. It is for this reason that UNHCR must routinely monitor welfare indicators and track trends over time".* (p. 8)

h. **Protection (Improving camp security)**

Forced displacement from Rwanda

 This was a particular problem in Goma and other refugee camps in Zaire (now Democratic Republic of the Congo). The presence in the camps of armed elements of the genocidaires and their control of the camp populations made security a major and, for a long time, intractable issue. Camp security was addressed in Study II (Adelman & Suhrke 1996) as well as in the Synthesis Report (Eriksson 1996) and their recommendation are repeated in Study III. (pp. 168-169) The recommendation made by Borton, Brusset and Hallam (1996) is for UN Peace Keeping Missions to be given *"authority and appropriate means"* (p. 169) to protect camps, for the host government to play an active part in the reduction of violence and policing of camps as well as controlling their size and distance from the border. This recommendation continues by indicating the need for humanitarian organisations to implement measures to avoid security risks.

Forced Displacement from Darfur (Sudan) into Chad

Refugee protection, a key area of concern to UNHCR, features strongly in the evaluation. (Bartsch & Belgacem 2004) The major area of concern was to move the refugees away from the border and away from the Janjaweed incursions. *"The situation in the border areas demanded decisive action on the part of UNHCR"* (p. 7) and the rate of relocation quickly outstripped the speed at which camps could be effectively developed resulting in the sites available being overcrowded. In addition the coming rains would make relocation more difficult.

The issue of camp security did, however, contain some of the same elements that had been encountered in Goma, as the evaluation reports that *"The presence of combatants among the refugee population has been suspected since the earliest interaction with the caseload".* (p. 10) The presence of combatants, or ex combatants, may reasonably be expected in any situation where refugees are fleeing from a conflict; in the case of Chad, however, the evaluation reports *that "At the time of the mission (evaluation), UNHCR was in the process of negotiating a Memorandum of Understanding with the Chadian authorities covering the deployment of security personnel to safeguard the civilian character of the camps."* (p. 10) The type of action recommended in the JEEAR was being taken in Chad within the first six months of operation.

Borton and Eriksson (2004, p. 78) recognise the action taken by UNHCR in Chad as an example of the implementation of the recommendation made in the JEEAR.

Three major protection concerns for Bartsch and Belgacem (2004) are sexual and gender based violence (SGBV), unaccompanied minors and registration. UNHCR considers registration to be a protection issue, but as in the JEEAR, this was directly linked to food distribution and will be considered under that heading. SGBV was considered a priority issue *"especially in the light of the appalling reports from Darfur, both before and during flight".* (p. 10) The evaluation indicates that

> *"in some camps such as Kounoungo, gender-related protection considerations were included in the lay-out and camp design by involving the refugee community in the planning process"* but this was not the case in some other camps. (p. 10)

This indicates deficiencies in coordination, performance and accountability, resulting in best practice not being communicated and replicated.

Unaccompanied minors, referred to in the evaluation as "separated children" are another protection issue which reflects deficiency in some of the areas of concern raised in the JEEAR evaluation. The UNHCR evaluation of its Chad operation is categorical in stating *"Only limited interventions have taken place to identify cases of separated children."* (p. 11) A serious coordination deficiency is identified when it is reported that *"it was not clear whether contact had been established with ICRC to set up an effective tracing mechanism."* (p. 11) A deficiency in accountability is similarly identified in the statement that *"No accurate figures for the number of unaccompanied and separated children amongst the refugee population in Eastern Chad is available"* (p. 11) and refers to the cases reported to be *"most likely a serious underestimate".* (p. 11) The desperate search for family members by some refugees is documented in "Living with Refugees". (von Planta & Atkins 2004)

Recognising these deficiencies to be of strategic concern, the evaluation recommends that *"A comprehensive protection strategy should be developed....... which addresses the physical security of refugees, their legal protection including registration and documentation, as well as their material and social protection."* (p. 24) The causes and manifestation of insecurity in the refugee camps may differ from those in the forced displacement from Rwanda a decade earlier but nevertheless they remain serious concerns.

i. **Distribution (Food issues and registration)**

Forced displacement from Rwanda

The topic of Distribution (Food issues and registrations) contains five specific recommendations made following observations made by the evaluators (Borton, Brusset & Hallam 1996) concerning the methods of distribution, particularly at the beginning of the emergency response, rather than the availability of food for relief actions. The influx of Rwandan refugees, particularly into Goma, was sudden and rapid and in order to distribute food before registration was completed group leaders in the camp were used as *"an expedient mechanism for food distribution."* (Borton, Brusset & Hallam 1996, p. 169) As they observe, this strengthened the power of the group leaders, many of whom were from the previous political leadership in Rwanda, and amongst whom were genocidiers. These leaders were able to manipulate and divert food, with the result that groups such as the elderly and female headed households had difficulty in accessing food rations. The move to

individual registration and food distribution directly to families was seen to be against the interests of the leadership, who also to a large extent controlled the security of international aid workers in the camps. Borton, Brusset and Hallam (1996) do, however, observe that some organisations were able to move to direct distribution more quickly than others. The report, observing that the failings of the distribution system, rather than a general lack of food, were the major cause of malnutrition, is critical of organisations for focussing on the introduction of supplementary feeding programmes rather than on addressing the problems of the distribution through group leaders. (p. 169)

The first recommendation (p. 169) concerns operational guidelines for food distribution including distribution, directly to families rather than through group leaders, and distribution to women where necessary. The provisions of this recommendation have been included in the Memorandum of Understanding (MOU) between UNHCR and WFP (United Nations High Commissioner for Refugees & World Food Programme 1997 Paragraph 2.2) and the intervening years have seen a proliferation of guidelines, handbooks, manuals and other reference material produced by various organisations. Two of note in the context of refugee and IDP camps which contain sections on food distribution are those of United Nations High Commissioner for Refugees (2007) and the Norwegian Refugee Council (2004).

The second recommendation (p. 169) concerns deployment of registration/enumeration specialists with UNHCR Emergency Response Teams, as in the case of the emergency response to the influx of refugees from Darfur into Chad considered below. As will be seen, however, this institutional solution was not without its own set of problems.

The third recommendation is that *"formal food needs assessments should be carried out early in an emergency".* (p. 169) Working under the MOU between UNHCR and WFP (United Nations High Commissioner for Refugees & World Food Programme 1997) and utilising the UNHCR/WFP Joint Assessment Guidelines (United Nations High Commissioner for Refugees & World Food Programme 2008) such needs assessments are now regularly carried out.

The fourth recommendation concerns *"the establishment of Supplementary Feeding Programmes (SFPs)* (p. 169) *in refugee camps"* as a substitute for addressing problems in the general ration supply. Specific guidelines are now available (United Nations High Commissioner for Refugees & World Food Programme 1999) for the establishment of SFPs in refugee and IDP camps.

The fifth recommendation is that *"The costs of milling cereals supplied by WFP as part of the general ration should be included within the Internal Transport Storage and Handling (ITSH) costs and therefore paid automatically by donor organizations".* (p. 170) Such provisions are included in the MOU between UNHCR and WFP:

> *"If whole grain is provided local milling capacity must be available, and the ration should include compensation for milling costs (normally 10 per cent up to 20 per cent, if justified), if these costs are borne by the beneficiaries. WFP is responsible for mobilising the necessary resources for milling and will provide milling facilities to the beneficiaries where feasible".* (Paragraph 4.3)

Forced Displacement from Darfur (Sudan) into Chad

These are the issues that form the main story line of the documentary "Living With Refugees" (von Planta & Atkins 2004). In this documentary, an African journalist, Sorius Samura, accompanies a refugee family as they travel from the border area close to Sudan to a refugee camp deeper inside Chad and struggle to survive for several weeks waiting to be registered and to receive assistance. The harshness of Chadian Government officials responsible for registration and the apparent helplessness of UNHCR officials is contrasted with the relative compassion of a traditional leader and the willingness to share scarce resources by fellow refugee families. Bartsch and Belgacem (2004) also highlight problems encountered with registration; one of which is identified as being an exclusively camp based approach to assistance, leaving *"pockets of altogether unattended needs outside the camps especially amongst the spontaneous arrivals".* (p. 8)

As was recommended in the JEEAR evaluation, staff with expertise in registration were deployed with the Emergency Response Teams (ERT) and left clear recommendations on the continuation of the process (p. 11). The evaluation reports staff interviewed as saying that, *"the initial design of the registration system was well functioning but subsequently fell into disrepair, both on account of partner capacity and the lack of UNHCR Protection staff to maintain the system accordingly."* (p. 11) The registration activities had been

> *"entrusted to the Commission Nationale pour l'Accueil et Réinsertion des réfugiés (CNAR)"* who had *"instituted simple procedures for the pre-*

registration and registration of new arrivals including the recording of head of household and basic demographic breakdown on a simplified control sheet". (p. 11)

This indicates that registration and the subsequent distribution were designed to be by families as recommended in the JEEAR evaluation. However it is also clear that the system broke down due to a number of factors including:

- the identification and construction of camp sites lagging well behind the speed of arrivals,
- the lack of capacity and training of the government officials involved in the registration
- the lack of UNHCR staff to undertake training and oversee the process once the ERT had left.

The question is then open as to whether a short term distribution of basic food through group leaders, exactly what was criticised and advised against by the JEEAR evaluation, may have been a better approach. Borton and Eriksson (2004) recognise that

"in the early days of large-scale refugee influxes where formal registration has not yet been undertaken, UNHCR and its implementing partners have little choice but to make use of the group leadership within the refugees to organize initial distributions". (p. 82)

j. **The role of the media.**

Forced displacement from Rwanda

The media had played a major role in publicising the war and the humanitarian response in Bosnia in 1993 and 1994, coverage of which was largely made possible by the advent of satellite communications. This publicity of a humanitarian crisis, particularly by television news, was so effective that the term "the CNN effect" was coined for the way it raised public awareness. (See Robinson 2002) Never before had humanitarian suffering been so widely publicised or the successes and failures of the humanitarian effort received so wide a coverage. As Bosnia faded from the headlines, the attention of the world's media shifted to the massacre in Rwanda and subsequent humanitarian operations in the region. Borton, Brusset and Hallam

(1996) recognise that the advent of on-the-spot satellite broadcasting and the powerful nature of the images of the influx into Goma contributed to the massive response there, and may also have *"contributed to the lack of policy coherence by the media's focus upon the humanitarian story, rather than the more complicated and difficult-to-comprehend story of the genocide and the conflict."* (p.170) While the report does not make any formal numbered recommendations for this issue it does suggest that:

> *"A rigorous study of media coverage of humanitarian aid operations and the way it influences and is, in turn, influenced, by relief agencies should be undertaken. To increase its ability to inform policy, the study should cover more than one relief operation."* (p. 170)

The subject of the relationship between the media and the response to political and humanitarian crises has been the subject of academic research. (See CARMA 2006)

Olsen, Carstensen and Høyen (2003) examined the effect of the media on the response to seven humanitarian emergencies, two natural disasters, the Indian cyclone of Oct 1999 and the Mozambique floods of Jan 2000 and five complex emergencies Angola, Sudan, Balkans, DRP Korea and Afghanistan. They conclude that

> *"only occasionally do the media play a decisive role in influencing donors. Rather, the security interests of Western donors are important together with the presence and strength of humanitarian stakeholders, such as NGOs and international organisations lobbying donor governments."* (p. 109)

The relationship between the media and humanitarian assistance has, since the publication of JEEAR, received attention from humanitarian organisations, many of which now employ a cadre of media professionals to manage their relationship to the press, particularly in major humanitarian crises

Forced Displacement from Darfur (Sudan) into Chad

Bartsch and Belgacem (2004) do not make any reference to the effect of the media. However the documentary referred to above (von Planta & Atkins 2004) was either

in production or being broadcast at the same time as the evaluation was taking place. As the documentary features an interview with a UNHCR Public Information Officer in one of the camps, the UNHCR team would have been aware of the documentary and some of its content well before it was broadcast.

k.   Mitigating the impact on host communities.

Forced displacement from Rwanda

Borton, Brusset and Hallam (1996) observe that the presence of large numbers of refugees and displaced persons had an effect on local communities with both gains and losses being noted.   Some of the most important losses were to the environment surrounding large camps, including Goma, an environment on which the local population depended as a reliable resource. The humanitarian actors were slow to compensate the local populations which resulted in antagonism towards the refugees which limited *"the options for the wider management of the crisis".* (p. 170)

The first recommendation is that policies are put in place both to *"minimise and mitigate the effects of relief operations on the surrounding populations and their environment".* (p. 170) UNHCR has since published a set of environmental guidelines for use in refugee and IDP situations. (See United Nations High Commissioner for Refugees 2005)

The second recommendation is that funding should be available to provide rapid compensation to host populations adversely effected by refugees. (p. 170)

The third recommendation lists strategies that should be in place to minimise the negative effects of refugee concentrations and lists:

> *"providing food that requires little or no cooking; providing fuel for cooking; extending camp infrastructure and services (health care, water supply, etc.) to surrounding local populations; and rehabilitating physical infrastructure (e.g. roads and airstrips) damaged in meeting relief needs."* (p. 170)

(Borton & Eriksson 2004, p. 82) identify a number of actions taken between UNHCR, UNICEF, UNDP and WFP to minimise adverse local impacts of large displacements. They conclude, however, that:

> *"studies of the impact of refugee operations on host communities continue to tell a story of negative impacts on certain sections within the host communities, often the poorest least able to adapt, that are not offset by the economic opportunities and improved access to certain types of services resulting from the refugee operation."* (p. 82)

In this context Borton and Eriksson (2004, p. 82) also cite the UNHCR evaluation of the displacement into Chad which is discussed in the following paragraphs.

Forced Displacement from Darfur (Sudan) into Chad

Mitigating the impact on host communities is recognised by Bartsch and Belgacem (2004) at the beginning of the findings.

> *"Inevitably, the overpopulation of the camps beyond their assessed capacity aggravates the impact that such a large population concentration is bound to have on a very fragile environment as regards water supply, grazing land and firewood. While refugees could initially depend on an extremely generous welcome from the host population, the first conflicts over resources between refugees and locals around the camp sites have already erupted and are likely to become increasingly virulent. This competition for resources needs to be addressed at the earliest stage if outright conflict is to be prevented and, to the extent possible, both refugee and host communities must directly be associated with the design of the assistance programmes."* (pp. 7-8).

Clearly major and potentially catastrophic deficiencies existed at the time of the evaluation.

The problem of identifying and negotiating for sites is cited as a reason in at least one case.

> *"This stalemate in the negotiations for a site lasted for several months while in the meantime the numbers of refugees increased substantially, most arriving with large numbers of camels and donkeys, thus exerting extreme pressure on the local infrastructure."* (p. 9)

These deficiencies are attributed to a lack of foresight or of a stronger negotiating stance. The recommendation that follows is however bland and lacking in real

substance *"Potential conflicts with the host population over scarce resources need to be anticipated and factored into the planning.* (p. 24)

Conclusion

The difficulty in responding to a rapid movement of a relatively large population into an ecologically fragile and logistically challenging environment consistently results in the response being, to use Bartsch's phrase *"behind the curve".* It may be argued that this is inevitable and largely due to the lead time in mobilising financial, material and human resources necessary to mount an appropriate response. However this is simply to argue that the problems of donor funding and preparedness measures, humanitarian early warning and contingency planning cannot be solved. The recommendations of the JEEAR to address these concerns do not seem to be impractical and, if this is the case, then the problem must lie elsewhere. It is clear that in the time between the publication of JEEAR and the evaluation of the response in Chad (1996 – 2004) that a number of institutional mechanisms had been put into place designed to meet the requirements of the recommendations made. Deployment of ERTs containing specialists in registration, early nutritional surveys, distribution directly to families and an emphasis on moving refugees away from the border into camps are all examples. All appear to have been to some extent unfit for purpose when tested against the forced displacement from Darfur into Chad.

Chapter 5 - A comparison of the recommendations of evaluations of the humanitarian response to the Indian Ocean tsunami (2004) and the Haiti earthquake (2010)

The humanitarian responses chosen here for comparison have some marked similarities with each other but also considerable differences. They are certainly the two largest humanitarian responses mounted to natural disasters during the decade. Both received massive media coverage, financial contributions and attracted large numbers of humanitarian organisations. These two events are separated by five years and 17 days.

While the sources of both disasters were earthquakes the direct cause of the damage in Haiti was the earthquake itself, while in the area surrounding the Indian Ocean it was the resulting tsunami. As a result the damage and loss of life was contained in one country in the case of Haiti but widely dispersed over many countries in the case of the Indian Ocean tsunami. In Haiti the capital city was heavily damaged whereas in South Asia coastal regions were involved. Overall, however, it is considered that the similarities of humanitarian responses to large rapid onset natural disasters are the overriding factor that makes the comparison relevant and appropriate.

The recommendations of the Tsunami Evaluaton Coalition (2007) reports on the Indian Ocean tsunami are categorised and tabulated in one volume under 11 headings as indicated in the preceding Chapter 3 (Design and Methods). Each of these 11 categories is further subdivided into groups of recommendations some of which are used in the analysis below. The categories overlap considerably and as a result repetition of recommendations in more than one category is common. In order to facilitate the comparison the researcher has, wherever possible, minimised repetition which has led to the elimination of some categories. Where repetition has been unavoidable, the topic is covered fully under one category and briefly mentioned where appropriate in others.

The TEC evaluations are a detailed, largely post facto, analysis of the response published between July 2006 and January 2007 designed to promote change in the way the international humanitarian community approaches responses to future natural disasters.

The evaluations of the humanitarian response to the earthquake in Haiti vary considerably in style and as to how much emphasis they put on making specific recommendations.

- Binder and Grünewald (2010) evaluated the mechanism for the coordination of humanitarian action (the cluster approach) in place at the time the earthquake took place. The field work for this evaluation was, in fact, completed a few months before the earthquake.

- Grünewald, Binder and Georges (2010) undertook an evaluation, commissioned by the Inter Agency Standing Committee, of the initial humanitarian response three months after the earthquake

- Interagency Standing Committee (2010) undertook an evaluation at the request of its members.

    "*The report is written 6 months after the 12 January earthquake in Haiti and is concerned primarily with the response by IASC members to the disaster, but necessarily refers to the role of other key actors, including the Haitian population and Government, international militaries, the UN Stabilization Mission in Haiti (MINUSTAH), and regional entities six months after the event.*" (p. 1)

- DARA (2010) included a chapter on the humanitarian response in Haiti in a larger analysis of humanitarian responses worldwide written approximately one year after the disaster.

- Patrick (2011) reviewed evaluations available some18 months after the earthquake and produced a synthesis of *"some of the key emerging lessons"* (p. 1) taken from a review of 28 evaluations and mission reports.

- DARA (2011) included a similar chapter written about two years after the disaster in their subsequent annual publication.

These evaluations provide a series of "snapshots" of aspects of the humanitarian response from the time the earthquake struck up until about two years following the

event. In some cases they tell a coherent story of criticisms and recommendations either being acted upon or apparently ignored. This is particularly relevant as the evaluations are referred to as "real time" or "emerging lessons" or are part of a series of annual reports all designed to identify current problems and bring about immediate corrective action.

1. To make the comparison between the emergencies and between the results of the evaluations of the humanitarian responses, this chapter will first describe the events and the initial response and then be structured using headings derived from the TEC recommendations as indicated in Table 5.1. The headings used will be:

Table 5.1 Structure used for comparisons of the evaluations of the humanitarian responses to the Indian Ocean tsunami and the Haiti earthquake

	Main Headings	**Sub-headings**
1	The event and the initial response	
2	Policy coherence	
3	Funding issues	
4	Coordination	a. Assessment b. Information and its uses c. People and services d. Supporting quality
5	Actors other than humanitarian	a. Media b. Military c. Private sector d. Affected country governments
6	Accountability	a. Information flow and reporting b. Rights and protection c. Upward accountability and reporting
7	Linking Relief, Rehabilitation and Development, Disaster Risk Reduction and Early Warning.	
8	Ownership	

Source: Author

1. **The event and initial response**

Indian Ocean tsunami (2004)

According to Pickrell (2005) the Indian Ocean tsunami was a result of a magnitude 9 earthquake which took place at 0758 local time on 26 December 2004. The

epicentre was 250 miles south-south east of Banda Ache, Indonesia, some 30 kilometres below the seafloor. The earthquake was the fourth largest since 1900 and the largest since a 9.2 earthquake in Alaska in 1964. The resulting tsunami wave was up to 10 metres high on the coast line of Sumatra close to the epicentre and some four metres high in Thailand, Sri Lanka and Somalia. A total of 13 countries sustained damage. The same article published less than a month after the event, quotes 226,000 dead, 500,000 injured and five million homeless and without access to food and water.

The disaster received massive coverage in the press (Matthews 2009, pp. 26-29) both as a result of the large number of countries involved and the deaths of tourists from Australia, Europe and America on holiday in affected areas over the Christmas to New Year period. As a result of the human interest and concern large amounts of both private and public contributions rapidly resulted. Cosgrave (2007, pp. 10-11) observes that while initial assistance was provided by survivors helping their neighbours, assistance soon arrived from communities further inland which had not been affected. In some countries the national military rapidly provided relief, soon followed by international NGOs, foreign Red Cross agencies and the UN. He also observes that the scale of contributions rapidly overwhelmed the implementation capacity of relief organisations. As early as the 6th of January 2005 Agence France-Presse (AFP) reported:

> *"In Jakarta this week EU aid commissioner Louis Michel condemned the "beauty contest" underway, with world leaders vying to announce spectacular aid pledges, regardless of the actual needs or capacities of affected countries. "* (AFP 2005)

Haiti earthquake (2010)

According to Interagency Standing Committee (2010) the disaster was a result of a magnitude 7.0 earthquake which took place at 16.53 on 12 January 2010 with an epicentre less than 10 km below the earth's surface close to Lèogâne some 25 kilometres from Port-au-Prince the Haitian capital. An aftershock of magnitude 6.0 followed almost immediately and by 24 January there had been at least 52 aftershocks of magnitude 4.5 or higher. It is estimated that 223,000 people died with over 300,000 injured. Sciba (2011) indicates that of an estimated three million affected population some two million were displaced. While there was a large

presence of international organisations and NGOs in Haiti implementing relief programmes following the 2009 hurricane season, many suffered significant loss of personnel, equipment and supplies as a result of the earthquake, reducing their response capacity. However, these organisations, almost all of which were supported by national counterparts within Haitian civil society, were able to provide assistance within the first 72 hours. (p. 8)

The disaster received an immediate high level of media coverage, much of which was so graphic as to prompt the Washington Post, just four days after the earthquake to comment:

> *"The images coming out of Haiti are more graphic than those from recent natural disasters, and the on-going wars in Iraq and Afghanistan. It's not clear if this reflects the magnitude and proximity of the disaster, or some change in the willingness of newspapers and other media to accurately present the full horror of the earthquake that devastated the desperately poor nation"* (Kennicot 2010)

Partially as a result of this media coverage an unprecedented relief operation was mounted and by *"May 2010, over 1,000 international organizations had provided humanitarian assistance in Haiti"*. (Grünewald, Binder & Georges 2010, p. 7)

2. Policy Coherence

The policy coherence issues in the case of natural disasters differ markedly from those relating to conflict displacement discussed in the previous chapter. The role of the Security Council is marginal and, in Haiti, limited to a resolution extending the mandate of the United Nations Stabilization Mission in Haiti (MINUSTAH) (United Nations Security Council 2010) to cover its participation in the humanitarian response to the earthquake. While following both disasters many nations and private donors rushed to contribute, there was very little politicisation of the situations themselves. The policy coherence issues that emerged in the evaluations of both natural disasters concerned humanitarian rather than political issues of the type which dominated the evaluations of the forced displacements from conflict.

Indian Ocean tsunami (2004)

One set of policy coherence issues emerged from the enormous, sometimes excessive, funds that were rapidly made available for the humanitarian response. The main issue was that the appeals for the response to the Indian Ocean tsunami were better and more quickly funded than those of other humanitarian responses taking place at the same time. This was considered a threat to the humanitarian principle of impartiality under which a humanitarian response should be in proportion to the needs of the affected population. With a long history of Consolidated Appeals (CAPs) being underfunded and hence not meeting the assessed needs of populations in other disasters around the globe Flint and Goyder (2006) recommend that:

> *"The international community needs to consider whether it is prepared to give substance to the Good Humanitarian Donorship (GHD) principles by committing to a target that all people affected by disasters should be entitled to a certain minimum level of humanitarian assistance and, if so, whether the current appeal-based system can deliver the resources to achieve that."* (p. 40) and

> *"Humanitarian agencies need to recognise that a commitment to impartiality may be inconsistent with open-ended appeals, and may require reallocating funds already raised. Flexibility in the use of funds – in line with the principle of impartiality – needs to be increased for future appeals by allowing private and government donors to indicate… that their donation can be used for other humanitarian emergencies once either the appeal target or assessed needs have been met."* (p. 40)

In the situation of a surfeit of funding, concerns are raised regarding accountability and corresponding recommendations made. The need for *"improved tracking and reporting financial data"* is highlighted and *"there is a serious need to understand how the humanitarian dollar flows from original donor to actual beneficiary, documenting each layer, the transaction costs and added values."* (Flint & Goyder 2006, p. 42) In a situation where the availability of funds often outstripped the implementation capacity of the humanitarian organisations there was a concern that aspects of efficiency and effectiveness may have been less important than the need to be seen to be spending money and showing some results quickly. One specific concern was indicated by Bennett et al. (2006) who recommend that *"IASC should urgently introduce monitored guidelines requiring all agencies to report on the*

numbers and cost of visiting delegations". (p. 15) One effect of the ready availability of funding was that budgets for field visits from organisations' headquarters and by donors had been relaxed to the point where they had become visible for both the expenditure involved and as a hindrance to action on the ground. (Bennett et al. 2006, p. 38)

Another set of policy coherence issues which emerge from the humanitarian response to the Indian Ocean tsunami is made by Telford, Cosgrave and Houghton (2006) who recommend *"a fundamental reorientation" of the humanitarian community from supplying aid to facilitating communities' own relief and recovery priorities"*. (p. 118) The implications of this recommendation are discussed in detail at various points in this chapter. It is this policy coherence issue that carries over more directly to the humanitarian response to the Haiti earthquake.

Haiti earthquake (2010)

The policy coherence issues in the humanitarian response to the Haiti earthquake centre on the marginalisation and/or exclusion of the Haitian government, civil society and the affected population. All of the evaluations reviewed make reference to various aspects of this problem and the ways in which it limited both the effectiveness and efficiency of the response. As is discussed under various headings below the implementation of the humanitarian response in Haiti not only ignored the recommendations that emerged from the second set of policy coherence issues in the response to the Indian Ocean tsunami but resulted in a course of action diametrically opposed to them.

3. **Funding**

Indian Ocean tsunami (2004)

Flint and Goyder (2006, p. 8) refer to the response to the Indian Ocean tsunami as *"the most generous and immediately funded international humanitarian response ever. US$ 14 billion has been pledged or donated for emergency relief and reconstruction from international sources."* (p. 8) The same source, however observes that *"it is apparent that allocation and programming, particularly in the first weeks and months of 2005 were driven by the extent of public and media interest, and by the unprecedented funding available rather than by assessment and need"* (p. 8)

The ready availability of funding, particularly funding from private donors, *"sometimes exceeded agency and local capacities"* (p. 39) leading to the recommendation that private and government donors be allowed *"to indicate that their donation can be used for other humanitarian emergencies once either the appeal target or assessed needs have been met"* (p. 40) This recommendation is made in the context of the principles of improved impartiality of humanitarian response and of Good Humanitarian Donorship. (Good Humanitarian Donorship 2003)

Haiti earthquake (2010)

DARA (2010) observes that the *"exact amount of money donated to the Haiti response will never be known. According to OCHA's Financial Tracking System (FTS), as of 9 October 2010, over US$3.5 billion had been raised. However, significant donations have not been reported to the FTS"*. (p. 162)

DARA (2010) observes that *"Funding decisions were largely made at headquarters level and not based on needs assessments"*. (p. 160) This may have initially been a result of the delay in publishing the Rapid Inter-Agency Needs Assessment for Haiti. (p. 162) Grünewald, Binder and Georges (2010) observe that *"Three months into the operations, however, aid actors and observers alike started to raise questions about the appropriateness and relevance of funding."* (p. 35)

The quantity, disproportion and earmarking of funds for the response is criticised by Grünewald, Binder and Georges (2010), in a situation where

> *"many international NGOs, including the large ones, had reached the limits of their capacity to absorb emergency funding … At the same time, important areas such as Agriculture, Early Recovery and Education had either been under-funded or only received funds after a long delay."* (p. 35).

The lack of funding for longer term recovery, and particularly for permanent shelter, (DARA 2010, p. 160) leads to a focus on the inequitable distribution of funding within the needs of the one crisis rather than the inequitable distribution of funding between crises as seen in the case of the Indian Ocean tsunami. In one way or another, however, the distribution of available funding does indicate that the coordination of donors as well as the flexibility, given to and demanded of, aid agencies remains a problem. DARA (2010, p. 162) concludes that *"As with the*

tsunami, the challenge is for all actors to use resources effectively to meet immediate and long-term needs." (p. 162)

4. <u>Coordination</u>

Indian Ocean tsunami (2004)

a. <u>Assessment</u>

The recommendations made in the TEC Tsunami Evaluaton Coalition (2007, p. 11) focus on the need for a coordinated needs assessment led by OCHA and using both common standards and reporting formats. To facilitate this process, mechanisms for both staffing and funding of the assessment teams should be able to respond rapidly and should involve both international and local actors.

b. <u>Information and its uses</u>

Key recommendations centre on the design of information systems to be standardised with a view to being handed over to Government as part of the humanitarian exit strategy (Bennett et al. 2006, p. 69), on the refinement of the capacity of data systems to improve the quality of material they contain (de Ville de Goyet & Morinière 2006, p. 54) the ability to filter and analyse the data (Bennett et al. 2006, p. 69).

Bennett et al. (2006) recommends *that "low cost satellite communications, internet systems and GSM telephone systems need to be examined and made available to UN and UN partners at the national level".* (p. 67)

c. <u>People and Services</u>

The ready availability of funding and the intensity of media focus on the crisis resulting from the Indian Ocean tsunami led to a large number of high profile visits from agency heads and goodwill ambassadors. Bennett et al. (2006) are very critical of the number and effect of visits to the field and recommends *that "the IASC should urgently introduce monitored guidelines requiring all agencies and donors to report on the number and cost of visiting delegations."* (p. 15)

The rapid deployment of emergency response teams by humanitarian organisations, including OCHA, led to staff in important posts being replaced on relatively short cycles. The high turnover of staff, particularly those involved in coordination, is criticised and it is recommended that *"HR departments should endeavour to deploy long-term (at least one year) personnel in the field as soon as possible."* (Bennett et al. 2006, p. 16)

As a large number of humanitarian organisations responded to the crisis and each set up its own administrative and logistics systems, knowledge of and the utilisation of common services (under the Humanitarian Common Service 'matrix') by all actors is encouraged (Bennett et al. 2006, p. 16). Telford, Cosgrave and Houghton (2006) recommend that *"Where appropriate, and with the support of multilateral agencies, states should establish, and international agencies should be prepared to work through, common mechanisms such as consortia and trust funds."* (p. 149)

d. Supporting quality

Telford, Cosgrave and Houghton (2006) recommend that *"International agencies need to respect the role and responsibility of affected states as the primary authorities – be they national, provincial, local – in responding to natural disasters and ensuring risk reduction".* (p. 114) They comment on a humanitarian response, undertaken without sufficient respect for the role of the host governments and thus compromising longer term quality. The important focus of the recommendations is the enhancement of the role of these governments. Telford, Cosgrave and Houghton (2006) recommend that *"International agencies should respect and promote national coordination of all response activities"* and *"the UN should play its mandated coordination role… by developing a coordination model that supports national coordination efforts"* and that these should be extended to the regional as well as national levels. (p. 115)

Bennett et al. (2006) also recommend that:

> *"Leadership and coordination skills should include the basics of how to maximise the output of meetings. These skills should be promoted by all agencies, forming part of the induction training for operational staff, along with standard operating procedures."* (p.15)

Haiti earthquake (2010)

     a.   <u>Assessment</u>

Patrick (2011) observes that:

> *"Most individual agencies conducted their own needs assessments, but each followed different standards, methodologies and focus thus limiting the usefulness of the results for an overall analysis or strategic planning."* (p. 3)

> And that the situation in Haiti: *"Necessitated the Rapid Initial Needs Assessment for Haiti undertaken by the international humanitarian community which was quick to implement, but slow to publish. As such many of its findings were out of date by the time they were widely available."* (p. 3)

DARA (2010) report that as a result of the delay in publication, the results of the Rapid Needs Assessment *"were not seen by many donors before funding decisions were made"*. (p. 162)

The standardisation of needs assessment recommended in the evaluation of the response to the Indian Ocean tsunami had not taken place, or was not utilised. The recommendation for the rapid funding and deployment of a needs assessment team had been implemented but the delay in publishing severely limited the usefulness of the exercise.

     b.   <u>Information and its uses</u>

Technological development in the intervening years made it almost inevitable that the types of technologically advanced communications systems recommended by Bennett et al. (2006, p. 67) would be widely available in Haiti five years later. Haiti was a testing ground for new information systems most of which were internet based. DARA (2010) reports that one of their informants described this as *"sexy but doesn't necessarily work"*. (p. 164) The use of internet based systems in a situation where internet access was often problematic is criticised. Agencies that did have useful internet connectivity did not find the OCHA "Haiti One Response" website helpful and DARA (2010, p. 164) reports that several clusters used Google Groups and Google Docs.

c.   People and Services

The frequency of visits to the emergency particularly by donors   is criticised by Grünewald, Binder and Georges (2010, p. 10) recommending *that "they also need to ensure that they do not hold up aid agencies in their work"* and that donors and humanitarian organisations *"should… limit high level visits".* This is probably a more realistic approach than that taken by the evaluation into the response to the Indian Ocean tsunami.

Just as in the tsunami the deployment of long term staff for coordination was a problem in Haiti. Binder and Grünewald (2010) simply recommend *"Ensure that cluster coordinators are deployed for at least 6 months".* (p. 46) This is considerably more limited than the recommendation made in the case of the Indian Ocean tsunami which was for a broader range of staff for at least 12 months.

d.   Supporting quality

"*Developing a coordination model that supports national coordination efforts*" as recommended following the response to the Indian Ocean tsunami by Telford, Cosgrave and Houghton (2006, p. 115) remained a problem in the response to the Haiti earthquake. " DARA (2010, p. 165) observes   *"Clusters: the same old problems" and "Convening of cluster meetings in accessible locations, the over-use of English, the limited engagement of government and civil society and the quick turnover of coordinators highlight the need to discuss how to make the cluster system more effective"* (p. 167)

There is one surprising difference between the recommendations of the evaluation of the response to the Indian Ocean tsunami and those of the recommendations of the evaluation of the response to the Haiti earthquake that reflects the development of the cluster system over the intervening five years. In the case of the tsunami it is recommended by Bennett et al. (2006)   that NGOs should have *"adequate representation within coordination structures at all levels"* (p. 15) In the case of Haiti, however, the problem was the number of participants in the cluster structure. The Interagency Standing Committee (2010) observes that *"Clusters had to contend with a rapidly increasing influx of humanitarian* actors (the Health Cluster, for example, at one stage had 420 participating organizations), *and were therefore somewhat limited in taking key strategic decisions."* (p. 17)

At the same time the clusters were weakened by the *"failure to sufficiently involve the Haitian state or civil society"*. (DARA 2010, p. 160) The solution adopted to this problem was "*The lead agency, key cluster members and (in some cases) the government met as a baby cluster or strategic advisory group"* (Grünewald, Binder & Georges 2010, p. 33) to decide on strategy with the *"Separation of the two key functions of a cluster: information sharing and strategy development / technical guidance*" (p. 33)   In their recommendation for future development of this topic Grünewald, Binder and Georges (2010) are more specific *"While the strategic function should be exclusive (but transparent) and should involve only the government, the cluster lead agencies and a handful of experienced NGOs, the information-sharing function should be open to all actors in the field."* (p. 62)

5. **Actors other than Humanitarian**

Indian Ocean tsunami (2004)

The "other actors" specifically focussed upon in the evaluation of the response to the Indian Ocean tsunami are the media, the military and the private sector (Tsunami Evaluaton Coalition 2007, pp. 50-52) . The relationship with affected country governments is covered under a number of other categories but, as this features heavily in the evaluations of the response to the earthquake in Haiti, it will be covered under this heading.

a.   Media

Telford, Cosgrave and Houghton (2006) recognise the role of the media in influencing the way *"many actors make their initial funding and deployment decisions"* and it is recommended that *"media organisations formally appoint journalists with an interest in the sector as aid correspondents."* (p. 118)

The role of the media in communicating to the recipient population is also recognised and it is recommended by Bennett et al. (2006) that *"A common strategy should include guidance on the use of public meetings, broadcast media, newsletters and posters"* (p. 15)

The necessity of using the media regarding donor contributions to give visibility to donor support is recommended by Flint and Goyder (2006) *"…tracking and*

reporting financial NGO's data needs to be a priority, as does improving feedback from agencies to their donors and to the media." (p. 42)

The inclusion of media representatives in needs assessment missions is recommended by de Ville de Goyet and Morinière (2006) *"Initial assessment teams should routinely include selected mass media representatives."* (p. 66)

b.   Military

Despite, or perhaps because of, the extensive deployment of international and local military resources in response to the massive destruction caused by the tsunami, severe gaps in civil-military coordination were observed in the report on coordination. (Bennett et al. 2006)

> *"The use of military assets in aid delivery has been the subject of controversy within the humanitarian community for many years. In conflict environments it is particularly sensitive, and even in the wake of natural disasters the humanitarian community usually deems their use acceptable 'only as a last resort' and under the direction of the appropriate civilian authority. The unprecedented deployment of so many military forces in response to the tsunami disaster has again brought to the fore a large number of issues regarding the use of military assets and heralded the next stage in the debate on the civil–military nexus. Although dialogue between humanitarian and military actors has improved in recent years, it is often still characterised by a lack of understanding, institutional differences and mutual suspicion."* (p. 45)

As a result, several lines of approach are recommended for improvement. (Bennett et al. 2006, p. 16)

These may be summed up as:

- Promotion and publication of the existing guidelines, although the necessity for revision is recognised. (p. 46) The Oslo Guidelines (United Nations Office for the Coordination of Humanitarian Affairs 2007) were first published in 1994 and finally revised in 2007.

- In house and external training

- Joint exercises between humanitarian agencies and the military.

Bennett et al. (2006) also observes that *"some of the military interventions came at a very high price"*, (p. 27) a reference to the high costs often charged by the military to their own government's humanitarian budget for deployment on humanitarian missions.

c. Private Sector

Although included as a section in the recommendations the items included are rather unspecific except for Bennett et al. (2006, pp. 16,45) recommending that pro bono offers of goods and services should be coordinated by OCHA on behalf of *"the wider humanitarian community"* in accord with the provisions of UN Guidelines. (United Nations 2006)

From a reading of the coordination report, some more specific areas of cooperation become apparent.

- The technology for cell phones and satellite imagery, which *"emerged as important instruments of communication and coordination in the immediate stages of the emergency… was in the hands of the private sector"* and greater efforts by humanitarian agencies are encouraged to form appropriate partnerships (Bennett et al. 2006, p. 13)

- The role of the Special Envoy (Bill Clinton) is emphasised by Bennett et al. (2006) *"the Special Envoy will give greater attention to private-sector involvement in the recovery process, the improvement of regional early-warning and disaster-reduction practices"* (p. 31).

Bennett et al. (2006) concludes that *"we should learn how better to harness the considerable resources of the private sector"* (p. 89) and that to this end *"OCHA should also ensure that all major emergencies have a dedicated focal point for liaising with key private sector companies"*. (p. 16)

The affected country governments are not addressed as a separate heading in the "Other actors" section but several of the recommendations of the TEC reports are addressed to them. As the subject of the role and relationship with the government is such an important issue in the evaluations of the response to the earthquake in Haiti it is addressed as a separate heading here.

Telford, Cosgrave and Houghton sum up the issues well:

> *"International agencies need to respect the role and responsibility of affected states as the primary authorities – be they national, provincial, local – in responding to natural disasters and ensuring risk reduction. Similarly, states in high-risk regions have a responsibility as the primary duty-bearers in risk-reduction activities and natural disaster response. States should set standards and procedures for inviting, receiving and regulating international assistance."* (p.114)

These principles have since been recognised in the development of Disaster Law. (International Federation of Red Cross and Red Crescent Societies 2011b)

Haiti earthquake (2010)

Patrick (2011, p. 8) sums up the situation with respect to the role of the non-humanitarian actors well *"Humanitarian coordination should accommodate non-humanitarian actors, most notably military, private sector, host government and local community and civil society."* The wording indicates that this is an aspiration and not an observation of the accommodation of the non-humanitarian actors in the response to the 2010 earthquake in Haiti

a.   Media

DARA (2011) demonstrates the dilemma that a high level of media coverage presents to the international humanitarian community for while

> *"The earthquake mobilised a massive international response, triggered partly by the close proximity to the United States and Canada and high media*

attention" (p. 266) the same media attention demanded action in a timeframe shorter than that in which it is possible to do a credible needs assessment.

As a result:

"Haiti was a media-driven emergency. Harrowing images compelled action. Many donors attempted – insofar as possible in the immediate aftermath of such a major disaster – to base their funding on needs assessments. At the same time, many feel that major donors felt impelled to act before they necessarily had sufficient information." (p. 161)

One example does suggest that a recommendation made in an evaluation was quickly implemented.

- Three months after the earthquake Grünewald, Binder and Georges (2010, p. 63) recommend that *"In emergencies with intense media coverage, OCHA must provide the HC with a professional press spokesperson who can handle relations with the press."*
- Three months later the Interagency Standing Committee (2010, p. 9) report that *"Within the first weeks following the earthquake, the OCHA office was strengthened to provide dedicated support to the humanitarian community on… media outreach."*

If both these statements are accurate then OCHA may have already embarked on the recruitment of the personnel at the time the earlier evaluation was being written. Alternatively a cause and effect relationship may be inferred in which case this may be an example of the utility of 'real time evaluations' in identifying rapidly implementable solutions to current problems.

The use of local media in facilitating the participation of beneficiaries is highlighted by Interagency Standing Committee (2010) *"Strengthening community participation, through support to local media… will be essential, and communities must be supported to become active partners in the decision-making process based on informed choice".* (p. 19) This is an almost direct repeat of the recommendation made by Bennett et al. (2006, p. 15) in his evaluation of the response to the Indian Ocean tsunami. It therefore appears that 6 months into the humanitarian response the recommendation of the TEC had not been heeded and this method of communication had been insufficiently utilised.

The technological advances that had taken place during the 5 years since the Indian Ocean tsunami had, however, made new methods of communication widely available to both aid workers and the public at large. Interagency Standing Committee (2010) report an

> *"example of innovative communication techniques enabling the population to voice their concerns is crisis mapping via www.ushahidi.com which allows users to submit eyewitness accounts or other relevant information from disaster zones via e-mail, text or other online media. The international humanitarian community needs to learn from such initiatives and develop a robust strategy to enable effective dialogue with the affected people"* (p. 25)

The use of such innovative techniques was, however, not without problems and rapid data collection methods need a rapid mechanism for publication if they are to be useful. In the same way, as with the rapid needs assessment, it was delays in publication that limited the usefulness of the new technology.

> *"Novel information communication technology was used in the Haiti earthquake response including social media, crowd sourcing and user-generated content of assessments including mapping. However, serious delays in collating and sharing information on humanitarian agency activities were attributed to poor prioritisation of information sharing"* (Patrick 2011, p. 9)

Delays in sharing information are likely to be a continuing problem in the humanitarian world where caution is exercised in releasing information that may reflect in some way badly on either the host government or on the humanitarian organisation. The political considerations within United Nations agencies in particular are likely to continue to cause problems in this area.

The presence of the media close to emergency humanitarian action, delivered in far from ideal circumstances, can always have negative consequences. Grünewald, Binder and Georges (2010) report that, *"In particular, the airdrop of ready to eat meals, where hungry people had to fight for food rations, and its coverage in the media, were strongly criticized."* (p. 41) Here the criticism came from the local people and local NGOs, both as to the way the distribution was undertaken (which may have been necessary if the area was inaccessible by other means) and to the way it was covered in the press which was considered to be exploitative of the desperate situation of the affected population.

b.   Military

The United Nations Stabilization Mission in Haiti (MINUSTAH) was established on 1 June 2004 by United Nations Security Council (2004). Interagency Standing Committee (2010) report that:

> *"MINUSTAH… prior to the earthquake included a military component of up to 7,803 troops, 2,136 UN police, 464 international civilian staff, 1,239 local civilian staff, and 207 UN Volunteers. The capacity of MINUSTAH to respond to the earthquake was severely affected by the large-scale loss of life within its own institution, with over 100 staff killed and many more injured."* (p. 10)

In the same account Interagency Standing Committee (2010) also reports that after its mission had been extended by (United Nations Security Council 2010) both to increase the force level and to support the humanitarian response, MINUSTAH made its logistics resources available to the humanitarian community as well as its human rights component; the latter *"focusing on protection concerns, and using established contacts with local authorities and civil society networks to help identify beneficiaries and engage with local actors". (p. 10)*

MINUSTAH was not, however, the only military presence and *"Twenty-six countries, including Argentina, Canada, France, Russia, the United Kingdom, and the US, provided significant military assets in support of the earthquake response"* (Interagency Standing Committee 2010, p. 10)

Interagency Standing Committee (2010) justifies the use of military support to humanitarian aid in a situation where the humanitarian organisations are overwhelmed by the needs *"Arguably, the humanitarian imperative required the humanitarian community to involve other actors appropriately, including military forces"* (p. 26) and cites the Oslo Guidelines (United Nations Office for the Coordination of Humanitarian Affairs 2007) as giving necessary information regarding the *"considerations to be taken into account and the conditions under which this (involvement of military forces) should happen"* (p. 26)

Interagency Standing Committee (2010) indicates one of the reasons why this was not always done, *"Humanitarians required the support of the military in facilitating the transport and distribution of assistance, but were reluctant to risk undermining the humanitarian principles so central to their modus operandi by engaging too closely".* (pp. 1-2) It must be concluded that the recommendations made by Bennett et al. (2006, p. 16) had either not been implemented or had proved ineffective. The situation may, however, have improved over time even if some humanitarian organisations remained reluctant. Patrick (2011) writing some time later reports that *"Most agencies worked well with the US military, working against established protocols on the humanitarian community's engagement with military assets. However, a number of agencies were reluctant to work too closely with the military."* (p. 7)

c.   Private Sector

As in the response to the Indian Ocean tsunami, the private sector was ignored and this resulted in opportunities being missed and inappropriate aid being delivered. The criticisms of the response to the earthquake in Haiti are even more robustly made and recommendations appear more strongly emphasised.

> *"The Haiti humanitarian response operation missed some prime opportunities to work more closely with private sector actors, and lessons must be learned and acted upon in this regard. In a large-scale sudden-onset disaster such as Haiti, which is very visible in the media, it is inevitable (and indeed commendable) that the private sector would wish to engage and offer support. The humanitarian community needs to understand better the priorities and interests of this sector … In Haiti; humanitarian actors received an estimated $70 million in offers from the private sector. Many were pro bono, and many may have been unsuitable, but most clusters were unable to respond positively to these offers as bureaucratic systems and procedures for receiving and utilising such support had not been pre-established. This is an important lesson to learn for future responses, and more needs to be done between emergencies to strengthen this critical and relatively unexplored partnership."* (Interagency Standing Committee 2010, pp. 27-28)

That the private sector should be referred to above as a *"critical and relatively unexplored partnership"* (p. 28) five years after the response to the Indian Ocean tsunami emphasises how little had been done in response to the recommendation made by Bennett et al. (2006). The United Nations Secretary General issued a bulletin (United Nations 2006) regarding the acceptance of pro-bono goods and services which entered into force on 1 March 2006. There should have been adequate time for these guidelines and the recommendations of the TEC to have been incorporated into the planning and implementing methodologies of humanitarian organisations responding to the 2010 earthquake in Haiti.

d.   Affected Country Governments

In the evaluation of the response to the Indian Ocean tsunami there was little reference to cooperation with affected country governments possibly because the governments were, even where involved in internal conflict, relatively strong and stable and their capital cities had not suffered as a result of the disaster. The situation in Haiti was completely different. In Haiti, where the capital city (Port au Prince) had suffered major destruction by the earthquake, Patrick (2011) refers to the response as being made *"more difficult by severe underlying vulnerabilities that existed in Haiti including systematic poverty, fragile governance, insecurity and the constant threat of natural disasters."* (p. 2)

As a result the tendency was increased for aid operations to be conducted outside government structures. DARA (2010) reports

> *"The UN Assistant Secretary-General of Peacekeeping Operations has sympathised with the government's post-earthquake frustrations, noting that the international community has a long history of weakening the national government by working with outside organisations: we complain because the government is not able to (lead), but we are partly responsible"* (p. 164)

It concludes *"Donors could have done more to promote government co-leadership of clusters"* (p. 165)

A year later DARA (2011) laments *that "Many of the lessons from previous major disasters were not applied. Donors should have done more to ensure Haitian authorities and civil society organisations were better integrated into the response and recovery."* (DARA 2011, p. 267)

Patrick (2011) is even more critical, writing:

"Largely unfamiliar with humanitarian natural disasters in urban areas and compounded by poor contextual understanding of Haiti's society and economy and of the capacity of key stakeholders, the humanitarian community's reaction was a classical response: self-contained, working outside government systems and reliant on imported material and personnel, supporting displaced individuals in internally displaced persons camps with food and non-food assistance." (p. 3)

Interagency Standing Committee (2010) recommends that

"In a disaster response, the affected Government, regardless of capacity and resource constraints, must be empowered to play a central role in the coordination and leadership of the humanitarian operation, and there is a need for more strategic thinking at the global level on how best to achieve this more systematically." (p. 26)

Clearly there was an assumption that the earthquake and resultant damage had rendered the Haitian government powerless and incapable of playing a role in the humanitarian response. This may, in some cases, have promoted the use of the oft repeated mantra of the government is unable or unwilling to justify immediate, unilateral and independent humanitarian operations outside government structures. The assumption of governmental incapacity may, however, have been unjustified. Interagency Standing Committee (2010) report that: "*On 15 January* (three days after the earthquake) *the Government established 6 working groups… to coordinate efforts in the sectors of health, food aid, water distribution, fuel and energy, reconstruction, and safety for temporary shelters*".(p. 8) Patrick (2011) reports that, *"Within days the government had made some important steps in resuming some core functions: making fuel available, repairing two of the four damaged electric plants, and reopening banks and paid public sector workers soon after"*. (p. 5) Humanitarian organisations should accept as an imperative the need to work closely with governments following natural disasters. Patrick (2011) concludes *"Identified weakness in Haitian government and civil society capacity should have highlighted, not negated, the need to work through and empower government to promote long term recovery"*. (p. 5)

6. **Accountability**

Indian Ocean tsunami (2004)

1. Information Flow and Reporting

1.   The recommendations on information flow (Tsunami Evaluaton Coalition 2007, pp. 4-5) focus on the methods and use of information from the responding agencies to the communities with which they are working. The dedication of staff to this purpose and the use of public meetings, media, newsletters and posters are suggested. The purpose of communication is to facilitate the community in participating in planning and implementing the response. The need for improved communication with women is emphasised as being particularly important in this context.

The Tsunami Evaluaton Coalition (2007, pp. 4-5) also emphasises the need for transparency with regard to financial tracking and reporting as well as the public availability of evaluations and proposes the use of conditionality of local tax exempt status as a mechanism to ensure that transparency requirements are fully complied with.

b. Consultation and Community Control

Also under the recommendations on accountability (Tsunami Evaluaton Coalition 2007, pp. 4-10) are a series of far reaching recommendations focussed on the fundamental reorientation of the international humanitarian community *"from supplying aid to facilitating communities' own relief and recovery priorities"* and should *"aim to empower affected people to articulate claims, demand accountability and to make their own choices."* (p. 5) It is also recommended *that "during disasters, the control of resources should be vested in local actors"* and that *"response actors should learn to stand back (or get out of the way) when they do not have the capacity or the endurance to understand or support people's own recovery efforts"* (p. 6). The use of cash as aid is encouraged as a means to allow people to assess and prioritise their own welfare needs. (p. 6) The importance of the representation of women *"claim holders"* and the provision for the needs of *"poorer and marginalised groups"* is emphasised. (p. 6)

While these recommendations are, in some cases, surrounded by caveats regarding *"culturally sensitive and context-specific approaches "* (Telford, Cosgrave & Houghton 2006, p. 113) they clearly go beyond the normal understanding of

"participation of the beneficiary community" and do indeed recommend *"a fundamental re-orientation"* of the humanitarian community.(p. 110)

c. Rights and Protection

The recommendations emphasise that a process of education on human rights should be incorporated into programmes which should strictly adhere to human rights conventions during implementation (Bennett et al. 2006, p. 75) and should include measures to protect the human rights of vulnerable groups. Cost should not be a barrier to the inclusion of the most marginalised whose rights need to be protected at all times. (Scheper, Parakrama & Patel 2006, p. 12)

d. Upward Accountability and Reporting

The Upward Accountability recommendations are based around two important concepts:

> i. Documenting the scope and magnitude of the local response systems and *"the role of remittances in supporting local response"* (Tsunami Evaluaton Coalition 2007, pp. 8-9)

> ii. The development *of "improved mechanisms to track how the 'humanitarian dollar' flows from the tax payer or contributing citizen to the beneficiary"* with an emphasis on the *"transaction costs and value added (or subtracted at each layer"*. (Tsunami Evaluaton Coalition 2007, p. 9)

Most of the recommendations on monitoring focus on the need for common and consistent accounting definitions and the development of a common reporting format which agencies can use for reporting simultaneously to a group of donors. (Tsunami Evaluaton Coalition 2007, p. 9) This is a reflection of the consequences of the multiplicity and strongly held independence of humanitarian organisations and of the sovereignty of donor governments, each reporting to its own public and parliamentary system.

Another recommendation under this heading, compulsory reporting, using a standardized system on the number and cost of visiting delegations, also responds to the problems caused by the multiplicity of responding organisations and donors.

The use of the Good Humanitarian Donorship Initiative is proposed as a possible route for standardisation. (Bennett et al. 2006, p. 15)

Haiti earthquake (2010)

    a.   <u>Information flow and reporting</u>

There are clear indications that there was a lack of information flow in the response to the earthquake in Haiti. DARA (2010) observes that *"There is no clear communication from either the government or many international actors as to what services camp residents can expect or what long-term shelter plans are being developed."* (p. 163) The same source also reported that *"the process of securing funding is characterised by a near-total exclusion of Haitian social actors".* (p. 163) Grünewald, Binder and Georges (2010) under a recommendation headed *"Improve communication with the affected population and ensure they are better informed"* recommends that *"Proactive communication with the population should also be used for expectation management purposes."* (p. 58) They highlight the need for the training of humanitarian workers, international and local in participatory responses. Far from the fundamental reorientation of the humanitarian community and the high level of beneficiary participation recommended in the evaluation of the response to the Indian Ocean tsunami, the effected population in Haiti were not even being informed of decisions that had been made without their participation.

    b.   <u>Consultation and community control</u>

It does not seem to have been a case of simply overlooking the need for participation, exclusion from participation is also reported. The IASC evaluation (Interagency Standing Committee 2010) is highly critical of the approach taken and quotes Grünewald, Binder and Georges (2010) saying that *"the affected population was largely excluded from the design and implementation of the response"* (p. 11) It recognises that similar observations have been made in regard to previous disasters but indicates that this was both surprising and damaging in the Haiti context:

> *"What is of more concern in Haiti is that the beneficiaries were easily accessible; there were no conflict or significant security concerns, few*

insurmountable logistics or linguistic barriers. And yet little dialogue with the affected community actually took place – and this in a context where humanitarians had relatively little experience of a problematic urban environment and had much to learn." (Interagency Standing Committee 2010, p. 25)

In his evaluation of the response, Patrick (2011) observes: *"Even the most devastated communities and governments retain capacities."* (p. 3) He observes that destruction of physical and material infrastructure does not in itself destroy communities who retain *"strong relationships, personal skills, organisational abilities, important norms and values, effective leaders and the ability to make decisions."* (p. 3) He recommends that humanitarian responders should *"slow down to allow meaningful engagement of community and civic leaders in the assessments who will add significantly to the quality and timeliness of results"* (p. 3)

c. Rights and protection

With regard to the need for improved communication with women, recommended in the evaluations of the response to the Indian Ocean tsunami, the response to the Haiti earthquake is highly criticised. DARA (2011) reports that *"Gender was not given the attention it deserved. Many donors and humanitarian organisations seemed to consider the needs so overwhelming that there was no time to address gender".* (p. 274)

d. Upward accountability and reporting

With regard to donor reporting and tracking of funds, DARA (2010) repeats the recommendation made following the response to the Indian Ocean tsunami that *"donors should encourage simpler, compatible reporting formats"* (p. 160) and DARA (2011) observes that *"Tracking aid flows was even more complicated by the huge number of private donors"* (p. 272) again a similar observation to that made some five years previously.

Referring to transparency and accountability DARA (2010) concludes:

"There is evidence that there have been too many actors, unclear communication, different priorities, lack of transparency on total

disbursements, little emphasis on participation and fostering ownership of Haitians in response planning and little promotion of a culture of accountability towards beneficiaries" (p. 167)

7. Linking Relief, Rehabilitation and Development (LRRD), Disaster Risk Reduction and Early Warning

These three topics will be considered together as they all require action outside the immediate lifesaving relief operation following a disaster. For many years a gap was identified between relief and development (Moore 1999) in which relief efforts continued and the development process necessary for beneficiaries to become once again self-sufficient were not put in place. Various terminologies have been used to describe the process that should eliminate this:

- Relief to development continuum (Borton 1994)
- Linking Relief, Rehabilitation and Development (Macrae et al. 1997)
- Repatriation, Reintegration, Rehabilitation and Reconstruction (4Rs) (United Nations High Commissioner for Refugees 2003)
- Mind the Gap (O'Dempsey & Munslow 2009)

In addition to linking into the rehabilitation and development phases of the recovery from a disaster, the relief operation also needs to be designed link to actions that will prepare for recurring natural disasters, limiting their effects and ensuring a future response capacity.

> *"The utilisation of development funds for disaster risk reduction has a high cost-benefit return. Many studies have indicated that disaster risk reduction is highly cost-effective: a dollar invested in disaster risk reduction can save two to ten dollars in disaster response and recovery costs."* (International Federation of Red Cross and Red Crescent Societies 2007, p. 1)

Whereas limited early warning of Tsunamis is possible that of earthquakes is not. However, disaster risk reduction activities can be carried out under a development programme to mitigate the effects of both these disaster risks.

Indian Ocean tsunami (2004)

de Ville de Goyet and Morinière (2006) propose that *"A pre-determined percentage of all future relief funding should be put aside for region wide preparedness for future disasters"* (p. 63) and that joint national/international assessments should be undertaken prior to emergencies happening in disaster prone countries. Telford, Cosgrave and Houghton (2006, pp. 115-116) recommend that in disaster prone countries host authority capacities should be mapped, institutions to manage disaster preparedness and response should be strengthened and that both international agencies and national governments should demand that these activities are adequately supported.

Christoplos (2006) acknowledges that links between relief and rehabilitation have been achieved but recommended *that "more consideration needs to be given to reducing risks of natural disasters and anchoring such strategies within national structures for social protection."* (pp. 80-81) Bennett et al. (2006) is rather more pessimistic observing that, *"If the build back better objective is to mean anything a more comprehensive strategy for sustainable livelihoods is required by those currently engaged in reconstruction"* (p. 61), recognising that appropriate infrastructure requires a sound economic base as part of a reconstruction programme.

Scheper, Parakrama and Patel (2006) emphasise the need for stable, long term partnerships in the response to natural disasters, *"The international community should have clear partnership strategies from the start in order to avoid glitches during the transition to recovery."* (p. 47)

Haiti earthquake (2010)

Largely linked to the virtual exclusion of the Government of Haiti from the response to the earthquake (referred to above), the evaluations reviewed are universally critical of the linking of relief, rehabilitation and development (LRRD) in the aftermath of the disaster. The increasing level of frustration of evaluators with the length of time it has taken for recovery to be initiated is obvious, as is the lack of impact of the newly titled 'real time' evaluations.

Already, just three months after the earthquake Grünewald, Binder and Georges (2010) highlight the need for the early inclusion of LRRD in the response strategy:

"In a sudden-onset disaster such as the earthquake in Haiti, what is known as early recovery often starts the day after." (p. 25) Therefore, early recovery is a component of almost every humanitarian sector rather than a distinct phase of its own. Three months later the Interagency Standing Committee (2010), however laments that *"The role of early recovery in any crisis, but particularly one on the scale of Haiti, is crucial, but continues to be somewhat misunderstood and therefore under-resourced."* (p. 28)

Writing a year after the earthquake DARA (2010) indicates increasing frustration with the implementation of an LRRD strategy in Haiti

> *"However, some are expressing concerns about the slow pace of recovery planning. The Brookings Institute warned in September 2010 that "the recovery process is not going well and reconstruction has barely started... recovery efforts on the ground have been slower than usual – slower than for the 2004 tsunami or the 2005 Pakistan effort"* (p. 166)

The same report may appear contradictory where it states *"Looking prematurely towards recovery, donors have been slow to acknowledge the on-going humanitarian crisis and mounting evidence of failure to provide adequate shelter or protection for the 1.3 million homeless displaced".* (p. 160) In fact, it is a statement that both relief and recovery must continue simultaneously, not a continuum, not a transition but meeting two immediate needs in the wake of a natural disaster.

Patrick (2011) writing 18 months after the disaster concludes that *"Recovery strategies should be articulated from day one and integrated into humanitarian programming from the start"* (p. 4) He reflects on the human cost of ignoring this important principle *"It was reported that much of the later negativity surrounding the response emanated from Haitians trapped in between dependence on humanitarian aid on a daily basis and the desire to recover livelihoods lost in the earthquake."* (p. 6)

DARA (2011), issued 2 years after the earthquake, sound almost desperate to communicate the LRRD message and close to declaring failure. *"Two years after the disaster, long-term recovery efforts are still inadequate."* (p. 226) *"Many of the lessons from previous major disasters were not applied. Donors should have done more to ensure Haitian authorities and civil society organisations were better integrated into the response and recovery."* (p. 267)

Referring to comments of those interviewed for the report DARA (2011) states:

> *"Organisations interviewed reported that support for the transition from relief to early recovery and longer-term development was lacking. Many donors preferred to support the emergency relief phase solely. 'Now there is a gap between emergency and rehabilitation,' affirmed one interviewee. 'It is very difficult to get funding for Haiti once the emergency has passed. Donors are not interested in funding rehabilitation and reconstruction,' noted another."* (p. 274)

Finally, reflecting on the repeated cycle of disaster and relief experienced in Haiti over a long period of time (DARA 2011) concludes

> *"Given the experience from the past, donors should have actively planned and engaged in creating more space for transition, development and humanitarian planning to be integrated into a long term vision that would have focused on building resilience and capacities of the Haitian people, civil society and government authorities. A clearer focus on how donors would support and facilitate a transition from relief to recovery to development (LRRD) and integrate longer term disaster risk reduction into plans was largely missing, and donors could have done much better at working with their Haitian government counterparts to achieve this.* (p. 278)

> *"Donors have largely missed the opportunity to integrate the response to previous disasters in the country to build local response and preparedness capacity, and have neglected longer term disaster risk reduction and longer-term recovery and resilience measures in the current recovery efforts"* (p. 267)

The development of this topic in the evaluations of the humanitarian response to the earthquake in Haiti indicates not only a lack of learning from previous disaster responses by the international community but also indicates a level of ineffectiveness in the first large scale application of the techniques of real time evaluation.

8. **Ownership**

In this research the term "ownership" is taken to indicate the involvement of direct beneficiaries and/or local civil society in the planning and execution of aid and development actions, as distinct from the involvement of the government of a disaster affected country.

Indian Ocean tsunami (2004)

The recommendations of the evaluations of the response to the Indian Ocean tsunami focus around the fundamental reorientation of the manner in which aid is delivered. This reorientation should include a move *"from supplying aid to facilitating communities"* (Telford, Cosgrave & Houghton 2006, p. 110) a *"focus on affected peoples' priorities rather than the institutional or bureaucratic preoccupation"* (of international agencies) (p. 111) and *"should aim to empower affected people to articulate claims, demand accountability and to make their own choices".* (p.115) As a key step to the realisation of these goals the authors recommend that, *"The international agencies should share information about their systems and practices with the affected population so that they can also participate in planning/programming."* (p. 119) Clearly the *"fundamental reorientation"* recommended necessitates not the participation of the affected population in the projects designed and owned by the humanitarian organisation, but rather a participation of the humanitarian organisation in projects designed and owned by the affected population. The essence of these recommendations is a fundamental change in ownership of humanitarian projects. Scheper, Parakrama and Patel (2006) express these principles as *"a need to rethink the end goal of humanitarian assistance and move from a service-delivery approach to a capacity-empowering framework".* (p. 44)

The use of direct cash grants for aid delivery is seen as one fundamental requirement for the realisation of more ownership by the affected population. de Ville de Goyet and Morinière (2006) recommend

> *"Empower the affected individuals or families to assess and prioritise their own welfare needs by using cash subsidies whenever possible… The need for thematic assessments would be considerably reduced if, when possible, the affected people were given the financial means to decide whether they want a better shelter, a boat, food or any other welfare item brought at high cost by expatriates. This approach would go a long way toward compliance*

with the Sphere principle of 'respecting the dignity of victims' in countries with active market economies, such as those affected by the tsunami." (p. 64)

Haiti earthquake (2010)

The recommendations made in the evaluations of the humanitarian response to the Indian Ocean tsunami appear to have 'fallen on deaf ears' five years later as the international community responded to the earthquake in Haiti.

Three months into the response Grünewald, Binder and Georges (2010) are pointing out that *"Communication with the affected population is a first crucial step towards more inclusive humanitarian assistance."* (p. 10) and to *"Go Urban" "moving from an individual to a community-based approach and by engaging in close cooperation with local authorities, civil society groups and development actors to provide the services needed together"* (p. 11). They identify reasons why local people were excluded including that *"humanitarian actors often assumed that after such a major shock there would be no local capacity to respond"* (p. 42) and *"Humanitarian actors should use French and Creole as their working languages"* (p. 57) rather than English. They also encourage humanitarians to *"Get participatory: Participatory approaches and consultation with the population and local institutions should be seen as a must, not as a constraint."* (p. 60) This was certainly an early warning of problems to come but a far cry from the almost revolutionary recommendations that followed the response to the Indian Ocean tsunami.

Assessing the situation six months after the earthquake Interagency Standing Committee (2010) makes a number of important points, including:

a.   In the immediate aftermath of the earthquake, the initial response was led by the Haitian population themselves, and countless live rescues and support to the injured were provided by local communities. *"Neighbours, friends, family, and strangers helped each other, saving thousands of lives, while the government and the international community mobilized their response and tried to overcome initial constraints".* (p. 7) It should be noted however that local communities also faced major constraints in their immediate response capacity, including the absence of electricity, widespread debris, limited first aid skills, and extensive trauma, which was compounded by a pervasive fear of additional aftershocks.

b.   Haitian civil society organisations (CSOs) were very active in providing immediate assistance to the affected and displaced population, despite the fact that they themselves were also extensively affected. *"Haiti had a vibrant civil society prior to the quake. Locating survivors among Haitian community aid organizations and supporting their efforts should also be an important component of the relief effort".* (p. 7) The extent of the presence of local and national non-governmental organisations (NGOs) in Haiti is underscored by the fact that the on-line directory of CSOs in Haiti already contained over 800 organisations (both national and international) prior to the earthquake.

Interagency Standing Committee (2010) concludes that there is:

> *"a pressing imperative to identify ways to engage better with affected Governments and civil society partners. The global humanitarian architecture must be critically reviewed to ensure that it is not implemented in such a way as to preclude such partnerships which are critical to the most effective response."* (p. 2)

In this last reference something of the sense of urgency expressed by Telford, Cosgrave and Houghton (2006) appears but once again stops far short of their call for a fundamental reorientation.

One year into the response and DARA (2010) continue to highlight the same deficiencies *"The cluster system was weakened by… failure to involve the Haitian state or civil society"* (p. 160) and the *"near-total exclusion of Haitian social actors"* from the process of securing funding." (p. 163)

Eighteen months into the response Patrick (2011) reflects judgementally on the response of the humanitarian community as a whole:

> *"Largely unfamiliar with humanitarian natural disasters in urban areas and compounded by poor contextual understanding of Haiti's society and economy and of the capacity of key stakeholders, the humanitarian community's reaction was a classical response: self-contained, working outside government systems and reliant on imported material and personnel, supporting displaced individuals in internally displaced persons camps with food and non-food assistance"* (p. 4)

As one of his emerging lessons Patrick (2011) identifies the need to *"better understand and support social and economic resilience"* which he concludes will

require *"a better engagement with people affected by disaster to understand their strategies"* if the humanitarian response is not to be socially harmful and counterproductive. (p. 7)

After two years of the response DARA (2011) simply reports that *"throughout the entire relief and recovery responses, Haitian civil society was largely marginalised and kept out of sight by the donors and the Haitian Government."* (p. 270) Possibly the only surprising aspect of this conclusion is the share of the blame being placed on the Haitian Government which, as has been discussed above, had itself been largely marginalised by the international community.

Conclusions

Some of the more extreme conclusions of the evaluations of the humanitarian response to the Indian Ocean tsunami may be ideologically justified and fully in keeping with, if not even an extension of, the principles of accountability to beneficiaries. (Humanitarian Accountability Partnership 2010) However, they were (and perhaps still are) too revolutionary to be considered practical for the humanitarian community to implement. The situation in Haiti after the earthquake was certainly not the same as that in the countries affected by the Indian Ocean tsunami, as has been discussed above, but as some of the evaluations reviewed indicate, the capacity of the government, civil society and the affected population was underestimated, underutilised and often ignored. It appears that it was convenient for the international humanitarian community, in its inexperience of urban disasters and ignorance of Haitian society, to continue an immediate lifesaving methodology far beyond any timeframe in which it may have been rationally justified. It is clear, however that this approach had a definite cost both to the credibility of the humanitarian organisations and to the Haitian people as opportunities for LRRD, DRR and sustainable livelihoods appear to have been squandered.

As an overall conclusion to this chapter, however, it is hard to improve on those made in two of the evaluations of Haiti reviewed above. DARA (2010) conclude that *"It is disappointing that many relevant recommendations from the Tsunami Evaluation Coalition and those in the last Human Response Index (HRI) report on Haiti appear to have not been heeded in the earthquake response"* (p. 166) while Patrick (2011) concludes:

"The humanitarian community needs to better explore how it can best learn from and implement previous lessons. The situation in Haiti and experience from other disasters tell us that lessons cannot simply be learned but must be continuously studied, revisited and reflected upon. Articulate and implement a robust communication strategy to ensure that key stakeholders are aware of previous lessons learned" (p. 6)

Having demonstrated that the assertion made by Ramalingam, Scriven and Foley (2009) that *"evaluations do not tell us anything new"* has a substantial basis in fact, this research will, in Chapter 6, focus on the subjective influences acting on the evaluation of humanitarian action.

Chapter 6 - Baseline Case Studies: evaluations undertaken between 1993 and 1999

Introductory summary of the chapter

In this chapter the researcher analyses nine case studies edited by Wood, Apthorpe and Borton (2001b), of evaluations undertaken between 1993 and 1999 and identifies influences on the work that were not controlled either by policy or methodology and/or which in some way limited the objectivity of the evaluation process. The influences identified are presented in an analytical framework which was later used in the analysis of research material gathered from interviews.

Introduction

Wood, Apthorpe and Borton (2001b) edited a compilation of the experiences of members of teams who had undertaken nine evaluations over the period 1993 to 1999 and presented them as case studies. In the forward to this publication MacDonald (2001) speaks of *"opening up a process that has for far too long remained hidden"*. (p. xx) Wood, Apthorpe and Borton (2001a) asked contributors to show *"What struck them as influential on the way they did their evaluations and what influenced the results they obtained"* with the objectives to "*sensitise evaluation practitioners, managers and students*" and to *"stimulate thought and awareness that can be applied at all times and in all places".* (p. 190)

Wood, Apthorpe and Borton (2001c) introduce the case studies as answering several questions:

- *"How are such studies carried out?*
- *Who are the evaluators and how are they recruited?*
- *What methods do they use?*
- *How do they arrive at their conclusions and recommendations?*
- *How do they cope when the agencies being evaluated press for their report to be less critical?*
- *How effective are they in identifying what went wrong, what went right and what lessons should be learned from the experience."* (p. 1)

The researcher observes that in providing answers to these questions the case studies reveal objective elements largely related to the policy and methodology of the evaluations. They also reveal subjective elements that cannot be covered

adequately, or in some cases controlled, by methodology or policy in the course of the evaluations. The researcher also observed that these elements, as well as others, were acknowledged by the authors as having in some way or other compromised the objectivity of the evaluation process. Elements that were identified using the criteria outlined in this paragraph were classified by the researcher as subjective influences and recorded in the analytical framework.

Table 6.1 Chapter headings indicating the case studies reported on by Wood, Apthorpe and Borton (2001b) and the year in which each evaluation was undertaken.

Chapter titles indicating the evaluation undertaken	Year Undertaken
Somalia: Towards Evaluating the Netherlands' Humanitarian Assistance (O'Keefe et al. 2001, pp. 19-38)	1993/94
Exploring the Swedish Emergency Relief Experience in the Horn of Africa (Wood 2001, pp. 39-58)	1994
Evaluating Sida's Complex Emergency Assistance in Cambodia: Conflicting Perceptions (Lindahl 2001, pp. 59-73)	1994/95
Doing Study 3 of the Joint Evaluation of Emergency Assistance to Rwanda: The team Leader's perspective (Borton 2001, pp. 74-101)	1995/96
Mission Possible: Six Years of WFP Emergency Food Aid in West Africa (Apthorpe 2001, pp. 102-121)	1996
Review of the International Federation of Red Cross and Red Crescent Societies' Tajikistan Programme (Wiles 2001, pp. 122-132)	1996/97
A Self-evaluation of my Experience Reviewing Australia's Official Assistance in Response to the 1997-98 Papua New Guinea Drought (Lea 2001, pp. 133-152)	1998
An Experimental and Inclusive Approach to Evaluation as a Lesson-learning Tool: Group URD's Work on the post Hurricane Mitch Emergency (Grünewald, Pirotte & de Geoffroy 2001, pp. 153-170)	1998/99
UNICEF-DFID Joint Evaluation of UNICEF's Kosovo Emergency Preparedness and Response (Telford 2001, pp. 171-189)	1999

Source: Author after Wood, Apthorpe and Borton (2001c)

The Case Studies (See table 6.1) are described by Wood, Apthorpe and Borton (2001c, pp. 13-18) and paraphrases of these descriptions are the source of the summaries below.

Somalia: Towards Evaluating the Netherlands' Humanitarian Assistance

This evaluation was undertaken in 1993 by the Netherlands Ministry of Foreign Affairs together with consultancy groups. The evaluation assessed

the Netherlands assistance to Somalia during the period 1991 to 1993 and explored whether it was possible to evaluate the delivery of humanitarian assistance in a complex emergency. In this context its concerns surrounded the difficulty of delivering the humanitarian response in complex emergency situations and the problems of accessing the projects themselves.

Exploring the Swedish Emergency Relief Experience in the Horn of Africa and Evaluating Sida's Complex Emergency Assistance in Cambodia: Conflicting Perceptions

In the early 1990s the Swedish International Development Agency (Sida), concerned at the way in which the funding of humanitarian action had cut into the overall development budget, decided to evaluate this growing expenditure. Wood, Apthorpe and Borton (2001c) suggest that levels of "*20 per cent or more of the aid program of some bilateral donors*" (p. 10) represented a watershed that triggered a demand for formal evaluations to be undertaken. Three studies were commissioned including the Horn of Africa and Cambodia which each form one of the case studies. The Horn of Africa study was undertaken several years after the projects had finished while the Cambodia study covered six years of assistance. A split in the evaluation team resulted in a majority and minority version of the concluding chapters.

Doing Study 3 of the Joint Evaluation of Emergency Assistance to Rwanda (JEEAR): The team Leader's perspective

This evaluation, consisting of four separate studies and a synthesis study, remained the largest and most comprehensive evaluation of humanitarian action until the TEC evaluation of the humanitarian response to the Indian Ocean tsunami of which Cosgrave (2007) authored the synthesis report. The administration of the JEEAR was complex. The evaluation was funded by the donor community, steered by a committee comprising the UN, Red Cross movement and NGO umbrella organisations and managed by a group of five bilateral donor organisations. This evaluation identified that humanitarian action must be evaluated taking into account the political

situation, that donors used humanitarian action as an alternative to military intervention and was highly critical of the coordination of humanitarian action in complex emergencies.

Mission Possible: Six Years of WFP Emergency Food Aid in West Africa

This evaluation was undertaken by a mixed team, led by independent consultants but also containing staff members of both WFP and a sister agency. The objective of the evaluation was to look at the reality on the ground of the provision of food assistance under conditions of conflict in Liberia. Due to difficulties in access to beneficiaries and the six year timeframe of the projects to be evaluated, the evaluation considered only the policy level of the operations but was able to identify conceptual misunderstandings and make recommendations for future operations.

Review of the International Federation of Red Cross and Red Crescent Societies' Tajikistan Programme

This evaluation was designed to evaluate both current projects and identify lessons that could be applied to future projects, as well as to justify further funding from the International Federation of Red Cross and Red Crescent Societies. The term 'external review' was used to describe this evaluation

> *"perhaps to make the evaluatory nature more acceptable but also because the agenda was clearly to provide clearance for further funding by the Federation to one of its member organisations".* (pp. 15-16)

A Self-evaluation of my Experience Reviewing Australia's Official Assistance in Response to the 1997-98 Papua New Guinea Drought

This case study covers an evaluation in which conflicts emerged between the commissioner and the evaluation team. While the methodology is described as 'commendable' the disagreements between the team and the

commissioner are presented as probably the most memorable aspect of this evaluation.

An Experimental and Inclusive Approach to Evaluation as a Lesson-learning Tool: Group URD's Work on the post Hurricane Mitch Emergency

The evaluation of the response of French NGOs to Hurricane Mitch is covered in this case study. As the title suggests Group URD adopted an 'experimental and inclusive' approach to the evaluation, largely at the request of the agencies working in the field, to include an inclusive, lessons learned approach enhanced by the use of techniques of networking and feedback. While succeeding in this objective, tensions arose between these and the more accountability focus demanded by the donors.

UNICEF-DFID joint Evaluation of UNICEF's Kosovo Emergency Preparedness and Response

This case study covers a joint evaluation between the United Nations Children's Fund UNICEF and the UK Department for International Development (DFID) of projects funded by DFID implemented by UNICEF in Kosovo. The case study focusses on the workings of the joint teams, showing how they worked effectively but also highlighting the problems that developed.

These case studies included some of the earliest evaluations of humanitarian action undertaken. O'Keefe et al. (2001), in their account of the evaluation of the Netherlands humanitarian assistance to Somalia undertaken in 1993/94, indicate that the evaluators had to face the doubts of the commissioning organisation for the evaluation as to whether "*humanitarian assistance could be evaluated.*" *"At the early stages of discussing the possibility of evaluating humanitarian aid there were severe doubts amongst IOB's staff about the feasibility of such an undertaking*. (pp. 19-20) Two major concerns were expressed

"Would the focus of the evaluation be largely on the implementation of activities rather than the results?... and the ethical issue that evaluators

might hinder aid delivery, that evaluators would 'get in the way' of those providing the humanitarian assistance." (O'Keefe et al. 2001, pp. 20-21)

Having observed that evaluation had been common practice in the field of international development programmes since the 1960s, Wood, Apthorpe and Borton (2001c) provide an analysis of the reasons why the *"application to the field of humanitarian aid… was comparatively slow"*. (pp. 10-11) They cite four basic reasons:

Attitude Humanitarian organisations resented the idea that people who had not been involved in the extremely difficult circumstances of the emergency should come and potentially judge their best efforts as not being good enough.

Technical difficulties In the *"highly dynamic situations involving rapid change"* (p. 11) under which humanitarian assistance is delivered, baseline data, comparable to that available in development situations, is not available.

Practical *"The pressure to respond rapidly and the difficulties in the operational context often result in poor and incomplete documentation by the agencies involved"* (p. 11)

Methodological The complex nature of humanitarian operations, often covering many sectors and including many organisations *"creates organisational barriers to initiating and undertaking evaluations."* (p. 11)

Wood, Apthorpe and Borton (2001c, pp. 11-12) indicate that these difficulties were *"gradually addressed and overcome"* over a period of thirty years, largely as a result of the increasing amounts of government money being applied to humanitarian action, leading to an increased demand for an accountability similar to that applied to development funding. While quoting some isolated studies undertaken in the 1970s, Wood, Apthorpe and Borton (2001c) identify the evaluations undertaken into the response to the famines in Ethiopia and Sudan in the mid 1980s as the time when *"evaluation procedures , as applied to development assistance, began to be applied to emergency assistance".* (p. 12) However, it was the *"1991 upturn in humanitarian expenditures"* (p. 12) that by 1993 resulted in "*an increased number of evaluations".* (p. 12)

Wood, Apthorpe and Borton (2001c) present the nine accounts as *"case studies"* that

> *"represent a sample from within the 1990s 'boom' in humanitarian evaluation. By virtue of the seven years spanned, the sample may provide glimpses of whether the 'art' of humanitarian programme evaluation has evolved during the decade and, if so, the ways in which it has evolved."* (pp. 12-13)

This research looks at the developments in the period 2000 to 2012 with particular reference to the more subjective factors reported.

In commissioning the chapters the editors (2001c) asked the authors , *"to reflect on their experiences and to discuss, in a primarily chronological order, the experiences they felt were most important and influential"* so as to "*help bring alive for the reader the debates and experiences the authors are presenting."* (p. 17). It is the analysis of these *"debates and experiences"* from the period 1993 to 1999 that the researcher has taken as the baseline data for this research into the subjective factors that influence the evaluation of humanitarian action and the learning process that should follow

An analysis, by the researcher, (see Table 6.2) of the experiences of the authors in Wood, Apthorpe and Borton (2001b) revealed a range of subjective factors reported by practitioners who participated in nine evaluations. These subjective factors can be classified under a number of headings, many of which are repeated by two or more of the authors, as is detailed under the corresponding headings below.

Several of those who contributed to the publication of the nine case studies used language that clearly indicates the way in which subjective influences pervade the evaluation of humanitarian action. MacDonald (2001) describes the evaluation of humanitarian actions as *"intense experiences, involving considerable judgements about perceived success or limitations of an action"* and of evaluators working under *"considerable time pressure and in the face of high expectations"* (p. xxi) all of which suggest that levels of subjectivity will be introduced to the independent, objective, thorough, process of evaluation.

Lindahl (2001) expresses the doubts of himself and other evaluators of humanitarian action as being *"challenged to think about the nature of the work they*

Table 6.2 Subjective factors identified from the analysis of accounts of evaluations undertaken between 1993 and 1999 reported by (Wood, Apthorpe & Borton 2001b)

	Subjective Factor
1	Subjective influences are perceived as universal in the evaluation of humanitarian action
2	Perceived influence of the sense of urgency surrounding humanitarian action
3	Perceived influence of the attitude, personality and interests of those undertaking the evaluation
4	Inter-organisational and Interpersonal relationships
5	Relationship between the organisation commissioning the evaluation and the evaluation team
6	Relationship between the commissioner of the evaluation and the project being evaluated
7	Relationship between the evaluating team and the project being evaluated
8	Relationship between the evaluating team with beneficiaries and other stakeholders in the project being evaluated
9	Relationships within the evaluation team
10	Interagency relationships in the field
11	Learning from evaluations and the learning cultures of humanitarian organisations
12	Inclusion of multiple objectives of "accountability" and "lessons learned" into one evaluation
13	Influence of the culture of evaluations and the "Western" origins of evaluation
14	The use of the OECD-DAC criteria

Source: Author

 are doing and whether they are really confident about the conclusions they reach." (p. 59) O'Keefe et al. (2001) has an interesting insight into the evaluation process as listening to the "*stories of the action"* told by "*those responsible for the implementation"*, of testing these "*against the beneficiaries' viewpoint"*, of "*the evaluation team being confident that they can tell a story from beginning to end"* and

the recognition that the story "*has to have a purpose, a moral*". They conclude that "*Evaluation of humanitarian aid is the writing of disaster fables.*" (p. 30)

One can comment that the evaluators displayed some lack of confidence in the objectivity of the process as they undertook these early evaluations.

1. **Subjective influences are perceived as universal in the evaluation of humanitarian action**

   a. **Evaluation perceived as an art rather than a science**

Lindahl (2001) observes that "*Evaluation of humanitarian assistance is both an art and a science. There are some 'hard' quantitative 'facts', but there are many impressionistic elements in such endeavours.*" (p. 59) A topic he returns to more dogmatically

   "*Evaluation tends to be more of an art than a science… However my own personal experience is that the art tends to dominate and that the temperaments, judgments and perceptions of the evaluators in the end strongly influence the outcome of the process.*" (p. 73)

Looking to the future he concludes "*But the bottom line is that the art will probably continue to dominate and the personal judgements of the evaluator will determine the conclusions and the recommendations*" (p. 73)

Noting this interesting observation the researcher, in the interviews conducted, looked at whether, in the intervening decade, evaluators considered their work to have become more scientific. (The results of these interviews are analysed in Chapter 7 of this research)

   b. **Perceived interest of politics, money and vested interests in the evaluation of humanitarian action**

   Wood, Apthorpe and Borton (2001c) identify "*bilateral donor organisations (from richer nations) that occupy a very influential position within the overall* (humanitarian) *system*". (p. 4) They observe that the same governments operate diplomatically "*around and alongside humanitarian systems*" and

may be *"seeking the victory of a favoured side or faction, and may involve the use of both covert and overt means".* (p. 7) The researcher observes that the political interests of donors (part of the structure in which aid organisations and evaluators function) limit the principle-agency of aid organisations to implement and evaluators to assess accurately the achievements of humanitarian action.

2. <u>Perceived influence of the sense of urgency surrounding humanitarian action</u>

The general sense of urgency that surrounds the whole process of humanitarian action, in planning, implementation and evaluation was identified in some of the case studies as a factor that introduced an element of subjectivity.

Lindahl (2001) observes from his experience leading a team evaluating Sida's Complex Emergency Assistance in Cambodia undertaken in 1994, *"Furthermore the crisis atmosphere in which the relief efforts often are carried out, and the urgency of life and death during the emergency, make evaluations of impact and effectiveness almost hearsay."* (p. 61) Particularly in the initial response to a humanitarian crisis, the need for action takes precedence and often overwhelms the record keeping and reporting structures.

Evaluations themselves are often driven by short deadlines both to ensure that the evaluation is seen to be *"timely"* and to be able to work in the field with the team that actually implemented the project before they are scattered to a variety of subsequent assignments.

 O'Keefe et al. (2001) in their 1993 evaluation of Somalia experienced a wide range of the problems of working in an environment where a sense of urgency is part of the nature of humanitarian action

> *"We also faced limits imposed by the nature of the phenomena to be examined. Much humanitarian assistance is immediately consumed and its temporary institutions dismantled; many of the people involved both as workers and as beneficiaries move on to other places and cannot be reached; in the fury of action, paperwork is frequently forgotten or inadequately executed so records are poor or non-existent."* (p. 28)

O'Keefe et al. (2001) also identified that the sense of urgency also comes from the commissioner of the evaluation *"The IOB* (an independent evaluation unit of the Netherlands Ministry of Foreign Affairs) *management team wanted to complete the work quickly to meet a deadline to present the report to Parliament"* (p. 35)

Wiles (2001) reports several areas in which the timeframe in which the evaluation was required to be completed influenced the evaluation *process "There was little opportunity in Geneva for discussion about the TOR and methodology before the work began"* and *"methods of working in the field were discussed while we were en route for Tajikistan"* (p. 126): *"it might have been useful to have had one or two more days in Geneva"* (p. 127). "*The schedule sometimes underestimated the time needed for interviews"* and *"it would probably have been useful to have allowed an additional day in each physical location"* (p. 127)

The researcher noted that several aspects of the evaluation process were compressed into timeframes shorter than those considered necessary by the evaluators, for reasons of the transitory nature of humanitarian assistance, political necessities and cost of the evaluation itself. The authors of the accounts of undertaking the evaluations also noted the high turnover of staff involved in the delivery of humanitarian action and the lack of baseline data as factors related to the urgency under which both projects and evaluations were undertaken.

a. Influence of high turnover of staff

Borton (2001) expresses his concern regarding

> *"the degree to which many of the key personnel involved in the initial relief phases in Ngara… in Goma and Bukavu… had already left to work elsewhere… We would need to put considerable effort into tracing and interviewing these individuals in other parts of the world."* (p. 80)

Telford (2001) writing of the evaluation of UNICEF's Kosovo Emergency Preparedness and Response undertaken in 2000 observes that,

> *"More meticulous and involved preparations were definitely needed. Yet, at this stage at least, they would have delayed the evaluation to the point that it might not have taken place. The evaluation had to be carried out rapidly so as to have access to the already reduced presence of key*

UNICEF sources in the region and at central locations (Geneva and New York). Staff were moving on, files (especially invaluable records of email traffic) were being lost, deleted or thrown out and in general memories fading." (p. 177)

The researcher observes that a difficult trade-off between good, adequate preparation and the continued availability of information in the field had a substantial effect on the implementation of two evaluations described.

b. Influence of lack of baseline data and poor data collection

Lindahl (2001) found the *"most important reason for the impressionism and conflicting perceptions in the Cambodia evaluation was the poor database on which to make assessments of the programme."* While recognising that this may be excusable in real emergency situations, he points out that *"much of the Cambodia experience was not really emergency assistance. It was more a long-term rehabilitation-cum-reconstruction phase undertaken in a generally non-crisis situation"* and that *"hope that there will be adequate reporting is one excuse for accepting ambitious TOR."* (p. 67)

O'Keefe et al. (2001) referring to the difficulties of the application of the OECD-DAC criteria to the situation in Somalia indicates that, *"effectiveness was difficult to judge not least because of the lack of baseline data"* (p. 22)

Wood (2001) identifies that *"It was recognised that in most cases we would not have a baseline from which to measure impact."* (p. 45)

Telford (2001) identifies one of the driving forces towards urgency in carrying out evaluations of emergency humanitarian actions as the fragility of records kept in these circumstances *"files (especially invaluable records of email traffic) were being lost, deleted or thrown out and in general memories fading."* (p. 177)

O'Keefe et al. (2001) highlighted the paucity of information available at the start of the evaluation *"The desk study, however, found little documentary evidence of results"* (p. 22) and Wiles (2001), noted *that*

> *"As often happens, no one location (headquarters, regional office, country office) had an easily located full set of programme documents"* (p. 127)

> Apthorpe (2001) reflects that *"Seldom do humanitarian evaluations have the luxuries of reliable (or any) baselines, reasonable needs assessments, up and running surveillance systems and so forth."* (p. 114)

The researcher observes that the lack of baseline data, recognised in six of the nine accounts reviewed constitutes a basic hindrance to objectivity in carrying out the evaluation of humanitarian action.

3. **Attitude, personality and interests of those undertaking the evaluation**

Five of the nine accounts of undertaking evaluations specifically note this as a factor limiting the objectivity of the evaluation of humanitarian actions.

Lindahl (2001) observes

> *"The impressionistic nature of the evaluations makes the results as much dependent on the composition of the evaluation team as on the realities on the ground: send in another team, and the results are likely to be different. As a result, learning is less than that desired and attempts to determine impact and effectiveness are mostly guesswork. In particular the cost effectiveness of the interventions is generally left unexplained.* (p. 61)

> And *"that personal judgements of the evaluators will determine the conclusions and the recommendations. Development assistance is not an academic exercise, but a political activity. As such it is an art of the possible."* (p. 73)

Another aspect of the personality of the evaluator identified by Lindahl (2001, p. 71) is that of

> *"finding a balance between faults and achievements. Some evaluators have a tendency to seek out the faults and identify the problems while others like to highlight the achievements. This is more an issue of temperament and personality than of perceiving reality differently."* (p. 71)

With specific reference to the evaluation team involved in the case study Lindahl (2001) observes *"Some of the team members could be considered stakeholders in the process"* and that as a result *"the evaluation could not avoid vested interests".* In this context he cites *"an apologetic view… of the aid system in Cambodia in general and of the United Nations in particular"*, "*preconceived ideas based more on personal affinities than on facts"* and " a positive view of *"the programmes that Sida might continue to support"* possibly motivated by *"opportunities for future involvement by the team members concerned".* He also cites that the team "*represented the implementing systems… with long term careers in these systems"* and that these affiliations resulted in the institutional assessments becoming *"biased and impressionistic, using different criteria depending on which team member happened to look at which organisation."* (pp. 62-63)

Apthorpe (2001) gives an interesting insight into the type of individual attracted to evaluation teams,

> *"These consultancies, known as 'missions' – a word familiar also in military, evangelist and espionage parlance – offer those who undertake them the adventure of the chase after a big important quarry, shadowy and elusive on the surface, yet solid and permanent underneath. Another motivating factor can be the pull of the drama (or even melodrama); inclusion in the plot as it unfolds and in a cast that, where the atmosphere is right, provides support and camaraderie, particularly in the long hours on the road, when you cannot get the ferry across, run out of petrol and so forth."* (p.104)

Apthorpe (2001) also observes that

> *"Evaluators report mainly on the basis of what they see as normative in the situation concerned. Drawing interpretative conclusions is often closer to drawing a picture than drawing a bucket of water from a well. Matters of style, school of thought, university discipline, gender, cultural background and so forth are as important – perhaps more important – than forensic logic that provides a solely a-technical, a-social and apolitical skill. In reaching conclusion, sensory skills in the divination of programme intent and effect account for much."* (p. 115)

Wiles (2001) who undertook the Review of the International Federation of Red Cross and Red Crescent Societies' Tajikistan Programme in 1996 comments on how the vested interests of members of the evaluation team impacted on the

evaluation process, *"Some of the team members brought approaches to the work that related to their employer's agendas, rather than specifically to the TOR."* (p. 128)   Similarly Wood (2001) referred to the "different personal motivations" of members of the evaluation team which included *"personal applied research interests"* and *"specific strategic commercial interests"* (p. 41)

Wood (2001) reflecting on the similarity between the published findings of the evaluation and the discussions of the team early in the process comments: *"Although, perhaps not noticed so clearly at the time, this seems in retrospect to raise questions about the influence of preconceived ideas on the fieldwork.* (p. 53)

Borton (2001) found the previous employment of one team member to be a problem *"and felt that his four years working in different parts of the UN had reduced his ability to be critical of it; a view that was shared by several other members of the team"* (p. 93)

Wood, Apthorpe and Borton (2001a) after reviewing the accounts of the nine evaluations conclude,

> *"In all cases the vested interests of the evaluators need consideration, along with their ability to place these on one side in order to achieve an independent evaluation. It is important that team members do not have personal or professional interests in the findings that the study reaches. Consultants trying to set up the next job and university based researchers trying to explore an old field area should be avoided."* (p. 197)

If, as described above, the personalities and interests of the evaluators has such a subjective effect on the outcome of an evaluation then it would be expected that great care would be taken in the selection of the member of an evaluation team.

Influence of the selection of consultants.

Apthorpe (2001) from the evaluation of WFP Emergency Food Aid in West Africa undertaken in 1995 observes the influence of the sense of urgency in recruitment of the evaluating team,

> *"Availability is all too often the determining factor driven by a perception of overriding urgency and failure to allow sufficient lead-time for the recruitment process. This threatens – or actually undermines – at the outset the credibility of the exercise."* (p. 104)

Wood (2001) also reports the primacy of availability for the fixed timeframe of the evaluation: *"Probably the key criterion for selection was availability"* and also observes that *"Recruitment never seemed to be openly competitive"* (p. 41)

In their conclusions Wood, Apthorpe and Borton (2001a) focus heavily on the aspect of the selection of consultants. "*Recruiting an evaluation team is a major task… to find people who are trusted and able to do the work to an appropriate standard*" (p. 197), they also highlight the necessity to give *"adequate time so that availability does not become the sole criterion by which people are chosen"* (p. 197).

4. **Inter-organisational and Interpersonal Relationships**

While some part of an evaluation is carried out by research and desk work, the major part involves multiple working relationships. Attitudes, perceptions, trust, suspicions and hidden agendas in these relationships all contribute to subjective influences on the evaluation process. Analysis of the accounts of the nine evaluations revealed several distinct sets of relationships which will be detailed individually.

Perhaps rather surprisingly only a few of these relationships are directly commented on in the conclusions drawn by Wood, Apthorpe and Borton (2001a). Where such comments are made they are referenced in the text that follows.

5. **Relationship between the organisation commissioning the evaluation and the evaluation team**

Perhaps one of the most obvious elements in this relationship is that of consultant-client. Apthorpe (2001) dryly observes *"Having a client is much better than not having a client if you want to make a difference to policy."* (p. 117)

Lindahl (2001) points out *"We all know that too critical review might put at risk the next job, while glowing praise tends to be appreciated by the client."* (p. 71) He describes the dynamics at work

"There is only a certain degree of intellectual freedom for an evaluation if it is to have any impact at all. Too much praise provides little learning and too much criticism tends to create rejection. Finding the right tone within the degrees of

freedom permitted, phrasing criticism in a constructive manner, and avoiding praise for the only reason of assuring the next job, is not science but art."(p. 71)

The relationship therefore affects the independent nature of the evaluation, particularly, as is usually the case, when the evaluation team is selected by the commissioning agency Apthorpe (2001) reports,

> *"Our assignment had been commissioned as an 'independent' evaluation. We were, however, a mixed rather than an independent team, made up of both outsiders and insiders, selected and paid (probably the FAO staffer was seconded) by the commissioning organisation.*
>
> *We were, so to say, authorised by WFP to be independent, without being independent in any other way. Our independent report is usually referred to in the humanitarian literature as 'WFP's evaluation'. Where an independent evaluation is needed, arguably this should best be done under the auspices of an independent body, not the client organisation."* (p. 120)

Unfortunately this is rarely the case.

Contractual terms imposed by the commissioning organisation also impinge on the independence of an 'independent' evaluation. Lea (2001) reporting on his, difficult experiences in Reviewing Australia's Official Assistance in Response to the 1997-98 Papua New Guinea Drought undertaken in 1998, indicates his surprise at being presented with a contract which allowed the commissioner (AusAID) to alter the team's findings *"to suit their interests and requirements",* even to the point where *"the commissioning agency could force the contractor to 'correct' rejected material even if it was factually correct or a legitimate interpretation."* (p. 135)

Lea (2001) describes in some detail the effects these contractual terms and the working relationship they imposed had on the evaluation,

> *"We established a good relationship with him, although he made it clear that he thought we were too direct in our critical comments.*
>
> *We certainly got the impression that AusAID staff did not expect any criticisms of the relief effort. They seemed affronted that outsiders should come in at the end of the exercise and make any comments that did not support the Australian political and bureaucratic opinion that the whole*

exercise had been a great success (a view not shared by many international agencies operating within PNG or by PNG officials). (pp. 143-144)

The relationship apparently continued to be conflictual and became "*the team's major problem.*" The drafting process went through many versions, "*The third draft generated great heat and little light within the section of AusAID mainly responsible for the relief operation*" and "*it was deemed to be in great need of refining*". As a result it took six drafts before the report was forwarded any further within AusAID. The team was informed it should "*dwell on the 'positive attributes' of the relief effort; delete all names; delete most of the end notes that provided examples and detailed justifications; delete comments that 'denigrated' AusAID; and 'make the criticisms more palatable'.*" As a result "*All team members were irritated by the process and thought that the redrafting resulted in too many compromises and the omission of some key criticisms, explanations and caveats.*" (pp. 144-145) In all the report had to be redrafted seven times by the evaluation team before it was accepted by AusAID. Even then the final report was available "*only on a 'restricted' basis within AusAID*" and was "*rewritten by a media consultant who retained in a gentler form the original meaning and intent of the team members. Criticisms were muted and the ''positives' of the operation highlighted.*" (p. 146)

He expressed his concern that "*there was reluctance (with one or two notable exceptions) to accept that there were shortcomings in the aid process and lessons to learn from the Review Experience*" (p. 148) and concludes that

> "*It is not enough for consultancy teams to work hard and be professional, independent, and concerned about accountability, transparency and learning lessons. Teams must also have an understanding of political and bureaucratic sensitivities within their commissioning agency and have the skills to communicate effectively and sympathetically with senior aid administrators.*" (p. 150)

The defensiveness of the commissioning organisation clearly added a range of subjective influences to the evaluation process.

Wood (2001) reports that feedback to the commissioning organisation and the finalisation of the report can be difficult. "*Being told where you have gone 'wrong' and fighting to keep your interpretations is very stressful.*" (p. 54)

Telford (2001) concludes in rather stronger terms *"Donor inputs, performance and agendas need to be explicitly addressed in evaluations. In particular, political and geo-strategic 'agendas' should be evaluated against humanitarian principles, just as much as operational performance."* (p. 189)

Lindahl (2001) provides a useful if rather idealistic set of solutions

> *"In the end, evaluations must be a shared responsibility and a matter of trust between the parties. The commissioning agency must trust the evaluators, and listen to potentially critical views without partiality and with a preparedness to learn and change. The evaluators must appreciate the conditions in which humanitarian assistance takes place, and be constructive rather than scoring criticism for its own sake. Both must avoid the easy way out: of loading and receiving praise as a reason to do business as usual."* (p. 73)

In their overview conclusions Wood, Apthorpe and Borton (2001a, p. 207) suggest that attention needs to be paid to *"getting the report and its findings accepted and used."* They recognise dangers both in too much criticism which will *"tend to close ears"* and too much praise which *"encourages organisations to feel there is no need to respond, and little learning will occur".* They emphasise the need to *"stimulate action rather than encourage people to focus on escaping blame"* and the *"benefits from giving attention to the presentation of findings in a 'saleable' manner"*

The researcher observes that the relationship between the commissioner and the evaluator is one of unequal power in which the commissioner has the power to pay or withhold payment for the work done by the evaluator. As such it is a relationship in which the evaluator must be both willing, and empowered by the commissioner, to "speak truth to power".

The terms of reference are a tool by which the commissioner of an evaluation defines for the evaluators the scope of the evaluation to be undertaken and are therefore an essential element in the relationship between the commissioner and the evaluator. Wood, Apthorpe and Borton (2001a) complain that *"This is very much a consultancy-client relationship where the team comes in to do the job that is already defined and is not meant to ask any questions about it."* (p. 202) The accounts of the nine evaluations demonstrate that the imposition of badly written terms of reference adds levels of subjectivity to the evaluation process.

Subjectivity introduced by the terms of reference

Lindahl (2001, pp. 63-64) points out several weaknesses in the Terms of Reference for the Cambodian evaluation "*which are not uncommon for complex evaluations*". He refers to being required to assess "*the impact on different ethnic, gender and age groups; the impact on special target groups such as refugees, internally displaced and demobilised soldiers ; and the impact on the economy and various sectors*" as well as "*an assessment of efficiency, effectiveness, sustainability and so on.*" He points out that none of these were unreasonable if sufficient time had been allowed. He observes that "*Such a discrepancy between what is required by the TOR and what can be achieved given the resources available (time, systems and others) is not unique to the Cambodian case.*" As a result he reports that "*We* (the members of the evaluation team) *made our own individual interpretation of the TOR and carried various conflicting views with us through the evaluation…. Not until the report writing did all this surface as an open conflict*" He identifies the causes being "*Limited time, a leadership style that emphasised getting on with the work rather than questioning the TOR, and the strong personalities of the team members*"

Telford (2001) generalises, "*The designers of TOR are often overly optimistic about what can be covered.*" Driven by perceptions of urgency, small budgets, conflicting interests and the "*one chance*" nature of evaluations, the designers of the Terms of Reference (TOR) for evaluations try to include more than can be fully undertaken. The result is an evaluation team constantly under pressure to satisfy the commissioning agency even at the expense of a thorough evaluation procedure. (p. 188)

Lindahl (2001) discusses how over ambitious TOR may be renegotiated and recognises that this will differ depending on whether the evaluation contract is put out to competitive bidding or not. If the contract is awarded without competitive bidding, and "*if the Team Leader has sufficient standing and experience, he or she should in theory be able to negotiate the Terms of Reference.*" However, "*if the evaluation is subject to competitive bidding the evaluation team is in an awkward position. If it argues in its proposal that the TOR are unrealistic, there is a clear risk it will not be the winner of the bid.*" (pp. 68-69) Lindahl concludes that

> "*In my own experience… there is a tacit understanding between the aid organisation and the consultants that we understand that you have to ask*

more questions in the Terms of Reference than realistically can be answered, and we will try our best, but we assume that you won't hold us responsible for not being able to answer them all. If this is to change, there is a need for courage by both commissioning agency and evaluation team" (p. 68)

Borton (2001) commenting on the TORs for the Rwanda evaluation observed that

"The TOR for all of the teams were very lengthy: having been obliged to take account of the comments made by the 38 organisations represented on the Steering Committee' par excellence these were TOR that had been 'prepared by committee'… and read much more like a shopping list of points to be covered." (p. 81)

In their conclusions, Wood, Apthorpe and Borton (2001a) highlight a number of issues concerning terms of reference. They recognise that the case studies show that *"humanitarian evaluation teams are often given excessive and unrealistic TOR"* (p. 201). They suggest two reasons why this is the case:

1. The application of *"normal evaluation questions in particularly new and difficult"* situations *"where much of the information needed is not available"* (p. 201)
2. When multiple agencies all contribute to the TOR or when *"the TOR are circulated to different sections within the commissioning organisation for comment and addition"* (p. 201)

As a result of this "TOR overload" either the evaluation team will *"ignore parts of the TOR or will try to cover everything by spreading their efforts more thinly and drawing conclusions on the basis of limited data, thereby suggesting false accuracy."* (p. 201) They identify a key problem as being the direct application of a *"consultancy-client relationship"* defined for development evaluation being insufficiently adapted to *"the much more complex humanitarian evaluation process"* and recommend that *"A more negotiated process* (of TORs) *should characterise such* (humanitarian) *evaluations".* (p. 201)

6. Relationship between the commissioner of the evaluation and the project being evaluated

The perception of the informants in the project being evaluated regarding the commissioner of the evaluation and the perceptions of the motives for the evaluation are only specifically mentioned once in the accounts of the nine evaluations. Wood (2001) writes:

"To some extent the fieldwork was facilitated by the perception of Sida by its partners. For the most part Sida was perceived as a benevolent and understanding organisation, not one that was threatening, critical or over concerned with detailed accountability-style evaluations." (p. 51)

However, these perceptions regarding the motive for the evaluation may be common. A discussion of "accountability" and "lessons learned" based evaluations is undertaken later in this research. As indicated above, individual commissioning organisations tend to be perceived as more interested in "accountability" or "lesson learning", and evaluations are reacted to differently by those being evaluated.

7. Relationship between the evaluating team and the project being evaluated

O'Keefe et al. (2001) emphasise the crucial nature of the relationship between the evaluating team and the staff of the project being evaluated:

"The evaluation team must rely on informants, largely those responsible for the implementation, to tell their stories of the action. These stories must be tested against the beneficiaries' viewpoint until the evaluation team is confident it can tell a story from beginning to end. The story, however, has to have a purpose, a moral. Evaluation of humanitarian aid is the writing of disaster fables." (p. 30)

Apthorpe (2001) reports that the relationship with the project being evaluated can begin in an almost hostile manner, *"Our sense of foreboding was multiplied when, arriving at Freetown airport on the date agreed, the country representative politely informed us that, because things were so busy, it would have been better if we had not come!"* (p. 108) The relationship apparently remained at least antagonistic throughout,

"Our message in our debriefing by the country director was that we were unable to conclude anything in particular with any certainty. Understandably, this was not well received. He had anticipated an endorsement of past and present operations, perhaps with some added lesson-learning pointers for the future. In his view our position could be misread as being over-critical of a programme that was running 'as-well-as-could-be-expected-given-the-difficult-circumstances'

We left noting the unhappy divergence that had emerged with the country director, and concluding that interests and policy were determining the numbers given to us.

Our methodology at this point could be described as being based less on scientific than of philosophical principles as to what was likely to be the case, and what we could say about it. Policy analysis orientations were already well on the way to crowding out anything more project linked." (p. 109)

Other parts of the same evaluation, however, were conducted in a more positive and productive environment:

"I suspect that where evaluation reports do produce such evidential findings, rather than offering definitive statements of what is, they are often just impressions of what probably is not the case. Exceptions may occur where external evaluators are greatly helped by insiders (as we were, particularly in Guinea), or where they are able to draw on their direct experience of the case, acquired in roles other than those of consultants and validated through a peer review of some sort. The process, conventionally seen as one that leads from evidential findings to conclusions, is often in actual practice more like the reverse, with evaluators in effect looking for evidential findings to confirm and illustrate their best expectations." (pp. 114-115)

Wood (2001) reports a very positive relationship but also observes the historical nature of the projects being evaluated and the financial motives of the informants: *"The general experience of the team with the persons interviewed was extremely positive. There was little obvious evasion or discomfort, although in some cases there were tensions about who should be involved in interviews.* (p. 50) He attributes this positive relationship to three reasons:

1. The projects had already been completed
2. Those who had worked on the project "*wished to tell the story of a task that had, on the whole, been successfully completed*". (p. 50)
3. That doing so "*might improve their prospects of future funding*" (p. 51)

Grünewald, Pirotte and de Geoffroy (2001) outline the process that was to be central in their 'experimental and inclusive approach' to the evaluation of response to the post hurricane Mitch Emergency undertaken in 1999,

> "*The objective was 'evaluation, capitalization and collective learning'. The initial reactions were generally positive, but often cautious, both features reflecting the typical reluctance of NGOs, especially the smallest, which had not been following the current trends in evaluation thinking and emphasis from accountability to lesson-learning. It took us some time and effort to persuade them that behind the 'challenges and risks involved in evaluation work' there is dramatic potential for improvement.*" (p. 153)

They point out the value of taking time to build relationships with the agencies being evaluated and other stakeholders, "*Networking proved essential for establishing relationships with the national actors from the NGO sector. In turn this affected the quality of the fieldwork.*" (p. 159)

They conclude,

"*Creating empathy between the evaluees and the evaluation team is a sine qua non for a successful evaluation. It is critical to understand the field situation through the eyes of the actors, to gain their trust and support and to share information with them in a transparent manner. This will facilitate the mission and help ensure that critical remarks and recommendations will be well received and seen from the beginning as constructive.*" (p. 169)

Borton (2001) describes at length the pressure put upon him by one of the organisations evaluated, regarding the "*strong criticism of UNHCR for its lack of preparedness for the refugee influx into Goma*" (p. 95) which included interventions by the High Commissioner with ministers of two governments funding the evaluation. He concludes wondering "*Had I been manipulated into letting off the hook an agency whose poor performance I believed had contributed to the loss of thousands of lives?*" (p. 97)

The researcher observes that this is an interesting case where the organisation being evaluated has a strong enough relationship with the commissioner of the evaluation to be able to bring pressure to bear on the evaluator. In this case the fact that the commissioners had heavily funded the operation being evaluated and hence heavy criticism of the operation may have reflected badly on their own judgement as donors could have been the decisive factor in the decision to put a high level of pressure on the evaluator.

8. **Relationship of the evaluating team with beneficiaries and other stakeholders in the project being evaluated**

Relationships in the field which are important to the quality of an evaluation extend beyond those with the project being evaluated to encompass all the stakeholders. Grünewald, Pirotte and de Geoffroy (2001) report on the experimental methodology they adopted,

> *"The result was a rather inclusive approach, which was very different from the normal top-down exclusive strategies of most evaluation teams previously seen in the region. This ensured prompt and sustained support from all stakeholders and greatly facilitated smooth and cost–effective logistics for the mission. We undertook regular feedback to specific stakeholders to insure their ownership of the whole exercise. This was especially important as there was a fear within the evaluation team that the wrong message could have been sent by the donor. In particular it was feared that a defensive rather than open dialogue could result due to fears that a classical accountability-type evaluation was planned, rather than one emphasising the cross-cutting issue approach*
>
> *This was found to be quite effective for achieving the minimum level of confidence among the 'evaluees' and bringing them into the evaluation process."* (p. 161)

The relationship with government stakeholders was also important as was the political stance of the government concerned

> *"Meetings also took place with government officials in both Nicaragua and Honduras to understand the situation from their perspective and to bring them into the discussion of learning lessons for better practice in the future.*

In Nicaragua, the state position was developing a discourse so politicised and against the civil society that little progress could be made in the discussion. In Honduras, the situation was far better and interesting points could be exchanged on public health and disaster prevention." (p. 162)

The sustained inclusive nature of the relationships with stakeholders appears to have had on-going benefits,

"This comprehensive multi-phase feedback process proved to be a powerful learning tool and an important catalyst for the overall ownership of the results of the evaluation by both the NGO community and the relevant state actors. As a result a group called a 'Quality Platform' was established along with an 18-month process for developing methodologies for ex-ante assessments. (diagnosis, situation analysis, needs assessment) and ex-post evaluation (impact assessment, end-of-project evaluation)." (p. 168)

O'Keefe et al. (2001), however, sound notes of caution with regard to interviewing beneficiaries,

"Discussions with the beneficiaries of ICRC's work had less to do with the programme than with their own personal situations and survival mechanisms. While Somalis are articulate, as a result of their oral tradition, perceptions are subjective and prone to distortion." (p. 29)

Wood (2001) reflects a similar caution but applies it more *widely "Local contacts proved to be of considerable importance and facilitated, but perhaps biased, the field visits and case studies we made."* (p. 49)

Borton (2001) reported different problems: *"the tendency for interviews with beneficiaries in the refugee camps and in Rwanda to draw a wider audience than the selected interviewee, with associated problems of confidentially and accuracy".* (p. 88)

9. Relationships within the evaluation team

Evaluation teams are often assembled by the organisation commissioning the evaluation, may have not met or worked together before and/or have conflicting

interests. Wood (2001) observes that, *"While the ability to work together was perhaps valued, it was not a deciding factor for recruitment."* (p. 41) However, he reports that the team worked well together:

> *"There were no conflicts or tensions in the initial meeting in The Hague and there was no development of camps or factions within the team. This was probably helped by the fact that three of us knew each other and I had worked with two of the team on previous occasions."* (p. 48)

Later, however, Wood reflects that the good relations may be attributable to less positive reasons. *"There were no major conflicts within the team during the fieldwork. The reasons for this may be as much to do with the limited time the team spent together as in the personalities of the individual members!"* (p. 52)

Wood (2001) observes*: "The team leader was asked to accept the post having had no say whatsoever in the team's composition and not knowing any of the other three, except through the publications of one."* (p. 42) The working relationships within the team may have to be built while the evaluation is taking place and may vary greatly from evaluation to evaluation. Following a difficult start Apthorpe (2001, p. 106) reports that, *"Thereafter the team atmosphere started, jerkily, to improve... Around our halfway point in Liberia, elements of team solidarity were clearly setting in."* He identifies some factors in the development of the team relationship, *"Good fortune on a mission helps, but also the pressure of outside circumstances, and the evolving creation of mutual respect and 'team space' within."*

Grünewald, Pirotte and de Geoffroy (2001) appear to have found relationships within a joint evaluation team, that included the major funder of the project and evaluation, to have been more difficult than those with the stakeholders on the field:

> *"Another important player in the development of this evaluation was the Fondation de France (FDF), a private foundation that became the main donor… However, the involvement of the main funding agency in the evaluation had a number of side-effects that had to be dealt with during the mission.* (p.155)

> The Mitch Evaluation Task Force was established in order to allow the evaluation to consider an inter-NGO-government framework. *"It brought together about 30 NGOs, five or six administrative areas (collectives) in France, government officers from the Ministry of Defence, Ministry of*

Finance and Ministry of Foreign Affairs and a representative of the Inter-
American Bank. (p. 155)

Grünewald, Pirotte and de Geoffroy (2001) report that this large task force
"*had several positive and negative repercussions.*" While increasing the
"*range of people and institutions involved… which increased both the scope*
of the debates and the chance for coordination" it also risked *"political*
manipulation" by the government components. They comment on the
advantages and disadvantages of a broad based joint evaluation,

"This NGO-government collaboration created an interesting precedent in
France and gave our evaluation process unexpected weight and authority.
However, switching from an inter-NGO evaluation marked by notable
discretion to a more system-wide mechanism was a real challenge and
changed the rules of the game. Both Groupe URD and those of us on the
evaluation team had to adjust without 'losing our souls'. (p. 156)

They also report an interesting variant in the art of 'pitching' the report,

"In addition, after the main team had been selected, it was decided that it
would be helpful to add a photographer to the team to ensure more creative
post-mission feedback. This was decided after having witnessed an
emotional presentation of an evaluation mission based on a slide show." (p.
158)

Arranging an emotional presentation is surely a subjective element.

Telford (2001) comments on the stress in the team as a result of the different
objectives perceived by the various members of a joint evaluation,

"Differences of approach or methodology did arise in the team and this
generated a degree of stress. The differences stemmed from the fact that
the objectives and expectations were not entirely identical for the two
organisations and their members of the team: a hard-hitting' accountability'-
focussed evaluation was expected by some staff of the donor agency, while
a less contentious 'lessons learned' exercise was sought by the
implementing agency." (p. 178)

The conflict between this pair of objectives is discussed separately as it affects
several of the working relationships involved in an evaluation.

Lindahl (2001) identifies the resolution of one conflict being made

> *"We had two rather different concluding chapters, one written by the team leader, with whom I disagreed, and one written by me, with which the team leader disagreed. Fortunately we respected one another enough and he had the courtesy to put my conclusions as an annex."* (p. 67)

The working relationships indicated above are not independent of each other and may result in conflicting interests. The evaluation team needs to manage all these relationships simultaneously if necessary information is to be acquired and accepted by a range of interested parties. This can be a time consuming process. Grünewald, Pirotte and de Geoffroy (2001) observe that,

> *"Maintaining multiple accountabilities to both a donor and a network of field actors is essential but not without difficulties. Our experience suggests that transparency and clarification are crucial. In this case, each time there was either a perception of a hidden agenda or a lack of clarity; it required a strong effort to clarify the situation."* (p. 170)

Internal/external participants

Apthorpe (2001) reported the tensions of team members employed by the organisation being evaluated,

> *"Our UN team members were emphatic that, whatever this was, it was something for head office only. They feared that if the mission transgressed into such a protected area they would be placed squarely in the firing-line on their return, after 'the independents' had left."* (p. 110)

Similarly, Telford (2001) observes of a member of the evaluation team who was employed by the agency whose project was being evaluated *"However she had to cope with the position of being 'judge and judged' as a staffer of the implementing agency being evaluated."* (p. 181)

(Wiles 2001) comments on the effects of having staff from the organisation being evaluated seconded to the evaluation *team "The participation of the senior staff of the Tajikistan Red Crescent Society (TRCS) was important, not only for grounding the review but also for internal TRCS reasons"* (p. 126), however, he concludes *"The inclusion of staff and national society persons in an evaluation*

needs to be carefully thought through, and they should have carefully defined roles" (p. 131) .

In their conclusions Wood, Apthorpe and Borton (2001a) recognise the *"origins of team members are… a debatable matter"* and that whereas *"in some cases a mixed team of insiders and outsiders is effective… in other cases an external team is needed."* (p. 197)

10. Influence of Interagency relationships in the field

Conflictual relationships between organisations working in the project area influence the outcome of evaluations, with different perceptions of reality being presented to the evaluation team. This is particularly a problem when both organisations are stakeholders in the project and one is the organisation commissioning the evaluation. Apthorpe (2001) reports that,

> *"Aspects of the WFP-UNHCR relationship in the region were freely described to us by members of both agencies as 'mini-complex emergencies within the larger complex emergency'. Stand offs between them and the resulting further tensions were manifest, and, at times, experienced directly by us."* (p. 111)

As Munslow and Brown (1999) have argued, a major part of the complexity in complex humanitarian emergencies is introduced by the institutional complexities and complexes of the humanitarian organisations themselves.

11. Subjective influences relating to learning from evaluations and the learning cultures of humanitarian organisations

Evaluators often express frustration at how little use is made of their findings and recommendations. Wiles (2001) is an example of this,

> *"One of our recommendations was that there should be a structured follow-up process to review the implementation of its recommendations and that a senior manager should be designated to manage the process. .....this did not happen. Some, perhaps 50 per cent, of the report's recommendations were subsequently implemented."* (p. 131)

Apthorpe (2001) is rather more analytical in respect of the way in which an unwelcome finding was ignored by the commissioning organisation and reflects on the subjective nature of organisational learning.

> *"Why should WFP want to accept this bit of informed social learning when it believed it faced no crisis, had nothing to be particularly defensive about or to look hard to find? A senior manager's reputation, an organisation's official discourse and standing could be needlessly damaged if care were not taken to keep the independent flies off the elephants back, with their fidgety and, to the elephant, insignificant social facts. Why this particular piece of social learning ought to have been received, deserved to make a difference, and yet didn't is another matter. Organisational learning involves much more than just hired hands. It is an internal social process though in rare cases it may be triggered by an external intervention."* (p.113)

Wood, Apthorpe and Borton (2001a) present a rather dogmatic conclusion regarding the learning culture of humanitarian organisations,

> *"Finally, serious efforts are also needed to ensure that humanitarian evaluations are better used. Too many studies are not utilised effectively, and some are barely read beyond the draft stage. Follow-up tends to be patchy and informal, and there is often a lack of clarity about the role of the evaluation team and the evaluation department in ensuring action on the findings of such studies. While accountability and lesson learning outputs from an evaluation will need different forms of follow up, in both cases this should not be left to chance. Key lessons need to be identified and methods for ensuring their adoption established through training courses, staff development manuals and monitoring procedures along with agreed action routes.*
>
> *Overall we would argue that a win-win scenario can be established where lessons about evaluation are learned so that better designed and conceptualised humanitarian evaluations with adequate resources and sound methodologies, can be undertaken in situations where learning is sought and the structures to facilitate this are in place. However this must be a continual process as the range of humanitarian evaluation work is rapidly expanding and the gap between practice and guidance is growing."* (p.210)

Reflecting on his experience in West Africa, Apthorpe (2001, p. 121) writes, *"Indeed whether evaluations – rather than suitably composed seminars or workshops – offer the best route for policy lesson-learning, as distinct from policy lessons-to-be-learned, is another matter to be explored."*

12. <u>Subjective influence of the inclusion of multiple objectives of "accountability" and "lessons learned" into one evaluation</u>

As Telford (2001) observed, *"The designers of TOR are often overly optimistic about what can be covered."* (p. 188) One aspect of this is that many commissioners of evaluations require that the same evaluation both examine the accountability of the implementation of the project and the lessons learned from its implementation without realising the conflict of perceptions that this often sets up in the organisation being evaluated. In the environment of humanitarian action 'lessons learned' is somewhat a euphemism for 'what went wrong and how can similar projects be done better in the future'. Accountability, on the other hand is about efficient and effective use of resources, and openness about 'lessons learned' is often perceived as an admission that resources were not used in the optimal manner. In the highly competitive world of humanitarian action it is feared that such an admission may jeopardise future funding and the position of the organisation.

Lindahl (2001) points out what he refers to as a *"mixture of agendas"* in the TOR and identifies these as *"lessons for its emergency assistance"* and *"an aid audit focussing on impact assessment"*. He concludes that all the objectives in the TOR can be justified but *"combining them in one (short) evaluation has serious drawbacks as they tend to compromise one another."* (p. 64)

Telford (2001) observes of the joint evaluation of UNICEF's Kosovo operation:

> *"While it had originally been a DFID initiative, the evaluation soon became a joint effort. However, joint evaluations are especially susceptible to multiple expectations, and to being 'all things to all people'."* (p. 175)

> *"In this case the exact nature of such a 'joint' endeavour was never completely clear and the precise roles and participation of the respective organisations were never spelt out to all parties. While not sufficient to derail the evaluation, this degree of vagueness caused problems with some UNICEF staff and managers when it came to presenting the conclusions.*

Some UNICEF staff hinted that they saw the evaluation as a donor-motivated if not imposed, initiative." (p. 175)

"The significant and on-going DFID funding of UNICEF capacity-building in emergency preparedness and response, definitely tilted the power relationship during the evaluation in DFID's direction and raised the question as to whether it was a DFID inspection of UNICEF's emergency capacity." (p. 175)

Telford reflects that these differences generated stress within the team as discussed above and greatly complicated the writing of the report and drawing of conclusions, *"Finally, the inherent tensions of being 'all things to all people' (multiple expectations on the evaluation by the various potential readers) suddenly became glaringly obvious. Panic nearly set in."* (p. 184)

Reflecting in a more general way on attitudes towards evaluation, Wood, Apthorpe and Borton (2001a) write,

"Another important step in improving the approach to humanitarian evaluation is to change the attitude towards evaluations in donors and humanitarian organisations. Instead of evaluations being feared and seen as sticks to be used to beat an unsatisfactory organisation or as hurdles to be jumped, evaluations must be looked upon in a positive light, as opportunities for learning. Gaining from evaluations rather than losing from them should be the aim. Evaluation departments should be seen as critical in determining the future, and the future success, of the organisations undertaking humanitarian assistance. As units generating ideas and approaches for the future of these organisations, we would suggest that evaluation activities be relocated and integrated more closely into the learning and operations structure of an organisation, rather than being out on an 'accountability limb' as so many of them appear to be." (p. 210)

Here we see the link between changes in the subjective "attitude" and "perceptions" and the possibility of "objective" improvements in organisational structure and methods of operation.

13. Influence of the culture of evaluators and the "Western" origins of evaluation

In their introduction to the case studies, Wood, Apthorpe and Borton (2001c) point out that

"The authors of these studies are drawn from several parts of the world, although by no means do they form a fully representative sample… none… is from Africa, Asia, South America or North America. Overall the contributors are predominately white Northern males of 'mature' years (40-70). Only two are women. The Anglo-Saxon cultural zone is the origin for almost half of the authors". (pp.16-17)

In describing the pressure put upon him to accept the post of Team Leader, Borton (2001) remembers how *"My weak French language ability, paralleled by my belief that the Team Leader had to speak fluent French to work in Rwanda and with the francophone agencies, was downplayed".* (p. 80) He also observes that *"the fact that none of the team members was African was the cause of subsequent criticism of the team and the source of some embarrassment to members of the team"* (p. 85). He attributes the reason for this as being mainly due to the *"highly polarized nature of allegiances in the Great Lakes region"* (p. 85) which would have made the inclusion of possible candidates counterproductive. He does not, however, suggest that the non-African cultural bias of the team affected its work.

14. Subjective influences introduced by the use of the OECD-DAC Criteria

It may appear counterintuitive that the criteria designed to bring objectivity and consistency to the process of evaluation, are themselves a source of subjectivity, however, it appears that in two aspects that is the case.

a. Attribution of Causality

The attribution of causality to observed impact is reported as a cause of difficulty in the evaluation of humanitarian action. Lindahl (2001) observes that

> *"A particular methodological dilemma in an evaluation is the issue of attribution… The problem is compounded by the fact that many agencies tend to operate in humanitarian assistance. Hence assessing impact by a*

specific agency is next to impossible" and *"leads to guesswork and personal judgement".* (p. 70)

He concludes

> *"I believe that there is a general tendency to attribute too much to external assistance, and to underestimate the endogenous factors operating in the study locality"* (p. 70) and that *"too much of an emphasis on self-attribution in an attempt to show 'its' results is counterproductive".* (p. 71)

Under pressure from the TORs to demonstrate the impact of a project operating in a complex and changing environment where the beneficiaries are themselves agents of change, project staff and evaluators are at risk of assigning attribution on the basis of *"guesswork and personal judgement"* (p. 70) rather than an evidence base.

b. Cost effectiveness and value for money

Cost effectiveness has been defined as a combination of the two OECD-DAC criteria of efficiency and effectiveness, with value for money having a similar definition (Department for International Development 2011, p. 4).

Perhaps one of the most obvious questions the commissioner of an evaluation would want answered is whether the project has been cost effective and value for money. Lindahl (2001) observes that:

> *"While the TOR for evaluations routinely ask for assessments of cost-effectiveness, the reports equally routinely gloss over this issue with statements such as 'cost-effectiveness could not be assessed'… We should not fool ourselves into believing that we can undertake accurate cost-effectiveness assessments of most humanitarian assessments.* (Quoted as in the text, the second assessments may be a misprint) *First, effectiveness is very difficult to assess because objectives are difficult to quantify. And there is the problem of attribution."* (p.72)

> As with the attribution of causality, pressure to answer the question of cost effectiveness or value for money when inadequate information is available risks *"fooling ourselves"* into accepting subjectivity.

The above analysis of the nine case studies of evaluations of humanitarian actions undertaken between 1993 and 1999 has provided a historical framework of 14

subjective influences. This research was undertaken some 14 years after the last evaluation in the baseline case studies was completed and hence it was necessary to establish whether similar subjective influences continued to be encountered in undertaking current evaluations. To this end the researcher interviewed humanitarian professionals currently involved in the evaluation of humanitarian action regarding the subjective influences that they currently perceive as influencing their work. The results of these interviews are the subject of the following chapter.

Chapter 7 - Analysis of data collected from interviews

Analysis of sample

The results presented in this chapter have been obtained from 22 elite interviews conducted with highly experienced humanitarian practitioners who together have 396 years of documentable experience in the humanitarian field. The 22 participants include nine independent consultants while the remaining 13 participants are all actively involved in the evaluation of humanitarian action either at a policy or managerial level and are working with eight different humanitarian organisations.

All the participants had been involved in some aspect of the evaluation of humanitarian actions. The sample group consisted of people currently employed as shown in table 7.1

Table 7.1 Profile of participants indicating current employment and the average years of experience in each type of organisation

Number of participants	Current Employment	Code used in text	Average length of experience within humanitarian sector Years
2	Government Donor Organisations	GD	17.5
9	Independent Consultants	IC	23.8
2	International Organisations	IO	11.3
4	Non-Governmental Organisations	NGO	9.6
5	United Nations	UN	17.2
22	**Total participants Total Years**		**396**

Source: Author

The sample shows an experience range which may be expected to be typical of those working in the humanitarian sector where many begin working with NGOs and International Organisations before moving on to take up better paid and more stable positions in government or the United Nations. Some then become successful

independent consultants based on their already extensive experience. As may be expected, the independent consultants had the highest average experience in the humanitarian sector. The shortest experience of any participant was 6 years (working in an NGO) and the longest 32 years (an independent consultant).

The nine participating independent consultants, who regularly undertake or have undertaken evaluation of humanitarian projects, all had extensive experience with various humanitarian organisations before becoming consultants. (See Table 7.2) In several cases participating independent consultants have previously worked for more than one type of humanitarian organisation

Participants are identified by the code shown in Table 7.1 indicating the type of organisation for which they currently are employed and are numbered within that category so quotations from individual participants can be collated. This coding was used to identify whether the current working environment of the participant influenced the responses they made during the interview. To ensure anonymity, all references to organisations by name have been eliminated from statements made by participants. Details of location at the time of interview are not given as in some cases this would identify the organisation for which the participant worked and may in some cases enable an educated guess at the identity of the participant themselves.

Table 7.2 Previous experiences of the independent consultants who participated in the research indicating the areas of humanitarian action in which they had extensive employment before becoming independent consultants

No.	Extensive previous employment
3	Academic
2	International Organisations
6	Non-Governmental Organisations
4	Government donor
5	United Nations

Source: Author

The sample, while relatively small, included participants from the major types of organisations working in the humanitarian sphere and took advantage of the breadth of experience of a number of highly experienced independent consultants.

The sample was strongly unbalanced in terms of ethnicity (only three participants (13%) did not originate from an OECD country) and gender (only five participants

(23%) were female). Unbalanced in this context, however, does not necessarily mean unrepresentative in an environment which has been dominated by males from traditional donor countries. This long term bias of gender and country of origin can be demonstrated from the information given in the short biographies included in Wood, Apthorpe and Borton (2001b, pp. xv-xviii), of the 14 authors (three editors and 11 contributors) of the baseline case studies, 12 were male, two were female they are described as being of *"mature years (40-70)… a range of nationalities across the Western donor countries is included – France, Sweden, Netherlands, Ireland and Australia as well as the UK."* (p. 17) Another indication of the bias of gender and origin was given in the interviews undertaken for this research. Participant UN3 indicated that in the early 90s most evaluators of humanitarian action could be profiled as white males who spoke English.

The interviews were undertaken between July and October 2012.

<u>Structure in which the results will be presented</u>

In Chapters 4 and 5 the researcher demonstrated that the key findings of evaluations undertaken on one emergency were often not applied in the responses to future emergencies. This demonstrated a distinct lack of learning by humanitarian actors from the evaluation material and the resulting repetition of recommendations in evaluations of similar emergencies separated by extended periods of time. The interviews with the 22 participants have been analysed against the 14 types of subjective influences, identified from the baseline case studies, to determine if these are still recognised as being important at the time this research was undertaken. The interviews were also analysed to identify any subjective influences, perceived by participants, which had not been identified from the experiences of evaluations undertaken between 1993 and 1999.

The results obtained from this research are presented in a structure which broadly follows the 14 types of subjective influences, identified in Chapter 6. This has been done to allow the closest possible comparison between the two sets of observations. Material obtained as a result of the interviews that did not fit directly into the structure used in Chapter 6 falls into a number of categories.

1. Material pertaining to the general topic of subjectivity in evaluation. These have been placed at the beginning of the structure.
2. Material which went into additional detail is more clearly presented as sub headings within the structure used for Chapter 6

3. The additional subjective influence that emerged during the interviews, not noted in the 1993 – 1999 data, was the emergence in the past decade of "real time evaluations" conducted early in the project implementation and designed to apply immediately lessons being learned to modify the course of the project as it progresses. This has been added as No. 15 at the end of the structure.

The results are presented as below:

1. **Subjective influences perceived as universal to evaluation**

   a. Subjective influences perceived as a part of all evaluations of humanitarian action

   b. Is evaluation of humanitarian action perceived as an art or a science

   c. Perceived influence of politics, money and vested interests in the evaluation of humanitarian action

2. **Perceived influence of the sense of urgency surrounding humanitarian action**

   a. Influence of the high turnover of staff

   b. Influence of the lack of baseline data and poor data collection

3. **Perceived influence of the attitude, personality and interests of those undertaking the evaluation**

   - Influence of the selection of consultants

4. **Inter-organisational and interpersonal relationships**

   - Political Influences

5. **Relationship between the organisation commissioning the evaluation and the evaluation team**

   - Subjectivity introduced by the terms of reference for an evaluation

6. **Relationship between the commissioner of the evaluation and the project being evaluated**

7. **Relationship between the evaluating team and the project being evaluated**

8. **Relationship between the evaluating team, beneficiaries and other stakeholders in the project being evaluated**

9. **Relationships within the evaluation team**

   - Internal/external participants

10. **Influence of interagency relationships in the field**

11. **Subjective influences relating to learning from evaluations and the learning cultures of humanitarian organisations**

   a. Dissemination

   b. Selectivity for public domain

   c. Implementation of recommendations

   d. Influence of evaluation being perceived as an administrative requirement or "Box ticking exercise"

   e. Perception of ownership of an evaluation by the organisation being evaluated

   f. Influence of the readability of the report and the means of presentation

12. **Subjective influence of the inclusion of multiple objectives of "accountability" and "lessons learned" into one evaluation**

13. **Influence of the culture of evaluators and the "Western" origins of evaluation**

14. **Subjective influences resulting from the use of the OECD-DAC Criteria used for the evaluation of humanitarian action**

   a. Causality

   b. Efficiency and value for money

15. **Real time evaluation**

Several participants, including some of those working outside the United Nations, made reference to the UN Evaluation Guidelines (United Nations Evaluation Group 2005), indicating that these were often included in the briefing documents prepared for evaluators. Where this document gives explicit guidance, it is quoted at the beginning of various sections of these results to give an indication of established standards.

1. Subjective influences perceived as universal to evaluation

Participants recorded in this node 15/22 – 68%

Each interview started with a very general question asking the participant what they perceived as the role of subjective influences in the evaluation of humanitarian action, leaving the participant to define what they perceived as subjective. In the

few cases where the participant asked the researcher for a definition, the participant was asked to define the term 'subjective' themselves, as understanding the perception of the participants was part of the research. This initial question resulted in a number of different but somewhat linked perceptions of the role of the subjective in evaluations.

NGO4 said, *"My old mentor used to tell me that evaluation is about asking questions as if you are not afraid of the answers."* This immediately raises a subjective factor that there may be some questions that should remain unasked, and hence unanswered, because of a perception in the organisation that they may produce what would be considered as "the wrong" answers.

IC7 also used the word 'fear' in relating how the whole environment in which aid is delivered has changed. In the past, the implementer was *"stuck off in the bush"* with access to resources but with little contact back to headquarters and *"just got on with the job and achieved what you wanted to do"*. The environment now has become one of *"supposed accountability"* with *"somebody watching it, somebody being able to say 'this is what is happening in the field'*. The result has been a rapid growth in the information demanded by headquarters, fuelled by a fear that there may *"be a mismatch between what is actually happening on the ground and the expectation of what could or should be happening on the ground".* Again here we have a fear of finding unexpected answers that may be perceived as "wrong".

IC9, who had a background in academia, focussed on the issue of the mismatch between headquarters policy and field action in terms of *"denial"*. When the policy does not fit the situation on the ground the field team may deny knowledge of the policy developed far away in headquarters. The participant indicated that it was hard to tell whether this was a result of miscommunication or denial. The same participant suggested that good managers are continually *"adjusting, adapting to all sorts of things and they are evaluating as they go along".* Locking evaluation to one special activity in the *"project cycle was a disaster… making a big break between evaluation and management was not conducive to learning".* The participant suggested that this separation is increasing the perceived distance between headquarters commissioned evaluation and the necessities of field operations. GD2 suggested that the current role of evaluation increases the perception of the distance between headquarters and field

> *"the big push on evaluations is to focus on the supply side of evaluation – how to raise our game, how to produce bigger, better, fancier evaluation*

reports and very little of the demand side of evaluation how to… give evaluation evidence that will translate all these findings into practical policy. I think that area is very much neglected."

From these answers the researcher identifies that multiple fragmentations within humanitarian organisations have separated functions into "silos", headquarters, "the field", managers and evaluators, each with their own perceptions which often work against rather than with each other.

GD1 took the approach that the subjective element came from a lack of definition of what was an evaluation and hence the resulting misuse of the word. *"I mean one thing is that some people call and exercise evaluation, when it is not an evaluation whatsoever. There is a lot of mysticism about it, like monitoring, project completion reports, quick and dirty reviews, they all appear under the label of evaluation which is an unprotected label."* The same participant went on to explain the roots of the misunderstanding *"Anyone can call themselves an evaluator, it is not a profession which is... very well registered, like auditors"* although the participant did also indicate that a positive role is played in training and setting standards for evaluators by the American and European evaluation societies to address this problem.

 GD2 also challenged the idea that objectivity and subjectivity should be considered as alternatives in humanitarian action, rather they should work together and the real question should be *"to what extent should humanitarian interventions be based on evidence and to what extent should they be based on other factors."* He justified his position with examples,

> *"anything new or innovative obviously wouldn't have the necessary evidence to support it". "Evidence itself may be highly contextualised, which needs to be acknowledged when using it in (another) intervention"* and *"evidence is part of a much broader mix of factors that need to go into an appropriate humanitarian intervention".*

The researcher observes that in the opinion of GD2 the complexities of humanitarian crises and the rapidly changing work environments necessitate that objective "evidence" needs to be mixed with other less objective, perhaps subjective, elements in order to be able to design and implement a humanitarian response. Hence, the subjective in the opinion of GD1 becomes a necessary and important part of the process rather than an undesirable interference.

a. <u>Subjective influences perceived as a part of all evaluations of humanitarian action</u>

Is subjectivity only a concern in particular circumstances or is it a general phenomenon recognised more universally by those who conduct evaluations of humanitarian actions? This universality of subjective influences was a recurrent theme in many of the interviews undertaken, although participants looked at the pervasive nature of subjectivity in several different ways. IC3 simply said, *"It's right through the entire process…. What we regard as evidence and sources of evidence are highly subjective… It runs right through the whole process."* IC2 said, *"Just about everywhere along the line"* and identified as the crucial element *"the questions you ask, or the questions you don't ask."* Here the researcher identifies a further indication that being subjectively selective about exactly what is looked at in an evaluation is recognised as an important factor in the process of evaluation. The concept that reasonable questions may be feared to give answers that the organisation does not want to hear undermines the objective methodology and outcome of some evaluations.

IC1 having recognised the need for *"a strong, robust case for where the evidence points firmly to a particular analysis or interpretation"* went on to recognise that *"there comes a point where one applies a subjective assessment and where one's personality can come into it."* She continued to quantify her position as 70 to 80% should be objective but for the last 20% subjectivity comes in. IO2 referred to the subjective as *"basically applying different criteria to look at things."* The researcher finds a close link between the views expressed by IC1 and IO2 and those expressed by GD2 above, indicating that a mixture of objective and subjective thinking is necessary in the planning of humanitarian action. Here a similar mixture is indicated as being necessary in the evaluation of humanitarian action. If subjectivity is indeed a necessary element in the evaluation of humanitarian action and presented by some as a tool to be used in the process of evaluation, this must lead to the consideration of how "scientific" these evaluations can be.

b.   <u>Evaluation art or science</u>

Lindahl (2001, p. 66) characterised the evaluation of humanitarian action as *"more of an art than a science"* and enlarged on this statement indicating that*, "There are clearly ways in which evaluations of complex processes, such as humanitarian assistance, can be made less impressionistic and more scientific. However my own personal experience is that the art tends to dominate."* This research looked at whether the evaluation of humanitarian action is still perceived in this way?

A similar description was an approach taken by some participants to explain why subjectivity comes into play so often.   UN1, actively engaged in evaluations said, *"There is always a subjective element in evaluations… I believe that evaluations are more of an art, but it is not an objective science".*   UN2 said, *"I acknowledge it up front that subjective influence are at play and I have always found that any attempt to turn evaluation into some kind of hard science to be quite misguided…. It is the systematic use of anecdotal evidence."* IC6 nuanced the distinction, terming evaluation as a skill *"If I have to choose between an art and a science, possibly it is more of an art but I would prefer to call it a skill, a mentality"* indicating it as being somewhere between an art and a science. This should, however, be contrasted by the response of IC9; asked whether evaluation was more of an art than a science, simply replied *"Yes, yes, yes, absolutely!"*

Responses such as these, from both practitioners and evaluators, indicate that the perception of evaluation of humanitarian action as at least partly an art and certainly less than purely objective research, continues to be widespread.

c.   <u>Perceived influence of politics, money and vested interests in evaluation of humanitarian action</u>

Politics, money and interests were all cited in various places in the accounts of evaluations undertaken before 2000 reported by Wood, Apthorpe and Borton (2001b) as well as by Apthorpe (2011). The same theme was also cited by participants in this research as the reasons for widespread subjective elements. It was well summed up by IO1 who said that *"Evaluation is very much a function of the political and economic context in which it operates."* Being more specific he said *"The thing that most influences the involvement of objectivity or subjectivity… is where the money is coming from and what influences where the money is coming*

from." NGO2 indicated that subjectivity could not be avoided as *"the actions of all humanitarian actors are politicised to one extent or another."* NGO4 describing undertaking an evaluation said,

> *"You are going to be   walking in a minefield because you have a lot of political interests and you also have a very opaque system for giving and receiving feedback on what's happening within and around an evaluation exercise."*

IC7 summed it up, *"Political sensitivity is where there is a sort of intellectual subjectivity."* IC8 graphically described humanitarianism as having some of the characteristics of a "religion" which

> *"drives people to have a huge stake in whatever the findings are; to prove their particular mind set, orientation, beliefs and values which may or may not be what the reality is, and is certainly not objective."*

The perception of the influence of political interests by practitioners and evaluators indicates this subjective influence is highly pervasive and requires more detailed consideration. The subject of the effects of vested interests on evaluation of humanitarian action is discussed more fully under the heading of "Political Influences".

The researcher observes that there is a very general perception of subjectivity in the evaluation of humanitarian action, not just as a philosophical truth but evident in the practical undertaking of evaluations.

2. Perceived influence of the sense of urgency surrounding humanitarian action

Participants recorded in this node 17/22 – 77%

In the experiences of evaluators in conducting evaluations of humanitarian emergencies in the period 1993 to 1999, the problems caused by the urgency or perceived urgency of the environment in which humanitarian action is undertaken and evaluated were identified as a source of subjective influences on the evaluation process. Beck (2006) identifies several reasons why the urgency of humanitarian action limits the objectivity of evaluation:

- *"Data and information may be more difficult to come by: for example, there is a high turnover of staff working in humanitarian action, which may make it difficult for evaluators to interview key informants.*

- *Humanitarian action by its nature is often planned quickly, and objective statements and indicators may be missing from planning documents.*

- *Humanitarian action takes place in disordered conditions, leading to rapid change in circumstances that makes context difficult to be sure of later, and also meaning that many assumptions about 'normal' social and physical conditions may no longer be justified."* (p. 15)

Participants in this research identified continuing influence of a pervading sense of urgency in both the implementation of humanitarian programmes and their evaluation. NGO2 characterised the first month or six weeks of an emergency as *"manic and crazy"* and highlighted the *"special skills"* necessary to cope. In more measured tones, a participant working with an international organisation said *"it is in the nature of an emergency to be there fast"*. An approach to programme design, management and evaluation in an environment which was summed up by UN4 as *"Yes, when people are under pressure and people are working 24/7, sleeping two or three hours during a 24 hour period, yes, things will go wrong"*; is certainly going to invoke subjective elements.

Not all participants, however, agreed with this description. IC3 thought this generalisation *"overdone"* and a *"caricature"* of humanitarian action. NGO4 suggested that urgency *"is not always an accurate description"* and *"can become a little bit of a fig leaf in some situations"*. She enlarged on this point saying that not all humanitarian programmes are *"highly stressed… chaotic and highly fluid"* a description she believed only characterised *"sudden onset, life threatening emergencies"* and not slow onset emergencies characterised by *"deteriorating environment and deteriorating indicators"*.

There was rather more agreement regarding the sense of urgency, or at least compressed time availability, surrounding the evaluation of humanitarian actions. UN3 referred to the evaluation of humanitarian actions as *"quite demand driven"* often undertaken at the request of the agency head or other directors who *"want the product sooner rather than later"*. IC1 said, *"Agencies underestimate the time it takes to set up an evaluation… they make a start on day zero saying we want an evaluation and have one in a month or two, it is just not on",* and characterised this

as a *"consistent problem".* This point was also made by IC3 who indicated that it was not confined to the evaluation of humanitarian actions but applied to evaluation in general. *"I would say evaluation in general – there is not enough time. Certainly in the case of humanitarian evaluation, but I would not say specifically to humanitarian evaluation".* He went on to specify areas in which this applies as being the evaluation of *"impact"* (defined by Development Assistance Committee (1999, p. 31) as the *"real difference the activity has made to the beneficiaries")*

IC3 also identified the *"time actually dedicated to doing the job"* as an important area that can subjectively influence an evaluation. NGO1 specified the role of the budget available for evaluations as a controlling factor in reducing the time an evaluation team could be deployed on a specific project, sometimes reducing the time available to well below that necessary for a quality evaluation. IC2, however, did not feel that the lack of time added a subjective level but did *"detract from the quality"* and that not having enough time *"to interview a sufficient number of respondents"* may produce a lower quality evaluation. The issue of subjectivity comes about when conclusions are drawn on the basis of insufficient information. As NGO2 put it

> *"the worst case scenario is when you very hurriedly commission an evaluator who you don't know anything about and they don't know anything about the organisation, and you know you have to have an incredibly strong terms of reference to enable that to work… and that is not the case"*,

These are circumstances in which evaluators challenged by lack of time and pressured to complete extensive TORs are likely to make subjective assumptions and conclusions.

In the experiences of evaluators undertaking some of the earliest evaluations of humanitarian action described by Wood, Apthorpe and Borton (2001b), one distinct effect of the sense of urgency surrounding humanitarian action was identified: the influence of the high turnover of staff. During this research two participants also identified the lack of baseline data or poor data collection.

a) Influence of the high turnover of staff
Telford (2001, p. 177) identified the rapid turnover of staff in humanitarian operations with the risk of the loss of email traffic and files as one of the

elements in the sense of urgency in evaluation. A number of participants in this research indicated that this continues to be a significant problem.

NGO3 said *"the wave of first responders… stay on for a couple of weeks perhaps three months max… often the first group already has left by the time the evaluation has started… you miss out on the group or you need to track down all these people".* IC2 referred to this as the *"micro aspect"* the *"extremely high turnover of aid workers"* meaning that each wave has to *"learn from scratch".* He did, however, empathise with the humanitarian workers concerned, *"the high turnover is understandable… people burn out… they do it for a number of years when they are young… most drop out after five or six years".* GD2 expressed a similar view characterising *"a lot of people who implement humanitarian interventions, particularly with NGOs but also on the UN side"* as *"first timers"* or with *"a couple of short term postings in a totally different context"*

IO2 identified the problem in terms of *the "people who would benefit most from that learning (the evaluation) are the ones that actually are leaving the project".* IC7 said *"They all agreed to it, they liked the feedback and then the guy left. So what; you know nothing will change!"*

IC7 identified a different serious problem that results from the high turnover of staff and the difficulty of recruiting suitable replacements. *"They recycle people through the system and you know these people, very often they are corrupt, but you also know they're useless or they've been disciplined by their employers previously and you see them again and again in different settings in different countries!"* Such a situation of distrust on the part of the evaluator and unreliability on the part of a major source clearly presents a problem in producing an objective evaluation of a programme.

b) <u>Influence of the lack of baseline data and poor data collection</u>

IC6 focussed on the topic of the lack of baseline data in an emergency operation as potentially bringing a subjective element into the evaluation of humanitarian action. *"One of the biggest problems I often encounter in my technical field is lack of baseline or… the wrong baseline, namely an assessment. An assessment is obviously not a baseline."* She observed that in evaluations *"often non-recipients are not interviewed enough. If you don't*

have a baseline then at least you can have that as a kind of control group". The same consultant continued to criticise the quality of data collected during the response to the emergency phase

> *"Malnourishment data is very often – sometimes very regularly - collected but it is shitty, no value at all. So hard data beautifully portrayed with no validity has no use… but you have to evaluate the hard data. Don't take the hard data as a source of information without doing quality control."*

NGO2 spoke more fully of working in the complex world of emergency response where there was only limited information with which to make an evaluation of the resulting actions.

> *"The fact is that we are certainly not capable of establishing rigorous baselines. I would argue that you can't really do that in many development contexts let alone in humanitarian contexts. I think it comes down to what we mean by baseline data. We work in a very unscientific fluid world… what is crucial is that we understand what information we have, what it can tell us and what it can't tell us. Not to worry about it if it is weak or if it is subjective or if it is biased or incomplete. But we need to understand that it is incomplete or biased."*

In NGO2's statement the researcher observes an interesting concept is proposed, that of being objective about subjectivity, recognising it when it occurs and working with the uncertainties it creates.

Reviewing the responses of participants interviewed for this research regarding the topic of the sense of urgency surrounding humanitarian action, the researcher observes that there is a general recognition that the sense of urgency does affect the evaluation of humanitarian action. There is a real urgency regarding the initiation of humanitarian action which may be overstated or used for political ends. The removal of short term staff deployed at the beginning of the emergency presents difficulties to the evaluators and poor data collection and lack of realistic baselines provide an environment where some speculate reconstruction may be the only resort. The researcher, however, makes a distinction between a rushed evaluation as staff are leaving and the emergency phase is coming to an end and a

well-planned "real time" evaluation undertaken in the emergency phase to provide immediate corrective action to the project.

3.   Perceived influence of the attitude, personality and interests of those undertaking the evaluation

Participants recorded in this node 19/22 – 86%

"*Evaluators must ensure the honesty and integrity of the entire evaluation process. Evaluators also have an overriding responsibility to ensure that evaluation activities are independent, impartial and accurate*" (United Nations Evaluation Group 2005, p. 8)

Several of the authors who related their experiences of undertaking evaluations in the 1990s referred to the effects that the subjectivity of the evaluator can have on the evaluation. Lindahl (2001, p. 61) refers to the *"temperaments, judgements and perceptions of the evaluators"* as *"strongly influencing the outcome of the process".* Apthorpe (2001, p. 115) identifies *"Matters of style, school of thought, university discipline, gender, cultural background"* as being important influences. Wiles (2001, p. 128) relates that some team members' approaches related more closely to *"their employer's agendas rather than specifically the TOR."* Wood (2001, p. 53) raises questions regarding the *"influence of preconceived ideas on the fieldwork"* and in their conclusions, Wood, Apthorpe and Borton (2001a, p. 197) observe that *"In all cases the vested interests of the evaluators need consideration".* This research looked at how far these concerns continue to be perceived and accommodated amongst humanitarian workers at present.

UN3 immediately identified an extensive list of subjective factors he considered important

> *"age, gender, nationality, language, previous work experience, familiarity with the country, familiarity with the people working in the programme, previous positive or negative experience with a type of project, political orientation, and academic discipline",*

all of which he believed may colour the evaluation of a humanitarian project. NGO2 indicated *that "It is a huge problem, we have got some very contemporary examples… of evaluation products being simply not useable because of an*

evaluator's interests and agendas". IC5 said, *"One of the main sources of lack of objectivity is when the evaluator comes with some preconceived ideas… of what the issues are and how they've been dealt with"*. IC9 said that he believed that 100% of the subjectivity came from the attitude and personality of the evaluator. IC1 identified one way in which subjectivity *"comes through"* to such an extent is on reading an evaluation *"you could tell who the evaluator was even if you had not read the name on the cover"*. Perhaps rather happily IC6 said, *"I have probably been very lucky and I haven't encountered these people"* and indicated that in her opinion it was the client not the evaluator who was *"the biggest headache!"* As an explanation she added that she had only worked in small teams or undertaken rather specific evaluations alone.

Three participants commented on this aspect in ethical terms. IO1 commented *"In some cases you have consultants, I think, who are more ethical and evaluate based on what they find. In others… consultants shape what they find according to the people who are commissioning them"*. IC3 said, *"It comes down to the usual things of courage, integrity those kinds of words and… how much you are willing to sacrifice for the sake of so called objectivity."* IC7 explained his personal policy on this topic,

> *"I think you have to be very honest where you know your bias is positive or negative. I won't take on bits of work where I do feel I've got baggage. There are some jobs I won't go near simply because I have been involved in a past process, I've been negative or positive about it, and I don't feel I can be unbiased in my approach to the work. So I steer clear of it".*

As a follow up question to participants who indicated that the individual evaluator was a major source of subjectivity the researcher asked whether "If the same project was evaluated by two different evaluators would they have the same or very similar results?" UN5 focussed on the expertise and background of the evaluator, *"What you find is what you are an expert in… so the same terms of reference will have different outcomes depending on who is chosen to do it"* IC9 was categorical, *"Oh no you won't"* but explained that this was not a problem specific to evaluation of humanitarian action "if you send two anthropologists separately into the same village they'd come up with different pictures of the village". Furthermore, he indicated that the anthropologists would be conducting research which is expected to be more rigorous than an evaluation consultancy.

Having demonstrated a continued perception amongst participants that the attitude, personality and interests of evaluators did indeed add a recognised subjective element to evaluation, the researcher questioned how evaluators were selected.

<u>Influence of the selection of consultants</u>

"Evaluations should be conducted by well-qualified evaluation teams…. Evaluators should be selected on the basis of competence, and by means of a transparent process." (United Nations Evaluation Group 2005, p. 14)

As the attitude, personality and interests of evaluators are broadly perceived to have a major influence in introducing elements of subjectivity into the evaluation of humanitarian action, the selection process of the evaluators may be expected to be rigorous. Several of the participants commented on the importance of the selection of the evaluator. UN1 stated, *"The selection of consultants who can carry out a mission is a delicate one, not only because of the independence side, but also because of the competence and accuracy side"*. NGO2 said, *"It is not just the product but the whole process is changed when you have a quality evaluator… we build up a collection of trusted people"*. He continued, *"The quality of evaluators is a constant issue… there isn't a huge pool of the right sorts of people even though vast amounts of money are going into consultancies on evaluation"*.

Despite the recognised importance, it appears that the reality is often less impressive. UN3 summed it up *as "a very random process. We don't actually have any"* (consistent process for selection of consultants). He went on to describe how some sections had committees who decided who got contracts. For bigger evaluations a steering committee would be formed, including the evaluation manager, other concerned staff from within the organisation and some outside stakeholders. One of the functions of the steering committee would be to review the list of potential consultants. UN3, however, concluded saying, *"Now in my experience, I would say that I would normally be able to convince a steering committee of who they should take on."* Even with a fairly representative steering committee one person basically decided who should conduct the evaluation.

IO1 spoke of the difficulties of recruiting through public web sites which they *"try to avoid as we then get swamped with CVs"*. He continued to describe a structured process of using web sites such as ALNAP (which needs an individual registration)

to *"ensure that we are actually getting a more diverse and rich applicant pool."* Applicants are then reviewed by *"a committee of more than three people"* who rate the applicants using a *"consultancy matrix"* so that the final decision *"is made collectively and not made by one person"* Even so he asked *"Is it a real transparent process? Or is someone cherry picked that they know is going to basically kiss ass and say what they feel the commissioning entity wants to hear?"* IC5 summed it up *"selection of evaluators is so haphazard, it is difficult to speak about really!"*

Several participants commented on the regularity with which the same consultants were used. NGO1 said that selection *"at the moment is based on recommendations or a database of consultants that we have".* She continued,

> *"I think there is a tendency to use consultants we have used before because of reputational issues. I do have some concerns about using the same consultants all the time because I think they do tend to come with a set way of understanding an evaluation… so I think it is helpful to sometimes bring in somebody new who can be slightly more objective".*

NGO2 said,

> *"Evaluators are chosen that have an empathy for what your organisation is trying to achieve both programmatically but also through evaluations that are trusted enough, trusted enough to be challenging. Now whether that makes them objective or not is a completely different point."*

The sense of urgency, examined above, extends into the recruitment of the evaluators. UN3 said, *"Normally they want the products sooner rather than later and of course the (here he named a well-known consultant) of this world are not sitting around waiting for us to call them".* IC15 speaking of the urgency for recruitment said*, "Probably good evaluators are in demand and are already booked up",* and IC7 said,

> *"They want you tomorrow, they assume you're sitting around waiting for them to call you and as if you can get on a plane immediately and do it. They are unrealistic about the amount of time involved in reading, travelling, meeting people, interviewing, and writing up. They*

are looking at the bottom line… and you know an extra day here and there is trying to be cut off. Very often they will go for cheaper people or cheaper products, then perhaps they get what they pay for."

The researcher observes that considerable misgivings regarding the selection of evaluators for humanitarian action exist across a wide spectrum of the humanitarian community.

4.   Inter-organisational and interpersonal relationships

The accounts of evaluations undertaken between 1993 and 1999 (Wood, Apthorpe & Borton 2001b) indicated that the relationships built, or damaged between various participants in the operation and evaluation of humanitarian action had an effect on the evaluation and learning experience process. IO1 said *"but the thing that most influences the involvement in subjectivity or objectivity is going to be where the money is coming from and what influences where the money is coming from."* IC1 said, *"I would say that one of the key factors that influence the evaluation function or process is the money and the power that is associated with money."* The importance of the complexity of the flow of aid resources has been referred to in (Wood, Apthorpe & Borton 2001b, p. 8) where a diagram attributed to Borton was reproduced. (See Figure 7.1)

From Figure 7.1 it can be seen that the flow of resources is complex. Each link in this chain represents not only a channel for the flow of resources but also a relationship of interest and accountability. These relationships are both organisational and personal. The head of fundraising in a large humanitarian organisation once, in mentoring the researcher in donor relations and stressing the importance of interpersonal relationships, said, *"Always remember that organisations do not give to organisations, people give to people".*

Figure 7.1 Principle resource flows and routes between donors and beneficiaries in the humanitarian system.
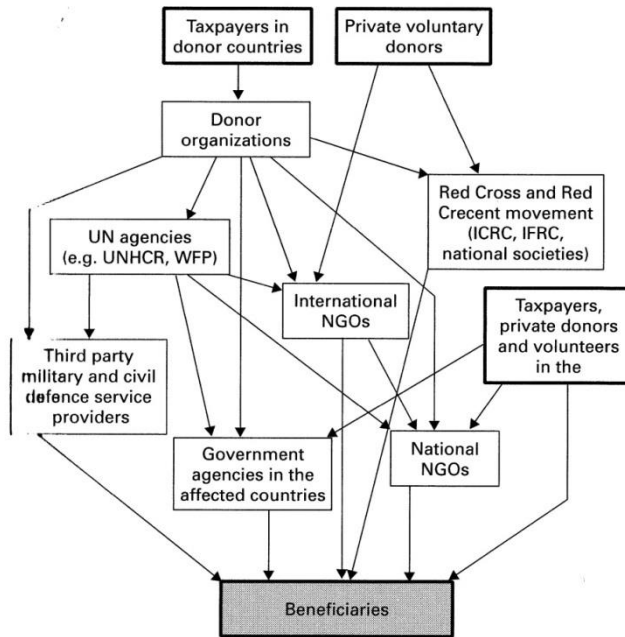
Source: Wood, Apthorpe and Borton (2001b, p. 5)

The chain of relationships involved in the evaluation of humanitarian action reflects some of the complexity of the flow of resources and may be represented as below in Figure 7.2.

Figure 7.2 Principal relationships involved in the evaluation of humanitarian action
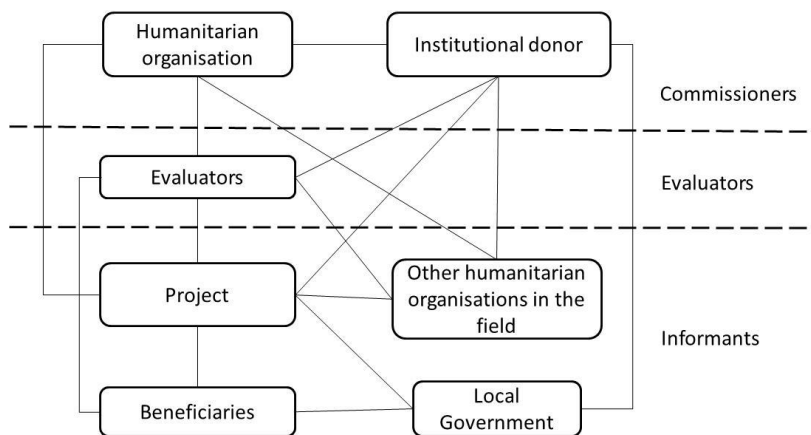
Source: Author

Each of the relationships indicated above will have a dual character, the institutional relationship and the personal relationship between the representatives of the two organisations who relate at an individual level.

Some aspects of the relationships are specific to the parties involved but others are of a much more general and pervasive nature. The latter type was described by participants in the research in general terms and will be reported here in the same manner. In some cases this interaction with participants opened doors to more detailed and specific discussions which are reported under separate headings. One example of this was the influence of politics and other vested interests.

Political Influences & vested interests

Referring to the internal politics surrounding evaluation, UN3 said,

> *"I would not go as far as to say that they (senior management) like it but they acknowledge the need for it because it makes it look as if we are taking a self-critical look at what we are doing and trying to improve ourselves. In a performance orientated culture that is important and they know they wouldn't get away with not doing it".*

> Here the respondent is indicating a perception that the evaluation of humanitarian action is seen by top policy makers as something inevitable, enforced on the organisation from outside rather than being their own tool for policy development. The same participant went on to say, *"evaluations do play quite an important symbolic purpose both for the organisation but also for the governments that fund us"*. This was said in the context of the symbiotic relationship between humanitarian organisations and the government departments that fund them both wanting to demonstrate transparency and a balanced approach to showing success, while at the same time learning from past actions. Here the evaluation is seen as a tool to generate these perceptions. Reinforcing this idea, IC1 simply said, *"It's something you have to do to keep the funding coming"*. Emphasising the all-pervasive nature of political influences and vested interests, IO1 listed the vested interests of the implementing agency, the entity requiring the evaluation, the donor and other stakeholders such as local politicians.

The selection of programmes for evaluation also has an important political influence, highlighted by IC2 when he referred to a *"politically commissioned evaluation"*, saying, *"The minister wanted an evaluation so the department did one"*. However he went on the clarify the mechanism that followed, indicating that the minister (in this particular government) would leave it to the evaluation department to take it from there and he would not be involved in the process at all. While this sounds like a good example of non-interference it leaves the question open regarding which other important programmes were not evaluated because the minister was not interested? The effectiveness of evaluations commissioned at the wish of politicians was commented on by a participant working with a government donor organisation: referring to situations in which critical evaluations reach parliamentary debate, *"the minister will say, 'yes, yes I will take care of it' and then parliament is satisfied and they never follow through after."* Explaining the reason for the approach of politicians who must have an evaluation but then fail to follow through he said,

> *"And so you have to educate those people by reasoning. But then again it's difficult to educate politicians in parliament because they like to deconstruct the world into very concrete and acceptable bits and pieces and the world is not such that it can be done that way."*

GD2 referred to the need of politicians to cater for the "Daily Mail" test, referring to a right of centre British tabloid newspaper which aims to reflect and influence the views of its typically conservative British middle class readers.

Not only are the programmes to be evaluated sometimes selected by politicians but politicians also have a direct effect on the type of aid being given. IC2 referred to an incident where portable hospitals were dispatched to a disaster area as

> *"a decision by the prime minister against the advice from the humanitarian department (of the government donor). They checked with OCHA who said, 'we've got plenty of those, don't waste your money on it'. That is such a waste of resources, but that's political, we are going to show that we are doing something".*

NGO4 referred to doing evaluations as sometimes *"walking in a minefield"* because of *"political interests and opaque systems"*, of the need to *"be able to understand the political sensitivity within the organisation"*, of *"untold or unexpressed content"* and the need to *"find a minimum common denominator that does not lower the quality of what you are going to do but is able to push the exercise forward"*. However IC9 warned that *"political sensitivity is where there is a sort of intellectual subjectivity"*.

Some participants gave a very negative view of some evaluations. NGO2 said *"It's a promotional activity quite often"* and UN5 speaking of one example said, *"you have a project, you initiate an evaluation with basically half the purpose of putting yourself in the limelight and pulling everybody else down… so you bring in these experts and you do this evaluation where you get flying colours… so it becomes a promotional tool rather than an objective evaluation"*.

In summarising the political dimension IC8 said, *"I've rarely come across a subject that is more opinionated and ideologically driven than humanitarian work."* Whereas IC6 focussed on their own solution, *"So there are interests, everybody has an interest, it's very dynamic. An evaluator has to learn to balance everything including his or her self."*

5. Relationship between the organisation commissioning the evaluation and the evaluation team

Participants recorded in this node 20/22 – 91%

"The relationship between the evaluator and the commissioner(s) of an evaluation must, from the outset, be characterised by mutual respect and trust". (United Nations Evaluation Group 2005, p. 13)

This relationship was identified by Apthorpe (2001, p. 117) as important as a subjective influence when he observed from the point of view of an independent consultant *"Having a client is much better than not having a client if you want to make a difference to policy"*. The inference of this comment being that preserving the relationship between the evaluator and the client is fundamental and recognising that trying to make too much of a difference to policy may completely disrupt the relationship and be counterproductive. The research looked at whether evaluators

still feel that the pressure to maintain the relationship with the commissioner of evaluations added subjectivity to the evaluation process.

UN3 spoke at length on this subject.

> *"And here we come to the old thorny question of are independent evaluators really independent, or are they looking for their next contract and probably inclined to be relatively sympathetic towards an organisation because if they are too critical they maybe… won't be used again?"*

He answered his own question saying*, "I have never really seen any empirical evidence that could prove it one way or another and I have tried asking independent consultants. One answered, 'I can't entirely deny the fact that this might have some kind of influence on the way I go about my work'."* UN4 indicated that a lot depended *on "the terms of reference that had been given"* and whether *"they were told this thing is broken and needs fixing… you may get the sense that it is because they would like to have the next follow up work they want to please the commissioner of the evaluation".*

Some independent consultants interviewed in this research were quite outspoken on this issue while others were rather vague. Certainly the most outspoken exchange was with IC9

Question: How much control have you experienced by the people commissioning (the evaluation)?

IC9: *Oh – 100 per cent*

Question: Is there such a thing as an independent evaluation?

IC9: *Oh no never! Independence is very relative and not an absolute term. I used to love the idea of thinking I was an independent person but of course one isn't.*

Question: Is the consultant looking to them (the commissioners) for future contracts

IC9: *Exactly yes.*

IC3 said, *"It comes down to contractual arrangements, he who pays the piper calls the tune"* and UN5 said *"there's a lot of consultants hanging around and if you are desperate enough you will be willing to do whatever you are told or asked to do".*

NGO4 recognised the problem but also recognised that it was not specific to humanitarian evaluation. She said,

"This is always going to be an issue and is not specific to humanitarian enterprise… I don't think that the fear of not getting the job next time is particularly more acute in humanitarian evaluation than it would be in any other type of consultants' job. I think this is a consultant specific issue not a humanitarian specific issue."

NGO2 was more nuanced explaining that consultants *were "paid by the organisation to produce a product and the commissioner has some degree of control over the content"*. However, she qualified this *"strong consultants… say this is what I have seen and observed, you may not like it but that's it. Strong consultants do that and agencies accept it… I suspect they push back less strongly if they are concerned about future earnings."* IC7 who has experience of taking a strong position explained *"I know for a fact that I'm often given jobs because they want somebody who is going to be very honest. He said the question is whether, when you are doing the evaluation, you are able to say, 'back off, I'm not having you in these meetings'. They will introduce us and we will say, 'thank you very much, you can leave us now, we will have these meetings without you.'*

The regular use of the same consultants was highlighted as a danger area by two participants. IC5 cited one organisation which *"tends to have a group of companies that it uses"* (for its evaluations) and that there were questions *regarding "to what extent the company wishes to not rock the boat even if the boat needs to be rocked… because they are keeping an eye on future contracts."* While the participant felt this may apply to companies, he went on to indicate that he had not seen *"evidence of… individuals toning down what they want to say because of future contracts".* GD1 related a situation where *"the press was getting onto the evaluation department saying… you always use the same consultants and they are not critical because they've got lots and lots of things to lose".* He continued to relate that an *"investigation went into that matter and there was no proof found whatsoever".*

IC1 used phrases such as *"I felt pretty independent. One of the skills of an evaluator is to be very conscious of the different audiences when you are writing… you could say tailoring it to the audience, but it is ensuring that you minimise the strong reactions that you could get".* He went on to relate a case where criticism of a large UN agency led to him coming *"under terrific pressure to change that assessment",*

including the head of the agency writing to the minister of the donor government which had funded the evaluation. Also on the topic of the developing relationship between the evaluator and the commissioner of the evaluation, IC4 spoke of discussions within a team of evaluators regarding whether the *"satisfaction of the client"* was the *"objective and the main indicator of success"*. They concluded it was *not "because a good evaluation will* (from) *time to time lead to dissatisfaction of the client… because they don't want to have bad results if they hoped for good results"* She related an occasion when this resulted: *"in the short term"* they were *"blacklisted"* for *"some years"* but *"after that we restarted good relations."* She concluded that *"maybe it's tricky time to time but at longer term it's the only way to remain credible."*

IC4 described the negotiation that takes place between the commissioner and the evaluators. *"In every evaluation there is a negotiation between… what the people, the payer and the evaluator will finally assume as the final product."* She set limits to what was negotiable *"we don't want, and this is our credibility, to really change the main and key important messages… maybe accept to change a little, secondary things but… make sure that on the main important issue you remain strong and firm".*

NGO3 explained the process in terms of obtaining the best outcome, *"it has something to do with… how critical can one be… in relation to what doesn't encourage an organisation… to learn".* IO1 identified three levels of pressure for change that may be applied by the commissioner. One level, disapproved of, is to change the *"evidence"* collected by the evaluators, a second level is to change the *"tone"* in which findings are presented so as to make them more acceptable and the third is in *"structure, just the clarity of how things are presented".* He concluded that *"it is important and imperative that the evaluation team is able to tell the story that they see needs to be told."* IC5 said his final fall-back position was *"one has to put in a note saying the agency does not agree with this and leave it at that."*

The definition of "independence" used by one organisation was quoted by IC9 as *"a team with no previous connections with the project concerned"* he pointed out it did not exclude *"previous connections with the agency concerned"* or any *"special knowledge of the project concerned".* He concluded that *"the word independent had done more harm than good because it deflects attention."* In his opinion independent has become another humanitarian buzz word which it is obligatory to use but which is open to a wide variety of interpretations. Such "buzz words", by

their wide range of non-standardised definitions, allow a high degree of subjectivity, where everybody can understand what they wish to be discussed, within a mask of objectivity.

<u>Subjectivity introduced by the terms of reference for an evaluation</u>

The terms of reference (TORs) for an evaluation are a document or part of a document prepared by the commissioners of the evaluation. *"The terms of reference should provide the purpose and describe the process and the product of the evaluation".* (United Nations Evaluation Group 2005, p. 10) This UN definition includes nine elements that should be included: context, purpose, scope, criteria, key questions, methodology, work plan, products and use of results, each of which is defined in detail in the document.

In his account of the evaluation of the UNICEF project in Kosovo, Telford (2001, p. 188) reported, *"The designers of TORs are often overly optimistic about what can be covered"* and emphasises the vagueness of the resulting TORs in the case of a this joint evaluation. As a result Telford (2001, p. 175) concludes that *"While not sufficient to derail the evaluation, this degree of vagueness caused some problems to UNICEF staff and managers when it came to presenting the conclusions."*

This research looked at whether there was now a greater degree of satisfaction with and confidence in the terms of reference provided to evaluators and whether the terms of reference provided to evaluators contributed to subjectivity in the evaluation.

NGO2 recognised a problem resulting from the need to evaluate many projects. *"I think quite often we try to do evaluations that cover far too many areas and so therefore it doesn't actually deliver for anybody… I think generally commissioners feel compelled to evaluate and they don't necessarily have clear ideas about what they want to know".* NGO4 attributed one of the main reasons that it has proved difficult to produce,

> *"higher quality evaluation outputs"* as being *"Ambitious terms of reference, or absolutely unfocussed terms of reference or any way poor quality terms of reference. You know Christmas list, wish list of terms of reference that ask you to look at… from cost efficiency to impact in 20 days in a huge programme."*

IC7 expressed his frustration when relating one experience with vague terms of reference,

> *"I mean we criticised them because they were all over… completely East to West and we said we can't with the time you are giving us... we cannot look at any of these issues in detail. We can skim them all but if you want detail in any of them you have to give us more time or less breadth and they said no, no, no!"*

He was able to identify the cause of the problem, which in a large organisation appears common,

> *"I found out later that those terms of reference had been… thrown around in house by three different business groups for months because they all wanted something different out of the evaluation… They could not work out a clear, single, unified terms of reference to please everybody so they bunged it all into the same document."*

GD2 presented a similar analysis,

> *"Terms of reference being far too ambitious, I completely agree. I've seen some ridiculous terms of reference. It is the nature of how, not just terms of reference, but any piece of work, is developed within an organisation. Whenever it goes round for consultation people only ever add stuff, they never take it away."*

He went on to describe *"all singing, all dancing, unbelievable terms of reference"* with *"hundreds of questions"* His advice: evaluations should have no more than three short questions and should be pared down. He blamed the *"inexperience"* of the person *"pulling together"* the terms of reference and *"trying to please everybody"*. The result he identified as *"evaluators themselves are very frustrated and just irritated by great long winded TOR they simply can't deliver on."* The same participant, GD2, considered the role of the evaluator in improving TORs. *"A good evaluator should be able to explain to someone (at the commissioning organisation) what they are proposing is not deliverable."* He identified the problem though as one of timing in a competitive bidding process. *"Do you try to explain that when the thing is going out to bid or when you have actually bid successfully? If you're challenging the terms of reference and another agency is saying 'Yeah we'll do it, you've got a problem haven't you'."* He concluded his answer listing a number of

evaluations and a wide range of organisations where he had witnessed the problem; evaluations *"trying to be all things to all men"*.

Looking for a positive way forward NGO4 who had been very critical of some terms of reference said she agreed with *"a number of donors"* that what was required was a focus in the design of terms of reference on how the evaluation was to be utilised and advocated, *"A utilisation focus approach to how you not only design your terms of reference, how you select the questions and how to interact with your field staff."* She indicated that the developing context of "real time" evaluations would make this essential.

6. Relationship between the commissioner of the evaluation and the project being evaluated

Participants recorded in this node 3/22 – 14%

In the published experiences of those who conducted evaluations of humanitarian action between 1993 and 1999, the relationship between the commissioner of the evaluation and the project being evaluated was only mentioned once in the context of the *"fieldwork being facilitated by the perception that Sida was. a benevolent and understanding organisation and not… threatening or over concerned with detailed accountability-style evaluations"*. (Wood 2001, p. 51). The researcher was interested to see how this aspect was perceived by participants in this research.

Two participants indicated that evaluations commissioned by donors were perceived differently. NGO1 said, *"I think people would get a sense of more urgency of more kind of stringent requirements when it's a donor asking for this information – because I think there is usually some kind of financial issue related to it".* NGO2 raised similar concerns, *"So for those evaluations that are being done for donors… organisations understandably have concerns over reputation and therefore longer term funding."* He contrasted this with an evaluation *"that had been commissioned for internal reasons and was more learning focussed".* NGO3 identified one donor as being one that *"looks more over your shoulder than another donor"* and that this would make more difficult the relationship between the evaluator and the project.

It is interesting that the participants who picked up on the effect of the relationship between the commissioner of the evaluation and the project being evaluated worked with NGOs and all related it directly to their relationship with institutional donors.

This may indicate that NGOs feel more vulnerable before institutional donors than the other types of organisations represented by participants to this research.

7. **Relationship between the evaluating team and the project being evaluated**

Participants recorded in this node 16/22 – 73%

The UN evaluation standards cite some basic rules for the relationship between evaluators and participants in the evaluation concerning *"conducting the evaluation and communicating its purpose and results in a way that clearly respects the stakeholders' dignity and self-worth"*. They also highlight the need to ensure the minimum inconvenience to participants and to respect *"people's right to provide information in confidence"* while informing them *"about the scope and limits of confidentiality."* (United Nations Evaluation Group 2005, p. 9)

Those who wrote their accounts of evaluations between 1993 and 1999, however, expressed much deeper concerns regarding the relationship between the evaluator and the project being evaluated. O'Keefe et al. (2001, p. 30) emphasised the reliance of the evaluators on their informants *"largely those responsible for the implementation to tell their stories of the action"*. Apthorpe (2001, p. 109) reported an evaluation where *"the relationship* (between the evaluator and the project manager) *remained at least antagonistic throughout"*. Grünewald, Pirotte and de Geoffroy (2001, p. 153) focussed on what had been at that time an *"experimental and inclusive approach"* to evaluation in order to improve the relationship and hence to get an overall better and more useful result from the evaluation process.

This research looked at how participants now view the subject of this important relationship between the evaluator and the project being evaluated.

The participants broadly took three approaches to discussing this topic. Some spoke of the continuing difficulties in building the relationship, some of how they work to build a better relationship, often focussing on an aspect of ownership, and some of how they worked despite a relationship that remained poor.

Of those who spoke of the continuing difficulties of building the relationship, several went on to describe how these were overcome. Some of the more experienced independent consultants who had been involved in evaluation of humanitarian action spoke of the changes they had seen taking place.

NGO1 related *"I think there is generally a level of suspicion or fear maybe"* and of one particular evaluation, *"I felt there was a little bit of fear about what we might reveal or some of the issues which might come up"*. UN5 likened the reaction of some projects to evaluation to a visit to the dentist, and feeling a conflict between knowing there is something wrong and hoping it would not be discovered because of the likely discomfort of having it attended to. NGO4 also used the analogy of a visit to the dentist.

> *"In the old days there used to be a lot of rigidness associated with evaluation because in the old days evaluation was perceived like… you walk into the dentist for a root canal… having your project dissected under a microscope… the teacher is coming to check if I've done my homework… I think we are now quite safe if we make a statement about the fact that we have moved away from that type of interaction".*

IC1 remembered,

> *"I have been in evaluations where material was withheld and that's as though the evaluator arrives and the shutters go down… The first ones I was involved in, you know, that discussion, 'Who the hell are you to come and question what we were doing? We were doing it to the best as we could. Who the hell are you?' I think… whereas now it's almost like it's an accepted, protective space. So that's a significant achievement over the last 20 to 25 years".*

UN5 listed a number of reasons why evaluators faced defensiveness from project staff. *"People have a pride in what they are doing and the organisation they are working with so they would be defensive about that"*. The risk to future career was an important element. *"If I'm doing a job and I am being evaluated then of course you are lots more careful because then you are nervous and how this is going to be reviewed and used and will affect you down the road."* Another element was the fear and vulnerability of the whistle blower.

> *"You have a system with very poor job security… you can be so confidential in the report… but it does not take very long to isolate and find out actually who said it… so people who are involved and can be affected by the evaluation would tend to be very careful.*

" GD2 remembered being himself *"on the other end of an evaluation"* and being given a *"full script that we were to try to stick to as much as we could"*. He speculated that this may happen to a lot of organisations *"particularly if it will affect their funding".*

Focussing on building a better relationship, UN3 said, *"We make a big play on the fact we try to work in a consultative manner, we are not there to judge or to assess individual performance. We are there to look at organisational effectiveness as a whole."* With a note of caution he added, *"However hard you try you can't get away from the fact that the word evaluation has some kind of connotation of judgement and assessment etc. We acknowledge that but we try to play it down as much as possible."* As part of his strategy to overcome this, particularly in an emergency operation he continued, *"we are a shoulder for you to cry on… so we try to give the sense that we are a channel for communication… and people are always positively inclined when they see what they have said to you is reproduced in the evaluation report".* In this goal of relationship building he recounted that when convinced that the evaluation was about *"learning lessons and… how we can improve our work for future planning": "there was a real change in their face and in their attitude, they became a little bit more relaxed".* He speculated that *"maybe that's where there is a sense of suspicion or the fear that the information is taken away from them and not given back to them".* To resolve this he *"builds into evaluations time with the team and partners to feedback initial findings so that there is a sense of inclusion and addressing issues early on before the final report is written."*

The theme of working with the informants on the results and recommendations of the evaluation early in the process was also referred to by other participants. IO1 recommends *"sharing of preliminary findings, conclusions and recommendations; that is presented in an oral form, ideally face to face with representatives of the key stakeholders… We are saying this is what we are finding, concluded and recommend, we want to share it with you and hear what you think."* He concludes that stakeholders should feel that the evaluation is being done with them rather than to them and as a result *"ultimately there will be more ownership of the evaluation and sustainability of its findings and recommendations".* IC4 was even more inclusive

> *"We try to start with a meeting with the key people from the programme to discuss the terms of reference and to explain how we are going to work… We always finalise our evaluation in the field mission with a debriefing*

session in which we present our preliminary findings and we discuss recommendations. And I would say this is really, really something very important… the chance for people to respond… to complete our understanding of the situation".

In formulating the recommendations she indicated *"It's much easier to think with the people that will be in charge of the implementation of recommendations, to start to think with them what is feasible and what would be the best."* IC7 referred to the *"proper debriefing session at the end of the evaluation involving discussions about what you are going to write, that's not even been written at that point".* GD1 went as far as to say that *"Most of the lessons learned that occur are during the evaluation process anyway and not because of a report sitting on a desk and people having to plough through the report".*

Some participants indicated that sometimes it proves difficult or impossible to build the sort of cooperative relationships discussed above. IC1 referred to situations where interpersonal relationships become dysfunctional, *"there is a personal element that begins to come in, you know, 'Oh sod him, why should I give him the benefit of the doubt?"* NGO2 referred to the possibility of project staff *"disengaging"* from evaluations, particularly if they *"have had a bad experience of evaluations".* She talked of *"disengagement with the process of evaluation and then a subsequent disengagement with the results of the evaluation, the report and the learning"* and obstruction as generally a result of *"disengagement rather than being more malevolent".* When dealing with these situations NGO2 and IC7 both referred to the necessity to triangulate carefully the information being received and of the analytical skills built up in an experienced evaluator.

8. Relationship between the evaluating team with beneficiaries and other stakeholders in the project being evaluated

Participants recorded in this node 18/22 – 82%

The UN Guidelines ((United Nations Evaluation Group 2005, p. 14) as Standard 3.11 states that *"Stakeholders should be consulted in the planning, design, conduct and follow-up of evaluations."* Subsequent paragraphs enlarge on this definition, emphasising the use of workshops, learning groups, debriefing and participation in field visits. The role of the learning group must include *"facilitation and review of the*

evaluation" as well as the *"dissemination, application of results and other follow up action".*

These components were central to the *"experimental and inclusive approach"* taken by Grünewald, Binder and Georges (2010, p. 153) in their evaluation of the post hurricane Mitch emergency, undertaken in 1999 and it is interesting to note that by 2005 this approach had entered mainstream thinking and standard setting for evaluations. As can be seen from some of the above quotations from participants, the methodology of building relationships with the project team has often been extended to include other stakeholders. In the specific arena of humanitarian action there is a strong current focus on the inclusion of and accountability to beneficiaries (Humanitarian Accountability Partnership 2010).

This focus was reflected in this research by specifically questioning participants concerning stakeholder participation, and specifically the participation of beneficiaries, in the evaluation process.

There was a fairly wide acceptance from participants for the need to evaluate humanitarian projects from the point of view of the beneficiary and to involve beneficiaries in the evaluation process. UN1 said *"intuitively it makes sense",* IO1 said, *"I think it is the reality".* IC1 referred to progress *"it's a hell of a lot better than it used to be"* and NGO2 said, *"There's a direction of travel that's positive."* IC2 encouraged evaluators to *"cast the net as wide as possible",* and IC4 stated, *"This is something that we impose on ourselves and our teams so we try to ensure that there will be at least some place, some room for the beneficiaries to give feedback and their opinions."* IC3 framed their reply in terms of a rights based approach referring to beneficiaries' *"right to have an opinion and the value of that opinion".* IC5 stressed the centrality of the beneficiary population in evaluation *"the object of the humanitarian action is to help the people on the ground and if they feel they haven't been helped or haven't been helped adequately I think that is pretty crucial".* He recognised there may be difficulties in deciding who to talk to and what the vested interests are *"but I mean that's the work of an evaluator to try and take those issues into account."*

IC1 referred to his own survey of evaluations published on the ALNAP website during the first couple of weeks of January each year, a sample he estimated at 30 to 40 evaluations. He then assessed each survey against *"whether beneficiaries were consulted and how they were involved (in the evaluation process)"* He reported that the results from the current year had shown 77% of the evaluations had

included the interviewing of beneficiaries in the methodology, however, a far lower number had systematically assessed accountability to intended beneficiaries. While recognising an improvement over the years he summed up these results as suggesting *"something about evaluation not keeping pace with the Humanitarian Accountability Partnership (HAP) Standards".* (Humanitarian Accountability Partnership 2010) NGO3 indicated that the involvement of beneficiaries in evaluation was *"vastly different from organisation to organisation or from evaluation to evaluation"* He continued, answering his own questions, *"how much has it become standard practice? Hard for me to say. Should it be standard practice? Definitely, I would say so"* He stressed that *"one has to be very clear beforehand what it is exactly that you are expecting from the affected population as such"* and *"the questions that were asked of them should be carefully considered."*

While participants may have widely recognised the need for beneficiary participation and some may have indicated an improvement in that direction, their assessment of the impact beneficiaries had on the evaluation process was less positive. IO1 said *"I would say that the beneficiaries probably have the least impact on influencing the report"* when compared to other stake holders such as *"the commissioner* (of the evaluation)*, the donors, the local partners and the evaluator".* IC3, asked about the involvement of beneficiaries, said, *"the short answer is relatively little, I think usually well no".* NGO2 said *"certainly there has not been enough weight put on the points of affected groups"* but he went on to indicate that *"the last few years has seen an awful lot more pressure put on NGOs to put the viewpoints of affected people at the heart of the analysis of their quality, which can only be a good thing".* He reflected that this was not just an issue for evaluation but had to start right at the planning stage with the *"perceived needs"* of the population. *"How participatory are we actually prepared to be when it comes to the choices over what intervention is the right one for that population at that time? I wish I had a nice solution for that but I don't".* IC3 identified the

> *"subjective mind-set of the people involved, not just the evaluators but those commissioning it. The vast majority would not consider any kind of genuine community based evaluation where you go to a community and say, You guys set up the whole thing, we want to know if it is worthwhile. We are not even going to do terms of reference, you tell us and work it out yourselves."*

IO1 spoke of a continuum of participation in different evaluations with a spectrum from *"very top down where beneficiaries are perceived or just observed, there is no*

involvement, to the other end where you ask the beneficiaries, do you guys think we should evaluate this and if so what exactly should we evaluate." The latter case being, he observed, close to the concept of empowerment evaluation advocated by Fetterman (2001).

GD2 spoke of his own institutional experience.

> *"There is certainly an acknowledgement within (the organisation) that we need to get much better at the impact at beneficiary level, involving beneficiaries not only in evaluation but in the whole programme cycle. We are starting that process. A lot of agencies talk about that but I've yet to see a decent evaluation process that actually involves beneficiaries at a meaningful level, not just as interview subjects. Beneficiary impact for beneficiary satisfaction: beneficiary satisfaction doesn't feature highly at all in the humanitarian evaluation process and I think it really ought to."*

He justified his position by comparison with public service providers in his own country.

> *"I can hardly imagine the provision of those services in this country without a very clear beneficiary feedback loop. You know it is so built into our psyche, the police force, health services, courts, social services, all have very strong beneficiary feedback loops and that is… largely missing in humanitarian work. I think that is a major failing and I think we could get much smarter at our humanitarian work if we tried to close that beneficiary loop."*

Several participants pointed out problems they perceived in pursuing greater beneficiary influence on the evaluation of humanitarian action. UN1 specifically highlighted the difficulty in the *"management of the expectations of beneficiaries in relation to limited resources".* IO1 pursued the same theme that the *"expectations of beneficiaries in Dadaab… Kosovo… and New Orleans after Katrina"* would be different *"and yet we have one set of standards usually against which we evaluate".* He connected the expectations and frustrations of survivors to level of subjectivity in their responses.

> *"For example if you take Katrina and the frustration and anger that many people impacted by the hurricane… had towards people who were trying to respond… I think that, in a way, can often have an effect on the subjectivity of assessments that were being done during Katrina."*

IC8 spoke of the mind-set of survivors, *"they are immediately motivated by the desire to look after themselves and their loved ones and communities and so by definition they are manipulating, not necessarily in a deliberately bad way, but in that self-interested way to do with survival"*.

NGO3 referred to the information obtained in many interviews, not just with beneficiaries as *"often perception based and in that sense subjective, but when you hear it from a number of people repeatedly"* he concluded it becomes more of a fact and less of a perception.

Ensuring the participation of a representative sample of stakeholders including beneficiaries was highlighted as a problem. UN1 pointed out *"the beneficiaries are not always a unified group; there are different subgroups and different relations of power"*. NGO1 cautioned against being *"co-opted by any one particular owner"*.

The inclusion of beneficiaries and other stakeholders in the evaluation process began as an "experimental and inclusive approach" by Grünewald, Pirotte and de Geoffroy (2001, p. 153) in 1999. The participants in this research, carried out 13 years later, recognise this approach as; necessary, increasingly, yet not systematically employed and that this important aspect of evaluation remains a "work in progress".

9. Relationships within the evaluation team

The authors who shared their experiences of the early evaluations of humanitarian actions undertaken between 1993 and 1999 (Wood, Apthorpe & Borton 2001b) reported a number of difficult relationships that had occurred within an evaluation team resulting in some compromise in the effectiveness of the evaluation process. Wood (2001, p. 42) reported leading a team when he had no say whatsoever regarding its composition. Apthorpe (2001, p. 106) reported a team that lacked cohesion till about half way through the evaluation, Grünewald, Pirotte and de Geoffroy (2001, p. 155) highlighted the relationships within the evaluation team as being more difficult than those with the stakeholders in the field while Telford (2001, p. 178) related how differences in approach and methodology in the team generated a degree of stress.

This research looked to see if similar difficulties continue to be encountered and the effect they have.

It would be difficult to imagine any field of human endeavour in which interpersonal relationships do not sometimes interfere with the effectiveness of a team and the evaluation of humanitarian action is certainly no exception. UN3 remembered situations where one team member indicated that they could not work with another because *"he's driving me crazy"*, and situations where there was *"a big falling out between people due to differences in methodology or overall approach"* or *"authoritarian leadership"*. NGO2 recalled being a *"part of evaluation teams where there have been incredibly robust conversations between evaluators based on differing analyses"* but also reported that these had been *"worked out between the evaluators"*. IC5 had *"heard of teams that have more or less fallen to bits because of personality differences"* and IC7 reported as a team leader, having a team member imposed on him who was *"completely out of their depth, could not contribute much, was increasingly withdrawn and falsified information in their output"*. NGO3 observed what he thought was *"a very strange combination* (of participants on a team) *and you know one could almost be sure that this evaluation team would run into trouble and I think they did"*.

It may be reasonably expected that the commissioners of evaluations would take care to assemble the evaluation team in a manner that minimised the risk of serious conflict. One method reported by a number of participants was to recruit the team leader and let him/her have a say in selecting the team. UN3 viewed this approach as *"another thing we might do"* but also said *"Most of our teams are actually chosen kind of ad hoc and we put them together"*, citing avoidance of bureaucratic procedures that would send the final decision on a larger contract to a committee outside the control of the evaluation department. *"Purely in pragmatic terms it is easier for us to hire consultants on individual contracts and then bring them together as a team"*. NGO1 indicated that assembling a team of evaluators who had previously worked well together as a process that had been tried but qualified the utility of this approach, suggesting that one or more members who are *"quite new to the team"* should be included so as to ensure a *"different viewpoint on a particular issue"*. IC2, looking back to a previous period in his career when he worked with an organisation that commissioned evaluations, said *"we would select the team leader and we would work with him or her to find the best team. We would not impose team members"*.

Four independent consultants, however, recounted more negative experiences. IC5 related *"in my case I have probably done 15 or 16 team leader roles and I can't*

remember a case where I was asked to be involved in the selection of team members". IC7 said

> *"Very often you won't know who the team is in advance… if you are lucky you get to travel together.* On one occasion he remembered "*We met at the hotel half an hour before we went for the first meeting. One had come from Mexico, one from the States, one from England and I had come from Geneva… We had no time to sit and talk about strategy or what we were going to plan. I'd been recruited four days before as team leader and had written an inception report with nothing more than emails from three other people: totally unprofessional.*"

IC9 said *"If I'm the team leader, I'm often, I'm normally the last person to be recruited and then I meet the team on the plane for the first time."* IC6 said from her experience there was

> *"no team building before except for some Skype calls, not necessarily they* (team members) *know each other, they usually have a lot of people they know in common"* She concluded *"I've never been team leader and I refuse to be. I want to have a longer life!"*

IC4 who regularly works with the same team said

> *"I always work within a team of people I knew before and we share a methodology. I know this is quite rare… but for me it is essential. We know each other… and we know we can work under pressure and there is a kind of hierarchy already in place and there is no internal dispute. Normally evaluators are individuals and work within teams they don't know and I heard about the difficulties."*

Several participants reported being members of evaluation teams that experienced difficulties in reaching a consensus agreement on the content and wording of the evaluation report. Participants identified different mechanisms for resolving this situation. IO1 cited the use of footnotes in the final report to identify differences in opinion, UN5 indicated the need for a *"partly democratic process leading to some kind of agreement or consensus"* but if this failed then there would be *"minority statements"* included in the report. IC3 also suggested the use of a *"minority report mechanism".* NGO2, NGO3 and IC9 all indicated that the final decision on wording was made by the team leader.

The researcher observes that many of the participants have experienced, or are aware of evaluation teams that have experienced, difficult internal working relationships and that several of the independent consultants report inadequate team preparation before the start of the field visit as common. Experience of mechanisms for the resolution of conflicts in a team regarding the wording of the final evaluation document report varies from the *"partly democratic"* with *"some kind of agreement or consensus"* to the dictatorial *"final decision on wording… made by the team leader"*.

The inclusion, in the evaluation team, of project staff members of the organisation being evaluated has been cited as a particular cause of conflict within evaluation teams.

<u>Evaluation teams with a mix of Internal and external participants as a source of subjectivity.</u>

The authors of the accounts of the early evaluations of humanitarian action (Wood, Apthorpe & Borton 2001b) identified problems that were encountered when evaluation teams included both external consultants and staff of the organisation being evaluated. Apthorpe (2001, p. 110) related the fear of the staff of the organisation being evaluated that *"they would be placed squarely in the firing-line on their return after the independents had left"*. Telford (2001, p. 181) refers to a staff member such an organisation as having to cope *"with the position of being judge and judged"*.

This research looked at the current perceptions of such mixed evaluation teams.

With a few reservations participants generally had a very positive attitude to mixed evaluation teams. UN3 gave four reasons for making use of mixed teams.

- First, was the cost, *"If you are going to have two staff members and one consultant it is going to be much cheaper than three consultants"*.

- Second, the value of internal knowledge. If a completely external team is used, *"you are paying quite a large amount of money just to learn about the organisation, how it is structured, how it functions"*.

226

- Third, the loss of knowledge, *"wholly independent evaluation teams take away an enormous amount of learning, none of which is retained by the organisation itself"*, as compared to the experience gained by members of the internal evaluation department while undertaking an evaluation with an experienced external consultant.

- The fourth, expanded on the third, the building up of the internal evaluation function, *"I will try to get the best and brightest staff members to work in this section. If I offer them desk jobs I am not going to get the best people. The best people want to actually get out there and do the work".*

NGO1 identified the role played by a team member with internal knowledge of the organisation. *"I think it is helpful to have somebody from within the organisation to apply the context and the clarity on certain issues because I think that when undertaking evaluations evaluators can make a lot of assumptions."*

IO1 explained the benefits of a team consisting of *"an external evaluator… they're leading it and are going to tell the story"*. Somebody from the local team will *"open doors… get access or meaningful access to some of the key stakeholders and areas of implementation"* He also added *"somebody from the regional office"* for a broader organisational picture. He, however, recognised that there may be *"a strong difference of opinion"* on such a team which would need to be reconciled.

NGO2 also stressed the advantages of mixed teams particularly in providing a reality check for the conclusions and recommendations of the external consultant. *"For me it is about the balance, internal and external.*

> *One of the great things is that if you can have a member of staff accompanying an independent consultant and they are jointly responsible for the evaluation… you find that… conclusions… and certainly recommendations not based on the reality of the organisation… don't come out in the report."*

He clarified that the staff member should be "no one from the immediate team (being evaluated) and identified that humanitarian workers are "hyper critical" and that although that does not necessarily "*make them objective, it*

does mean that they don't pull punches. In lots of ways what I've found is that if you do have that mixed team that the criticism is stronger from the internal source and praise stronger from the external source."

GD1 spoke of the need to protect local staff members who may work with the evaluation team in situations where criticism, particularly of government actions or policy may be dangerous. He indicated that they may be comfortable to work together with our team to help write the report but he or she may be less comfortable to present the report.

It is interesting to note that all the positive aspects of mixed teams have been related by those working for the UN, NGOs, international organisations and government donor organisations.   Some independent consultants, however, did express less positive views. IC7 indicated the need for confidential interviews without the participation of the organisation being evaluated *"We'll say thank you very much, now you can leave. We'll have the discussion without you"*. IC9 related a situation where whenever he went to see another organisation, the staff member of the organisation being evaluated *"came with me to protect the interests of the organisation or to report on me or something. It was unceasing, the interagency competition and rivalry".*

Clearly a special relationship of professional trust has to be built up between the internal and external members of an evaluation team.

10. Influence of interagency relationships in the field

Participants recorded in this node 5/22 – 23%

The relationship between organisations working in the field was commented on by Apthorpe (2001, p. 110) where *"Stand offs between them* (two humanitarian organisations) *and the resulting further tensions were manifest and, at times, experienced directly by us".* Similar tensions were referred to by IC9 above and this was not the only case.

IO1 expressed a concern that in some cases interventions were being evaluated by *"how (another organisation) feels about working with you in some context"* NGO3 expressed how interagency conflicts are hidden. *"On the surface I think agencies*

tend to be generally quite fine to each other… but when you start to dig deeper… it turns out there are many, many issues and lots of mistrust and competition… there is no question about that." IC5 spoke of the habitual mutual criticism of the UN and NGOs but said, *"I am not sure that's affecting the quality of the evaluation… It has to be taken into account and for evaluators to see whether the criticism is based on fact".* IC6 cited a case in which she found two databases maintained by different major UN organisations, which should have been cooperating, which held contradictory information.

The researcher observes that while inter-organisational relations in the field were referred to by a few participants as introducing a level of complexity into evaluation none considered this as a serious source of subjectivity, rather an irritant to be taken into account.

11. Subjective influences relating to learning from evaluations and the learning cultures of organisations

Participants recorded in this node 17/22 – 77%

In the process of institutional learning and the application of knowledge to the improvement of the delivery of humanitarian action, evaluation and the production of evaluation reports is only one component. In their conclusions to the experiences of those who carried out some of the earliest evaluations of humanitarian actions in the period 1993 to 1999 Wood, Apthorpe and Borton (2001a, p. 210) were heavily critical of the use made of evaluation reports, *"Serious efforts are needed to ensure that humanitarian evaluations are better used… some are barely read beyond the draft stage".* They criticised a lack of follow up and a lack of clarity regarding necessary follow up.

How evaluations are used and how the knowledge they contain is incorporated into best practice formed part of this research and is discussed under a number of topics. However, some of the participants, particularly the independent consultants, commented on the learning process as a whole, and several quite negatively.

IC2 said, *"I am not sure that we are so successful in terms of learning and changing behaviour… I will start by saying that if an agency calls itself a learning agency I say they are bluffing".* He added that setting up information exchange systems was not

learning and that learning only takes place if there is a *"political will to change"*. He concluded,

> *"Agencies do not learn, the system doesn't learn. It is evident that they do not change. They respond to evaluation, they either accept or not accept the recommendations, but they don't change. Next time they do the same thing... I wouldn't say it's an absolute set that systems don't learn but they certainly do it very slowly".*

IC3 said "*It* (evaluation) is an expensive but a very valuable data set. *It is incomplete in that the value has not been placed on utilising that data set and incentivising its use."* IC5 put it simply "*I would say that the other part of the evaluation, the learning cycle, is not developed enough".*

IC8 spoke at length, analysing the learning process that does take place in humanitarian organisations. He postulated that in humanitarian organisations learning is not a "*linear process or actually a logical or rational one*". He considered that learning in humanitarian organisations was difficult as the humanitarian endeavour is *"an art rather than a science and it is in the nature of art that it's beholden in the eyes of the observer and therefore there are many shades of opinion about its desired end".* He acknowledged that the accumulation of knowledge from evaluations was an important component but *"it takes an accumulation of those things and then some kind of overwhelming tipping point or trigger to actually turn that learning into a change of practice and that is not a linear process".* He said that there needed to be "*a great shock that necessitates change"* and indicated that this could be *"a scandal or some comprehensive failure that causes outrage or great disgust".* Saying that "*It's got to strike an emotional chord, it can't just be rational or logical, a lot of people have to suffer or die or some iconic person or image or belief has to be shattered".* To illustrate this point he referred to the situation in Goma in 1994 and the sexual abuse scandal in West Africa and the role of media in forcing change as crucial in imposing "*the incentives for change".* GD1 expressed some of the same sentiments, *"One evaluation doesn't make a change. It will be when the time is right, when there is a certain body of knowledge emerging, then those who have power to do something with the findings can't escape them anymore."*

IC8 also referred to the disconnect between those who do evaluations and those who deliver humanitarian assistance.

> *"The experts do it and they write big reports with many recommendations which has a disempowering effect… Eighty nine pages of recommendations* (referring to the Tsunami Evaluation Consortium), *who's going to own them and take responsibility? So therefore there is a gap between the elite evaluation industry and the plodders who are working away at the operations".*

IC7 referred to a similar *"disconnect"* between those who commission and undertake the evaluations and the *"desk officers, the operations guys, the ones who should actually be working with the product or the findings".*

Some participants concluded that evaluations had little effect in bringing about change. IC2 said,

> *"I am not sure that we are so successful in terms of learning and changing behaviour… we put forward the same lessons and recommendations for many, many years and it hasn't changed the world. So are we a success or not?"*

IC5 said, *"It is the nature of the agency, are they wanting to improve and learn or not?"* Asked if it was individual organisations or the whole institution of humanitarian organisations that had a fundamental problem with learning he said *"I think that may be correct, yes .To which you could say all these evaluations have been a complete waste of money".* IC6 said, *"Sure we make our living from doing them* (evaluations)*, but the long term value, the learning value, going back into the organisations to me is very questionable".* IC7 said, *"I'm not naïve enough to believe that we actually make a significant difference. I get paid for it but I don't actually make a significant difference".*

Two independent consultants did relate examples of more positive events. IC3 related *"a very pleasant experience… being told by a contracting agency 'We don't want firm recommendations – not even firm conclusions we want sign posts'… as to what the big issues were and the big questions that they, themselves could work on".* From the point of view of the evaluator

> *"It took pressure off ourselves… secondly it took a certain artificiality, out of the process because so much of it is artificially scientific… They took the signposts as key issues, as key questions that came out of the exercise and*

they, themselves, worked on it through workshops… in a very open and relatively honest manner."

IC5 related a situation in which the evaluators were feeding back their findings to a group of agencies. Some of the findings were negative and a few of the agencies were protesting.

"One of the agency heads got up and said, 'Look this is what we've asked for, this evaluation, we should be learning what mistakes have been made. We continue to make these mistakes, we should be listening, not just arguing with the evaluators about it.' But I have to say this is quite a rare event!"

The researcher observes that the expression of negativity by independent consultants regarding the effectiveness of evaluation and the learning process of humanitarian organisations is surprising both in its regularity and strength. Negative views are expressed in general terms indicating that they are perceived as the norm, the regular occurrence. The positive views, on the other hand, refer to individual examples indicating that they are the exception rather than the rule. IC5 defines his experience a *"quite a rare event"*.

The researcher then looked at the responses of the participants from organisations that commissioned evaluations to see if they viewed in a similar way the learning process.

UN3 related an occasion when a colleague told him *"when we were designing the programme one of the things we made a lot of use of was your previous evaluations of* (a similar programme area)" indicating that it was *"not unusual for somebody to stop me in the cafeteria"* and tell him they *"had read your new report and found it useful".* He recognised that this was not an everyday occurrence and that reading evaluations was *"down to individual initiative".* NGO2 also observed the individual initiative of reading evaluation reports, *"There are people that will read and people that won't read the report regardless of the content, regardless where it has come from".* He concluded *"I think it* (learning) *works on an individual level but it doesn't work on an organisational level.* IC8 also stressed the individual nature of change, indicating that *"one needs people of courage, the rule breakers, to make better rules or overcome bad ones and the character of rule breakers is one that requires a lot of courage".* GD2 referred to the need to be *"a bit more bloody minded"* in order to bring about organisational change.

NGO1 indicated that institutional learning was a work in progress. *"I get the sense that at the moment there is a big push by organisations to really think about learning… it is something we are all grappling with… we want to move beyond just talking about it to really ensuring that learning improves."* Specifically referring to the role of evaluations she said, *"There is maybe an issue around how we communicate the findings of an evaluation and also that an evaluation does not become in a sort of sense, it's only for the experts."* GD2 referred to the gap between evaluation reports and learning saying, *"almost by osmosis, we are all supposed to adopt these new findings and lessons."* And referred to (Hallam 2011) as the most recent publication on the subject of learning from evaluations.

The researcher observes that it is highly significant that none of the participants spoke of a clear, successful, example of systematic organisational learning and the only ones who did indicate a learning mechanism suggested that it requires a catastrophic situation to develop, courageous individuals or bloody mindedness to get beyond the assumption of learning by "osmosis". The researcher also observes that the situation revealed above is unacceptable and that there is a duty, as a result of this part of the research, to examine a novel process for the promotion of humanitarian institutional learning. IC6 made the same point, *"lots more can be done with existing evaluations, much more can be done"*.

Participants in this research spoke of a number of different aspects of the problem they perceived with the learning process in humanitarian organisations and in the humanitarian community as a whole. These perceptions are presented as individual topics below.

a.   Dissemination

NGO1 indicated a problem in this area when she said, *"I think there is maybe an issue around how we communicate the findings of an evaluation"*. IC7 highlighted the problem in more detail suggesting that

> *"five, ten, fifteen, twenty people might read the report but the organisation as a whole doesn't. People in the countries working on similar issues don't pick up and read something that might be significant to them in terms of their own programming – they don't even know about it."*

IO2 took a different view, *"the first thing is… to make sure that they (the reports) are available… In this last year, all of our evaluation reports are on our public website… So I think we don't have to worry should it be sent to this person or not to that person"*.   IC4 relates how an organisation opened what they termed an

"observatory" in Haiti where workshops were organised *"for discussion of the results of evaluations"* in which, *"when people agreed to share the results"* several evaluations on similar topics could be discussed. IC4 also referred to "*trying to write some briefings and articles… based on some evaluation reports"* that could be disseminated.

The researcher, however, views the approach to dissemination of "make it public, put it on the web" as in many ways an abdication of responsibility. Google, at the time of writing, returns over 14 million hits for the search words "evaluation Haiti" from which it is concluded that this is a case of "hiding information in full view". The researcher considers the more direct and targeted approach of "*briefings and articles… based on some evaluation reports"* reported by IC4 as likely to be much more effective.

b.   Selectivity for the public domain

This is an extension of the topic above, "dissemination". Here the focus is on the publication of evaluation data for wider public consumption rather than ensuring it reaches those who could directly apply it to planning and implementation of future humanitarian actions.   Although best practice, from the point of view of accountability and transparency, may be to make all evaluation information public, selection criteria are sometimes applied.   Some of these criteria appear more justifiable than others.

UN1 explained that, *"evaluations* (of his organisation) *are in the public domain. We do a little bit of self-censorship, particularly when it comes to assessing the actions of governmental counterparts".* This, he explained, was to avoid possible *"adverse impact"* on the organisation's ability to deliver protection and assistance. NGO2 reflected a similar mechanism *"There have been two reports, one that is published and one that is internal, or parts of which are held back from publication".* Again this was explained on the basis of possible hindrance to the organisation's operations. GD2 reported,

> "*We publish absolutely everything unless there is a national security dimension… or if it would affect our international relations with certain countries… in which case we would redact any sensitive information. These are the caveats we have. If it is just embarrassing to* (the organisation) *then we go ahead and publish."*

NGO3, IC6 and IC9 all related stories concerning individual organisations that they named (two organisation promoting accountability and one donor organisation) that had at various times suppressed evaluations which were critical of some of their work. The researcher has not been able to verify these stories so no direct quotations are made in this research. However, as this research is focussed on subjectivity, the fact that three highly experienced participants perceived the reputation for transparency of these organisations to be so tarnished is considered important. In a more general context UN5 related that

> *"If the report has been critical of the organisation and its leadership… they will basically want it to be forgotten as soon as possible, which means it will go on the most remote shelf in the archives and never be heard of again,"*

IC7 expressed similar perceptions *"When something comes back that's vaguely threatening or challenging to an organisation, suddenly you find… it didn't get circulated"*

From the above the researcher observes that while there may be justification for withholding information contained in evaluations pertaining to those outside the organisation who have the power to influence negatively the delivery of humanitarian action, there is far less justification for withholding criticism of the organisation itself. The researcher finds the approach reported by GD2 "*If it is just embarrassing to* (the organisation) *then we go ahead and publish"* to be the most appropriate.

c.   Implementation of recommendations

The presentation of recommendations for improvement may be one of the main reasons for undertaking evaluations but several others may be identified. UN3 clearly presented this view, *"Look you know, providing recommendations that are implemented is maybe one legitimate function of evaluations but there are at least six other reasons why you might want to do it".* He then listed four, organisational learning, individual learning, support for advocacy efforts and team building. He added, *"I think, to be honest, recommendations are often the weakest element of any evaluation report. They are often very simplistic,"* he quoted one such, which was very specific to the organisation he worked for: *"the organisation should do more for a particular group".*

IO1 described the system in place in his organisation, *"the management responds to all the recommendations whether they are going to be followed up or not"* indicating that the organisation has a choice as to which recommendations it considers appropriate for follow up and which are not. NGO2 quotes some recommendations received as being *"completely null and void and all it does is to serve to demonstrate the fact that the consultant doesn't understand the organisation"*.

Where evaluations result in weak or inappropriate recommendations a simple count of recommendations implemented is obviously inappropriate. However UN3 indicated that at least one major donor has, *"what I call the compliance approach."* Which he explained as, *"You do an evaluation, it produces a set of recommendations and then sometime afterwards you look to see whether they have been implemented or not and you come up with a sort of score sheet"*. He stated that one may come up with a figure of say 72.6% but *"It is a completely meaningless figure. What does it mean? Does it mean they were implemented in a sustainable manner? Where did you get the information from anyway?"* He described a methodology by which a questionnaire is sent to the field where there is a lot of subjectivity surrounding the term implemented. Those responding to the questionnaire will not say all were implemented but do not want to say only half were implemented. UN3 concludes *"So almost inevitably you are going to come up with about the 70, 71, 72 per cent figure. It kind of looks and sounds about right. It makes you look good without boasting too much."*

UN4, from experience of working in emergency response situations, suggested that many evaluations are unnecessary anyway. *"They tell you the obvious and the obvious parts were going to be closed (resolved) anyway… and the sophisticated or new recommendations are just left"*. He justified this saying that if *"you are awash in funding and so on you would implement them but otherwise, when the limelight moves away and the attention moves elsewhere, everybody is comfortable… it would be good to have A, B, C, D but we can leave them."* And *"that's the nature of the business. It's chaos. It's trying to bring order, some semblance of order, in a chaotic environment"*.

UN4 concludes *"I think generally we should have less evaluations. I think we should have more monitoring",* a topic to which this research will return when discussing "real time evaluation" where in the opinion of the researcher, the distinction between monitoring and evaluation becomes less distinct. UN4's experience, the researcher

observes, seriously questions the usefulness of the evaluation process where one is *"bringing a measure of humanity, always insufficient, into a situation that should not exist"* Rieff (2003).

d. <u>Influence of evaluation being perceived as an administrative requirement or "Box ticking exercise</u>**"**

Contrary to the wish of UN4 above that we should have less evaluations, the current movement is towards more, particularly a type described by UN1

> "*Then you have another type of evaluation… I call tick the box evaluations and the first ones to do so, but they are not the only ones, was* (name of major donor)*… If there is a contract beyond a certain amount they will oblige the operation to carry out an evaluation through an external consultant."*

This approach to evaluation moves the emphasis away from evaluation as a tool for learning and towards becoming simply a bureaucratic requirement. The title given by UN1 to this type of evaluation strongly suggests that the "operation" will have a subjective approach to the process with the goal of getting a report, politically engineered to show the commissioner what they want to hear, and so maintain the flow of funds. As this description was used by one of the first participants to be interviewed it was decided to include a direct question on the topic of *"tick the box evaluations"* in all subsequent interviews. The aim of this question was to identify if this description proved to be widely perceived. If so it could be concluded that the value of evaluation, specifically as a learning tool, was being undermined or as UN1 put it *"turning into a stale ritual".*

UN3 saw this development as an expression of the *"symbolic function"* of evaluation, *"I have always thought that one of the primary purposes of evaluation, if not the primary purpose, is a symbolic function."* UN3 put this in the context of, as quoted earlier, *"it makes it look as if we are taking a self-critical look at what we are doing and trying to improve ourselves and in a kind of performance oriented culture that's important".*

Several participants in the research identified donors as being the driving force in a developing "tick the box" motivation for evaluation. IO2 asked *"Do we have to do more evaluations because beneficiaries are asking for more evaluations or it's because donors are asking? Practically I think it's driven by donors".* He believed

that regular evaluation was *"a good practice"* but recognised that *"if you just… do it as a box ticking exercise then… you will just have ticked the box, made copies and consultants are happy and that's it"*. IC7 presented a similar analysis, *"I think very often evaluations are done… to tick a box, because donors are expecting it… it has to be done, therefore it's done but it's not necessarily done with any real intention"*.

IC8 recognised the pressure coming from donors but also presented a justification. *"So it's not so much ticking the box, it is the pressure of donors on implementers, it's the pressure of parliaments on donors."* As an example he cited that

> *"in Canada there's 100% coverage because of the public accounting act… you better evaluate properly… when needed… to have enough findings, enough information to provide the tax payers with what has happened with the money spent"*.

IC8 concluded, however, *"this kind of evaluation ritual is devastating because it's sucking up lots and lots of funding but it didn't create more information"*. Looking at it from the point of view of a donor GD2, when asked about donor driven evaluations leading to a *"tick the box"* approach said,

> *"Yes, I can see that, but certainly it's not the message we are trying to get out. Our key message… for any evaluation partner is that we go for quality not quantity… You know, half the time when people ask me about evaluation I just tell them not to bother because they haven't thought about the evaluation questions, how it's going to be used. In fact they get the impression we're not going to use it"*

In responding to this question two participants in the research identified the need to have a clearly defined purpose for each evaluation if *"ticking the box"* is to be avoided. NGO3 identified "*The question needs to be asked: Why are you doing an evaluation? What's the purpose of the evaluation?*" and as a result the need to be *"much more selective, or much more strategic in making choices of doing an evaluation"*. IC7 identified the problem as *"I think I find most people don't really know what they want and why they want it. So it is a box ticking exercise"*.

UN4 said,

> *"I think there should be an objective reason why you are doing an evaluation… there has to be a reason why you are doing it and not just automatically all the time doing it… but it is a requirement otherwise you will not get the next funding".*

IC2 referred to organisations that have "*a policy to evaluate everything or a certain percentage of our programmes*" and concludes *"That's when evaluation starts becoming a ritual"*. UN5 said, *"So it ends up having very limited value and it becomes something that you have to do because you have to do it, rather than something that has a meaningful approach to take things forward"*. He concluded *"the risk is of course that it has no actual significant purpose, it's part of the way you do business. Have done* (the evaluation) *and then you write the report, put* (it) *on the shelf."*

NGO2 took a more nuanced approach. He recognised that *"there's a sense that we should be evaluating – we must tick that box."* He emphasised *"the evaluation process for those being evaluated, being part of that evaluation, there is an awful lot of immediate reflection and although not documented well, that reflection turns into shifts in programming"*. He, however, recognised that *"We are still stuck in the trap of evaluation reports once written, box ticked, and being never looked at again"*.

Two independent consultants expressed their frustration with *"ticking the box"* IC6 said, *"If it's ticking off donor requirements it's, I think, the saddest part of evaluation for evaluators"*. IC7 answered from a different perspective,

> *"You know, we're trying to move away from ticking the box because everybody does evaluations, because it is so many millions of dollars. And we said… what are you going to do with this thing, where is the follow up?"*

From the responses above, the researcher observes that a "ticking the box" mentality is widely recognised throughout the humanitarian evaluation community. While it is recognised as a reality, largely, imposed by donors, it is considered detrimental to the fundamental purpose of evaluation as a learning tool. There is recognition of the motivation for greater accountability behind the donor pressure for more evaluations but GD2 also recognised that these efforts may become counterproductive. The relationship between accountability and learning is studied later in this research.

Perception of ownership of an evaluation by the organisation being evaluated

The concept of ownership of an evaluation was another issue raised by UN1 early in the research which was incorporated into subsequent interviews. UN1 discussed the issue in terms of,

> *"You don't always have ownership for every evaluation but you know if it is only an external party that has an interest then nobody in* (name of organisation) *cares about it. It is considered yet another top down exercise and nobody will pay attention… they will try to slip it under the carpet or do passive resistance to it".*

Participants who commented on this issue addressed it in terms of building a sense of ownership among the staff of the project being evaluated using the same mechanisms identified in 6 above. NGO1 identified,

> *"I think what is really helpful and what we try to build into evaluations is having some kind of time with partners or with the team we are working with to feedback initial findings so that there is a sense of inclusion".*

In a similar way IC4 identified the role of a workshop or debriefing session during an evaluation as giving *"the chance for people to respond"* to the evaluation findings. A *"second objective is to work together on the recommendations... to think with them on what is feasible and what would be the best."*

The researcher observed that the sense of ownership of an evaluation process resulted in a sense of ownership of the recommendations. Yet ownership is a perception that has to be built into the evaluation process by a conscious effort of the evaluation team, whatever the motivation for the evaluation itself may be. The project's sense of ownership of the evaluation may be the ultimate measure of success in the building of a good relationship between the evaluation team and the staff of the project being evaluated.

f.	Influence of the readability of the report and the means of presentation

Wood, Apthorpe and Borton (2001a, p. 189) identified, *"Another aspect of humanitarian evaluations that tends to be given insufficient attention is that of getting the report and its findings accepted and used".* Grünewald, Pirotte and de Geoffroy (2001, p. 158) *"decided that it would be helpful to add a photographer to*

the team… This was decided after having witnessed an emotional presentation of an evaluation mission based on a slide show". In light of these observations the researcher decided to include discussions with participants on their perceptions of the best way to present evaluation findings.

UN3 spoke at length on this topic saying,

> *"I just place a high degree of emphasis on readability. Sexy titles and attractive formats and an easy read, because we have all got choices as to what we read. We have all got more on our desks than we can ever possibly read so, for me, it's a kind of competitive market edge thing. You have got to make something sufficiently attractive that somebody will pick up your document rather than another document"*

He identified *"short paragraphs"* as important and that *"bulk has an intimidating effect",* recommending evaluation reports of *"30 to 35 pages long"* and being *"absolutely against… short reports and then lots and lots of annexes"* He commented on having "a fantastic designer" who transformed the appearance of evaluation reports that may have *"quite a wide readership"* for what was really a small cost. UN5 was even more drastic with regard to length *"We all know that more than four pages nobody reads anyway, so these 200 page academic look alikes they have no impact".* GD2 emphasised *"writing it in plain English, keeping it short"* and specifying these things in the terms of reference.

NGO4 returned to the original context in which this topic was raised by Wood, Apthorpe and Borton (2001a), that of being able to get past the barrier created by negative findings, referring to *"being able to find a minimum common denominator that does not lower the quality of what you are doing but is able to push the exercise forward".* NGO3, however, took a different view

> *"If an organisation did a crap job, will the evaluation tell them so or will it be… more relative in its critique in order to encourage learning… When I am expecting a harsh verdict or judgement and in fact... it's sort of well softly, softly… I am wondering actually so what is being done in order to clean it up."*

The researcher observes that presentation has two aspects that will impinge on the effectiveness of an evaluation. The first is whether people are going to read the report or whether they maybe "intimidated" by its bulk, language or presentation. A

short, less inclusive, report that is read will be more effective in bringing about change than a massive, much more inclusive, report that intimidates potential readers. In a world where so much information is obtained in "screenfulls" from the web UN5's four pages may not be too short for an executive summary and UN3's estimate of 30 to 35 pages for the whole report may not be unreasonable. The question of the strength in which criticism is presented is, however, a much more difficult issue to judge. Too hard and all that may be generated is defensiveness resistance and rejection; too soft and there will be little or no impact. It appears to remain an important part of the commissioner – evaluator relationship to find the space between rejection and irrelevance, always assuming, in each situation, that space exists at all.

12. Subjective influence of the inclusion of multiple objectives of "accountability" and "lessons learned" into one evaluation

Participants recorded in this node 18/22 – 82%

The Development Assistance Committee (1991, p. 5) states

"*The main purposes of evaluation are:*

- *to improve future aid policy, programmes and projects through feedback of lessons learned;*

- *to provide a basis for accountability, including the provision of information to the public.*"

The subjective nature of this dual aspect of evaluation came to light in the 2001 evaluation of UNICEF's Kosovo operation. Telford (2001, p. 175) reports that, *"While it had originally been a DFID initiative it soon became a joint effort"* (between DFID, the donor and UNICEF the implementing organisation). Telford (2001, p. 178) identified that,

> *"Objectives and expectations were not entirely identical for the two organisations and their members of the team: a hard hitting 'accountability' focussed evaluation was expected by some staff of the donor agency, while a less contentious 'lessons learned' exercise was sought by the implementing agency".*

The terms of reference were described by Telford (2001, p. 175) as *"vague"* which is clear from the TOR summary (p. 173) where *"The overarching purpose of this evaluation were: To assess the extent to which UNICEF's programmes met their*

objectives (accountability*) and to draw lessons from this assessment for future improvements* (lessons learned)…"

In their conclusions Wood, Apthorpe and Borton (2001a, p. 210) suggest that *"evaluation activities"* should be more closely focussed on learning rather than *"being out on an 'accountability limb, as so many of them appear to be."* Both indicate a perception on the part of the authors that accountability and lessons learned are different evaluation objectives that may be hard to reconcile. This research looked at current perceptions of this issue.

GD2 clearly identified how the two objectives of evaluation affect the design of the TORs,

> *"I agree, however, that the design of evaluation terms of reference for lessons learning look very different from ones designed for accountability. So the questions you would ask and how you would approach it* (the evaluation) *would be very different".*

GD2 recognised that evaluations are rarely as simple as such a categorical answer would suggest,

> "*Evaluation can be a messy business; you do evaluations for all sorts of reasons, including aspects of both lesson learning and accountability. You just have to try and fudge it and make it work… much of the time. Be aware of the conflict, of the contradiction between the two" (GD2).*

From this response the researcher observes that GD2 finds himself working in a very similar environment to that described by Telford in 2001. While clearly separating two possible objectives of evaluation at an intellectual level, GD2 recognises that this is rarely possible at the practical level and indicates a considerable level of subjectivity in the resulting *"messy business"*, having to *"fudge it"* and working with conflicts and contradictions of which he can only *"be aware"* and not reconcile.

IC2 identified a similar potential for subjectivity in the issue of accountability vs. lessons learned- including the influence of the objective on the questions asked. IC2 also added a third category of evaluation that has already featured in this discussion of results, the ritualistic (box ticking) evaluation.

"Why are you evaluating. Is it a ritual, is it a real desire to learn or is it a demand for accountability? That will already flavour… the questions you ask or the questions you do not ask". (IC2)

IC2 continued by explaining how the focus on either learning or accountability affects the evaluation. *"The ritual is usually standard, learning you go into the process, accountability you go more into the result."*

NGO1 identifies that *"There is sometimes that tension between the accountability and learning aspects"* and IC1 referred to it as, *"a creative tension… trying to do both".*

NGO2, IC4 and UN1 used interesting terminology in connection with lessons learned and accountability. NGO2 in referring to the lessons learned approach used the term *"peer review"*, by which the researcher understands an evaluation team as a group of practitioners with similar interests and experience as the project team whose work they are reviewing and whom they are advising. IC4, UN1 and IC9 conversely all used the term *"audit"* when referring to the accountability approach to evaluation with UN1 also using the term *"inspection",* both terms suggesting a level of enforced external judgement*. IC9 related *"that's how the evaluators were feared… the evaluators are coming, the police are coming, the auditors are coming."* To the researcher the use of these terms emphasises the polarisation between lessons learned and accountability and the very different power dynamics that may be perceived in the two processes.

UN3 observed that evaluations commissioned by the implementing organisation will tend to take a *"consultative approach with the emphasis on lessons learned rather than the emphasis on accountability",* while recognising that "*amongst donors* (accountability) *is the dominant mode of thinking".* This observation closely mirrors the "peer review" and "audit" analysis above. NGO1 expressed the confusion of purpose of evaluations.

"Is it for accountability, is it to prove to the donor that we have done what we said we would do, and had the impact that we said we would have, and is that the only reason why we seem to be doing it, or are we doing it for the purpose of learning and for continual improvement or both, for our own internal purposes as well as for the purposes of the donor".

Some participants in the research presented something of a compromise. UN1 said, *"Lessons learned and accountability are not completely separated",* IC1 spoke of the *"balance between accountability and lessons learned"*. GD1 stated that *"lessons learned and accountability are two sides of the same coin"* but also recognised that *"there is, of course, a trade-off between the learning aspect and the accountability aspect."*

IC9 identified the current political emphasis on accountability making it impossible to undertake a purely lessons learned evaluation. "*Accountability, now; nobody would dare say they weren't doing that or interested in that!"*

The researcher observes that the dual objectives of evaluation, accountability and lessons learned that follows directly from the OECD-DAC guidelines, continue to generate conflicts and contradictions when, as appears usual practice, both are combined into a single evaluation exercise. Some commissioners of evaluations attempt to avoid, or at least minimise, the conflicts and contradictions by emphasising one objective over the other, a form of *"fudging it"* which may, however, cause some confusion regarding the power dynamics of the evaluation process. However, the political reality of a universal emphasis on accountability cannot be avoided.

13. Influence of the culture of evaluators and the "western" origins of evaluation

Participants recorded in this node 12/22 – 55%

The nine experiences of undertaking some of the earliest evaluations of humanitarian actions (Wood, Apthorpe & Borton 2001b) were all undertaken by evaluation teams led, and almost entirely consisting of, individuals from traditional donor countries in Europe and North America. Of the 14 authors of these accounts 12 were male and all appear to be from traditional donor countries.

UN3 reported

> *"When I started doing evaluations in the early 90's with* (name of colleague), *it was considered completely normal and acceptable that* (name of colleague) *and I would go to Nicaragua to do an evaluation… both male, both white, both*

around the same age and neither of us speaking Spanish. That was considered not a problem. Most evaluators had that kind of profile."

Indicating a change had taken place since then UN3 continued,

"I think we are much, much more sensitive now. We are definitely always gender balanced and…we always try to make sure that there is at least one person on the team who can interact with people in their own language and can translate for others on the team."

As a result of the North American roots of evaluation and of the North American/European origins of many evaluators, the researcher recognised two culture based factors that may be perceived to influence perceptions of the value of evaluations: the western cultural origins of evaluation itself and the similar origins of evaluators. In this research participants were asked how they perceived "culture" to influence the evaluation process.

A number of participants commented on the "western" origins of evaluation. IC3 said *"Even the whole concept of evaluation as it is done is a very westernised idea"* NGO3 indicated, *"The culture related to evaluations is very Anglo-Saxon, I mean evidence based".* IC5 recognised the origins of the criteria used in evaluations but speculated on their universal basis, *"the OECD-DAC criteria – they seem to be western because they are done with a kind of western codified managerial… academic approach. But I am not sure that the content or the actuality of them would be different in a different culture".*

GD1 reported a growing opportunity for cultural diversity in evaluation practice.

"In the past 20 years we have seen the emergence of evaluation societies in Africa… in Asia… and in Latin America. Those societies are native born and they contain membership of local evaluators… they have had discussions on what extent do we have culturally sensitive evaluation in our region"

IC2 expressed a similar view, indicating that although western concepts and methods of evaluation have been imposed on Africa we are now,

"seeing evaluation departments being created in countries like Uganda, Zambia and some West African countries. The African Evaluation Association… is really talking about creating an African evaluation culture. I

have no idea… how different that would look but there is certainly a wish to do so."

IO2, however, held quite a different position,

"It's a fairly western set of criteria that are being used by evaluators and… most of the evaluators come from western countries and if they are not coming from western countries they… would have earned their credentials in western countries… I don't see, right now, an alternative to this practice which would be delivering results that would… improve the quality of evaluation."

The researcher observes that participants from across different types of organisations working in the field of evaluation of humanitarian action recognised the "western" origins of evaluation and the application of "western" criteria in its implementation, but differed in their perception of how these origins and principles may continue to influence the practice of evaluation into the future. The researcher observes that, while the nature of evaluation may continue to be largely dictated by the needs of donors two distinctly different influences may come into play.

- Traditional donors and implementing organisations are beginning to demand evaluation more focussed on beneficiary impact. IO1 acknowledging this, stated "*evaluation now should involve more of the anthropological approach, people who are really familiar with the local context and the local stakeholders".*
- The emergence of new donors as major players in the humanitarian scene. Will the Gulf States and possibly China adopt or adapt the existing models for evaluation or develop their own?

IC5 speculated that *"there probably is a different framework"*. In relation to considerations of impact and objectives he suggested that the way these topics are considered *"may be different". However,* this and the possibility that *"there isn't the desire to, probably to, look at results*" should not *"stop questions being asked"*

Two participants focussed their comments on the effects of cross cultural relationships between the evaluators and their informants. IC3 rhetorically asked *"How much are you culturally, class wise, gender wise, looking for certain things before you come to any view at all?"* Using the interview in progress as an example he explained *"We are communicating now on some implicit understandings as to what is acceptable for you and what is acceptable to me, what you expect from me*

and what I expect from you" and how this may be different if the conversation was with *"somebody in a Somali refugee camp in Northern Kenya".* IC6 used a hypothetical example of her interviewing an Afghan woman, she would have her preconceptions of the needs of the Afghan woman; and the Afghan woman would have her own perceptions of what *"a white blond woman"* wants to know. IC6 viewed these cross cultural experiences in terms of *"interests*" and the *"evaluator has to learn to balance everything, including him or herself".* IC6 interpreted the skill to operate cross culturally and to check everything against the *"coherence of the whole evaluation"* as an important skill required of an evaluator of humanitarian action.

14. Subjective influences resulting from the use of the OECD-DAC Criteria used for the evaluation of humanitarian action

Participants recorded in this node 16/22 – 73%

Early in the interviews conducted for this research participants started to refer to the criteria specified by the Development Assistance Committee (DAC) and known as the "OECD-DAC Criteria". These are defined in Development Assistance Committee (1991, p. 4) *"The aim* (of an evaluation of a project, programme or policy) *is to determine the relevance and fulfilment of objectives, developmental efficiency, effectiveness, impact and sustainability"*. The difficulties of applying these criteria to humanitarian evaluation was referred to by O'Keefe et al. (2001, p. 22) in their 1993 evaluation of projects in Somalia where they interpreted relevance as appropriateness, found efficiency "*difficult to define because intervention in disaster was immediately necessary irrespective of cost implications; and effectiveness was difficult to judge not least because of lack of baseline data and the disappearance of the beneficiary population through death or movement."* They were also only able to use rehabilitation as their measure of sustainability.

The OECD-DAC criteria were further defined for use in the context of complex emergencies in Development Assistance Committee (1999, pp. 30 - 32) and two more criteria of *"coverage"* and *"coherence"* added and defined. These are the OECD-DAC criteria applied to the evaluation of humanitarian actions and form a basic component of many terms of reference. NGO2 suggested that the automatic inclusion of all the OECD-DAC criteria into terms of reference can be counterproductive, *"I'm guilty of this as a commissioner as well… you get ten days*

248

both to travel and to speak to everybody and answer against six or seven criteria each of which could probably take…two or three days." The researcher observes that the universal acceptance of the OECD-DAC criteria for evaluation of complex emergencies makes it hard for commissioners of evaluations to justify the exclusion of any of them, therefore they find it far simpler to include them all. Over ambitious terms of reference are considered by many participants to be detrimental to the evaluation process.

Two elements of the OECD-DAC criteria were raised by participants as potentially introducing a subjective element to evaluations; the attribution of causality to the changes observed (impact) and the interpretation of efficiency in terms of "value for money" by some donors.

a. Impact and causality

These elements of evaluation are the subject of Terry Smutylo's satirical song "Output, Outcome, Downstream Impact Blues" (Patton 2002, pp. 153-154) which refers to *"impact"* as an *"obsession"* through which donors *"want to see, a pretty little picture of their fantasy"* in order to continue funding and urges donors to *"wake up from your impossible dream"*. Such a satirical song may appear out of place in academic writing; however the researcher believes that the position of the composer and performer as Director of the Evaluation unit at the International Development Research Centre in Ottawa justifies its inclusion. Obsession, fantasy and dreams are certainly words that emphasise the subjective rather than the objective nature of evaluating the impact of development projects. This research looked at the perceptions of how the inclusion of impact in the TORs for evaluations of humanitarian action was perceived by participants.

UN1 said *"One thing that perplexes me of the DAC criteria… is the issue of causation and impact"* He used the analogy of a randomised control trial as the evidence of a direct causal relationship and concludes that *"This is obviously something that is a measure that in the humanitarian world we don't have".* GD2 concluded that *"In the social sciences and certainly in the humanitarian world it's much more messy".*

IC3 clearly showed frustration when working with what he termed the "*higher level criteria… sustainability and impact you are in no-man's land and you are into the twilight zone. Basically we haven't a got a clue and where we think we've got a clue it's because of subjective analysis which you can't avoid anyway!"* IC5 expressed the view that, *"Those* (OECD-DAC) *criteria have been useful. I think the question of*

impact has always been problematic and I think causality is always problematic". IC7 also identified *"impact is the one* (OECD-DAC Criteria) *I have the biggest problem with"* GD2 differentiated between attribution and contributions

> *"Wishful thinking, meaningless… at an individual project level there's so much of what can be described as background noise: lots of other interventions, lots of other issues going on at the same time which makes meaningful attribution quite difficult. I've always struggled anyway with the distinction between attribution and contribution."*

IC1 indicated the cause to be *"the DAC criteria being pushed to too low a level beyond where they make sense. You can answer impact issues with a different kind of study, not with a project evaluation."* The researcher observes that the last two participants have highlighted the root of the problem: whereas the overall impact of the humanitarian intervention may be demonstrated in an evaluation, such as that of the Indian Ocean tsunami (Telford, Cosgrave & Houghton 2006), it is more difficult, perhaps impossible, to attribute impact in each of the multitude of smaller evaluations that have been undertaken at project level in the case of the intervention in response to the earthquake in Haiti.

   b.   Efficiency and value for money

The OECD-DAC criteria include these aspects but use the term *"cost effectiveness"* rather than 'value for money' a concept that participants in the research identify as a focus of one particular government donor.

> *"Efficiency measures the outputs – qualitative and quantitative – in relation to the inputs. This generally requires comparing alternative approaches to achieving the same outputs, to see whether the most efficient process has been used. Cost effectiveness is a broader concept than efficiency in that it looks beyond how inputs were converted into outputs, to whether different outputs could have been produced that would have had a greater impact in achieving the project purpose."* (Development Assistance Committee 1999, pp. 30-31)

NGO1 indicated that her organisation was actively pursuing the concept of value for money but had reservations. *"We are having some internal discussions about the whole value for money discussion and breakdown of unit costs. We are looking at*

developing our own internal policy of an understanding of what that means". She then indicated that the project was *"not just about the financial aspects"* but about the "*social, environmental and economic value"* that is "*delivered to the beneficiary*". The researcher recognises that this response comes from an organisation that receives money from the donor in question, therefore is compliant but with reservations.

Independent consultants, however, took a more critical line. IC1 used the most strident terms referring to value for money as *"a fraught mix"* adding that he was *"just appalled when suddenly* (name of donor) *was throwing its weight around and pushing value for money, I think it confuses things. You can see where it is coming from politically, protecting them against taxpayers."* IC6 said value for money is *"one of the most difficult things to do and it can be easily abused"* She continued to express how strongly she felt about the issue, *"I have started to take it out of my assignments and terms of reference"* and referred to value for money as *"a can of worms".*

The researcher observes that value for money is a complex issue that both those who have to convince donors of the efficiency of their projects in order to attract funding and those who undertake evaluations find difficult. When the economic, climatic and logistical context in which aid is delivered is taken into account, a simple division of cost by number of beneficiaries cannot be applied. If pressed to evaluate value for money, none of the participants in this felt they were able to do this in an objective manner.

15. Influence of the emergence of "Real time evaluations"
Participants recorded in this node 7/22 – 32%

The concept of real time evaluation (RTE) has only recently been introduced into the evaluation of humanitarian action. GD1 thought the first RTE to have been carried out in Darfur in 2005. Just four years after this, Cosgrove, Ramalingam and Beck (2009) wrote the ALNAP "Pilot Guide" to real time evaluation in which they define RTE as *"an evaluation in which the primary objective is to provide feedback in a participatory way in real time (i.e. during the evaluation fieldwork) to those executing and managing the humanitarian response".* As such, RTE is a form of evaluation, perhaps the first, that has evolved directly from the specific needs of humanitarian action to address the rapidly changing emergency environment in which humanitarian action operates. The concept of real time evaluations rapidly

gained wide acceptance amongst the humanitarian community and by 2013. The use of the ALNAP database of evaluations made while working on Chapter 5 of this thesis revealed that almost all the evaluations of the response to the earthquake in Haiti were termed "real time". How much did this reflect a genuine change in the methodology of evaluation and how much simply the adoption of a new terminology, a new buzz word? UN3 suggested that the latter may sometimes be the case.

> *"When we started to do real time evaluations the* (head of agency) *really liked it because it was action orientated… And this is kind of sexy… so he started to call evaluations real time evaluations even when they weren't real time because it's got more sense of dynamism, energy, concern for emergency performance."*

NGO3 was sceptical, *"I have never seen a real time evaluation because they are done after the fact… or most of them at least. I know very few that are really done in that sense, in real time."* GD2 also highlighted a note of caution, "*It's important to note the date on them (RTEs) and the period to which they refer. In many circles RTEs are actually just normal evaluations published a year or two after the event.*"

The participants who highlighted such negative aspects of RTEs were in the minority. NGO2 recognised that RTE had become a *"common tool"* but that it had to be applied in a timeframe which matched the context in which the project worked. He saw the problem as "*trying… to nail down standard operating procedures, but there is no standard operating environment*" yet recognised the advantage of recommendations that could be "*utilised in terms of learning immediately"*. IC7 related a positive experience of RTE bringing about immediate results in operations, *"We were having these discussions today and by tomorrow certain things have clearly been put in process to change".* However the experience was less positive when it touched on management structure. *"When we were looking at structure stuff there was much more blockage to wanting to change".*

NGO4 saw RTE as a real change in the role of evaluation.

> *"Real time evaluation is about countering the very extractive nature that used to be associated with the evaluation exercise… The emphasis is on the project team who is left at the field level… It's about giving the signal that they are the one in the driver's seat and you are just trying to boost this."*

IC8 took the emphasis from the evaluation to the learning *"It's not just RTE but also more real time learning… that was the most exciting part of it."*

In discussing the immediacy of the learning from RTEs with participants, the researcher recognised that the advent of RTE changed both the role of evaluation and of the evaluator. In some of the later interviews the researcher took the opportunity to question participants on whether RTE was in fact a form of monitoring and whether the evaluator became an advisor or counsellor to the project team. GD1 saw a distinct difference, *"monitoring is an internal process… RTEs are normally evaluations done by an external team."* He continued *"the thought behind RTE is, if you are having an external team reviewing and evaluating what is going on, people in the field are far too busy to do that themselves."* GD2 recognised that RTE *"is distinct from monitoring but I suspect… there's a high degree of overlap there. Evaluations always look beyond the log frames and indicators and monitoring indicators."* GD2 then identified the key question of an evaluation as *"Are we doing the right thing?"* He concluded that, *"|Any good evaluation that does attempt to address that question shouldn't be regarded simply as another monitoring tool"*.

When it came to discussing the possibility of a radical change in the role of the evaluator and the researcher suggested the word counsellor GD1 responded

> *"I wouldn't use the word counsellor because then you are becoming part of the solution of the problem and then if you are evaluating sometime later you evaluate your own counsel. So there goes your independence! So the real time evaluator is not a friend here to help you they are an independent… person who looks at the situation based on TORs does his work, writes a report and gives feedback."*

The researcher observes from the nature of the RTE process that the relationship between the evaluator and the project team may well be different from that when the evaluation is being undertaken post facto. This is an element that may well be a subject for later research and it may be that a changing role for evaluators in RTE is slowly emerging.

Chapters 4, 5, 6 and this chapter have presented the results of this research obtained from

- The comparison of the evaluations of the humanitarian response to the forced displacement from Rwanda to DRC and the forced displacement from Darfur into Chad (Chapter 4)

- The comparison of the evaluations of the humanitarian response to the Indian Ocean tsunami and the Haiti earthquake. (Chapter 5)

- Baseline case studies of the experiences of undertaking evaluations between 1993 and 1999 (Chapter 6)

- Interviews conducted with humanitarian professionals actively involved in evaluation (Chapter 7)

In the following chapter, when discussing Part 2 of the research the results obtained from the interviews will be synthesised with the framework developed in Chapter 2; that of trading of principle-agency with structure for operational-agency and as a result being instrumentalised.

Chapter 8 - Discussion and Analysis of the results of this research

In this chapter the researcher applies analytical frameworks to the results obtained from the research to provide a synthesis that forms the basis for the recommendations that follow in Chapter 9.

Part 1 of the research

The recommendations identified in Chapters 4 and 5 are discussed in the context of selected literature cited earlier in this work.

Part 2 of the research

In Chapter 7 data collected in the interviews are analysed using the framework of the 15 subjective influences identified in Chapter 6. In this chapter the data are further analysed using the principle-agency, operational agency model developed in Chapter 2. This synthesis of data collected, the results of the literature review, the baseline case studies, together with the experience of the researcher demonstrates the ways in which the evaluation of humanitarian action continues to be instrumentalised.

Contextualisation of the analysis

The findings of this research must be placed in the context of two documents that have played a major role through the preceding chapters: Wood, Apthorpe and Borton (2001b) from which the baseline case studies are drawn, and Ramalingam, Scriven and Foley (2009) whose identification of the perception that *"evaluations do not tell us anything new"* (p. 9) was the origin of this research.

Some of the influences identified in this research, particularly those relating to team selection, terms of reference, relationships within the team and the problem of follow up on the recommendations of the evaluation, are the same as those identified by Wood, Apthorpe and Borton (2001a, pp. 187-207). In order to ameliorate the resulting problems these authors made seven recommendations:

i. Recognition of the complexity of the context in which humanitarian action is delivered

ii. Clearer focus of evaluations on either accountability or lesson learning

iii. More input from the evaluators to the terms of reference

iv. Renegotiation of the terms of reference following an initial field assessment

v. Resources allocated to evaluations should be adequate to fulfil the terms of reference

vi. A change in attitude to evaluations from that of audit to learning opportunity

vii. Better mechanisms for the follow up of evaluations and the lessons they identify. (pp. 208-211)

This research has shown that, with the possible exception of the first, the same recommendations could easily be made at the present time, some 13 years after their publication. This provides another indication of the resistance of the humanitarian evaluation learning system to respond to criticism and recommendations.

In their analysis of the problem that *"evaluations do not tell us anything new"* Ramalingam, Scriven and Foley (2009) use a framework developed by Argyris and Schön (1978) in proposing that the recommendations made in the evaluation of humanitarian action rely heavily on *"actions to correct mismatches and errors based on practices, policies and norms"* (p. 10), a process Argyris and Schön (1978) refer to as *"single loop learning"*. Ramalingam, Scriven and Foley (2009) suggest that a more innovative approach would be evaluating, and making recommendations that change organisational *"practices, policies and norms"* (p. 10) or even change the *"overall organisational rationale and context"* (p. 10) of the organisations themselves. These levels of change are termed by Argyris and Schön (1978) as double and triple loop learning respectively. Ramalingam, Scriven and Foley (2009) propose that these levels of organisational learning would lead to higher levels of innovation and suggest this as the route for improvement of humanitarian action.

This research takes a different approach, by identifying a broader range of reasons why evaluations do, sometimes, appear to not tell us anything new. This research also analyses the way in which the subjective factors identified in the preceding chapters influence the evaluation of humanitarian action, the institutional learning process that follows and their ability to be an effective tool in bringing about change.

Discussion of Part 1 of the research

Analysis of the comparisons undertaken in Chapters 4 and 5 demonstrates that there is some repetition of themes throughout the evaluations of humanitarian responses spanning the period 1996 to 2010. Hence the proposition that *"evaluations do not tell us anything new"* identified by Ramalingam, Scriven and Foley (2009, p. 2) appears have some credibility. Our analysis presents four reasons why that repetition is present.

1. **Evaluations are used as a form of advocacy for change in issues outside the control of humanitarian actors**

   Advocacy is a tool used by organisations to increase their agency when they are constrained by structure. Evaluations may be used by both the evaluators and the commissioning organisation as advocacy tools, building a constituency for change by the repetition of the criticism and/or recommendation from many evaluations of a wide range of organisations. Although the humanitarian community does not have the authority to make the changes advocated directly, a large enough constituency can bring movement in the desired direction.

   The recommendations often appear to be expressions of principle-agency but may have an underlying motivation of increasing operational-agency.

   Some examples emerging from the research are advocacy for changes in the:

   a. Action of governments to prevent or ameliorate conflict: the principle-agency component is the reduction of human suffering but also present is the operational-agency motivation of improved access and security, enabling larger humanitarian interventions to take place. This is identified as the principle finding of the JEEAR by Borton, Brusset and Hallam (1996, pp. 161-162) where clear recommendations are made.

   b. Actions of governments as donors: the principle-agency component is a more predictable, flexible availability of humanitarian aid to beneficiaries, with the increased operational-agency that such funding would bring to humanitarian organisations. The inconsistency of the availability of funding

either excessive or too little is highlighted by several authors including: Borton, Brusset and Hallam (1996, p. 163), Telford, Cosgrave and Houghton (2006, p. 122) and DARA (2010, pp. 160-162)

2. Recommendations of evaluations are acted upon but changes in the humanitarian environment necessitate further recommendations and subsequent change.

Structure both facilitates and constrains. Agency and actors will work to maximise the facilitation while minimising the constraints. Two aspects of this appear in the humanitarian sphere: problems in coordinating humanitarian organisations and difficulties in cooperating with military actors.

a. <u>Coordination of humanitarian organisations</u>

Agency requires freedom of action. Humanitarian organisations look to coordination to give them a free space in which to operate, constraining other organisations from directly competing with them. At the same time, the organisation does not want to give up its agency to take advantage of whatever opportunities may exist for its own growth. A manifestation of this phenomenon is the amount of time in coordination meetings spent on information sharing, which organisations use to mark out their territory. Whereas coordination is usually justified on the basis of principle-agency (i.e. best service to the beneficiaries) it is used to maintain and where possible maximise operational-agency.

The emergence of a large number of diverse humanitarian actors taking whatever advantage they can from coordination mechanisms, while allowing themselves to be constrained by coordination to the minimum extent possible, is at the root of the problems encountered in coordination. Mechanisms established over a decade ago to encourage participation in coordination have been overwhelmed by the increased numbers of humanitarian actors responding to major emergencies. The emphasis on recommendations for improved coordination has shifted to cope with this situation. Some important elements of this are discussed in Interagency Standing Committee (2010, p. 17)

Coordination continues as a work in progress. As such it should be expected to feature in evaluations as an indication of the requirement for further development. The repetition in evaluations should be viewed as the necessary pressure to keep the process moving forward, perhaps not as fast as many would like, but nonetheless moving forward to overcome new challenges.

b. <u>Cooperating with military actors</u>

The role of the military in the delivery of humanitarian aid is complex, involving deep philosophical and ethical considerations. It is, however, most basically, the involvement of a very large, powerful, government controlled actor in the humanitarian scene. All humanitarian actors recognise, and often regard with awe, the operational-agency of the military while at the same time being cautious of, or even despising, what they perceive as the lack of principle-agency of the military in humanitarian affairs. The variety of roles in which the military can be deployed and the manner in which these impact on the delivery of humanitarian action confuses and disturbs traditional humanitarian actors.

- When deployed in peacetime with a purely humanitarian mission, particularly in response to natural disasters, the application of military capacity to humanitarian action can be achieved with little or no instrumentalisation. Lengthy and detailed provisions for the use of foreign military and civil defence assets in disaster relief are to be found in the Oslo Guidelines United Nations Office for the Coordination of Humanitarian Affairs (2007)
- In peace keeping roles it was the inaction of the military in Rwanda and Bosnia that humanitarians criticised. Humanitarians encouraged their presence in the Zaire refugee camps and in Kosovo as agents to allow improved humanitarian action. Both situations resulted in considerable controversy and are commented on critically by Cockayne (2006, p. 6) and Rieff (2000, p. 26)
- When deployed in combat situations, such as in Iraq and Afghanistan, the use by the military of humanitarian action in a hearts and minds role, for the purposes of force security, intelligence gathering and the achievement of the force commander's objectives, must be viewed as an act of complete instrumentalisation. The principle is stated clearly in

United States Army (2009, p. 13) An extreme example of this is the quotation from Lieutenant General Victor E. Renuart Jr. in Hollis (2004)

The result is polarisation within the humanitarian community. The more Dunantist organisations maintain their principle-agency by avoiding contact with the military and taking advantage of its presence where they can, while remaining at the greatest distance. The more Wilsonian organisations increase their operational-agency, taking the maximum advantage of the military's assets while keeping only the distance necessary to maintain their own identity. In behaving in this way, they accept a much higher level of instrumentalisation.

Repeated recommendations in evaluations for improved guidelines, as well as joint civil-military training and exercises, encourage some actors to engage more closely, but serve to increase the resistance of others. The recommendations remain necessary, however, as the role of the military changes in response to the global political-military situation. The relationship between humanitarian organisations and the military and the need for continued improvement is highlighted in Borton, Brusset and Hallam (1996, p. 166), Bennett et al. (2006, p. 45) and Patrick (2011, p. 7).

3. **Fundamental changes to the accountability structures in which humanitarian action is delivered**

Accountability is another example of structure which both constrains and enables the agency of humanitarian actors. The notion of accountability in humanitarian action has, however, been one of fairly rapid change.

- Prior to 1990 humanitarian organisations were relatively few in number and regarded largely as good causes warranting support, with relatively little accountability. Concerns regarding accountability were expressed in terms of a lack of accountability to, and the paternalistic attitude of humanitarians towards, beneficiaries. One of the earliest discussions of this is one of the main themes of 'Imposing aid'. (Harrell-Bond 1986)
- During the early 1990's increasing humanitarian budgets and a changing governmental culture resulted in the imposition of higher levels of accountability by donors on humanitarian organisations. Evaluation was used as a tool to accomplish this objective, a link made by Wood, Apthorpe and Borton (2001c, p. 12).

- By the mid-1990s concerns were being expressed that accountability to, and dependency on, government donors (upward accountability) should be balanced with accountability to beneficiaries (downward accountability). Returning to the theme of 'Imposing Aid' Borton, Brusset and Hallam (1996, p. 168) made a proposal that became known as the 'Humanitarian Ombudsman' and then provided the origins of the Humanitarian Accountability Partnership as documented in Humanitarian Accountability Partnership (2013d)

- More recently emphasis has been placed on a wider range of accountability, often focussing on the beneficiaries, but also including local populations outside the scope of humanitarian projects. A wider definition of humanitarian accountability is given by Humanitarian Accountability Partnership (2010, p. 1). At the same time, there is a trend amongst recipient states to take stronger control of, and demand accountability from, humanitarian organisations operating in their territory.

Humanitarian organisations feel their principle-agency being constrained more and more by accountability structures with which they must compromise in order to maintain or increase operational-agency. Humanitarian organisations have adopted coping strategies to maximise their operational-agency but in each case encounter some degree of instrumentalisation. For example:

1. Prioritisation of accountabilities: responding to whichever is the most powerful at any time, taking into account alliances being formed between those demanding accountability, e.g. donor and host government or host government and local population.

2. Playing-off one accountability against another to increase the power of the organisation, e.g. negotiating with the host government and local leaders to agree to a course of action and using the agreement to strengthen their negotiating position with the donor.

Such strategies remain undocumented but the two quoted above have been observed in the experience of the researcher.

4. Unimplemented recommendations

However generously reasons may be assigned for the repetition of recommendations in evaluations, there remains an obvious gap in institutional learning in the transformation of recommendations from evaluations into practice. Many recommendations have not been implemented and this failing is a major reason why the same recommendations are repeated time after time. The phrase "lessons learned" should be dropped from the humanitarian vocabulary until such time as this gap is closed, rather than allowing its continued use to suggest that no gap exists.

One criticism of the recommendations made in evaluations cited by participants in this research was that many of the recommendations were vague generalisations rather than specific points for action. This may be attributed to instrumentalisation of the evaluation process (as discussed later in this chapter) which constrains the agency of the evaluators to make honest direct criticism and clear practical recommendations. The highly politicised environment of evaluation of humanitarian action is one in which all the actors know the rules that guarantee the sustainability of the present system. Evaluators are constrained to operate within those rules or lose both their influence and income.

Discussion and analysis of Part 2 of the research

An analysis of the accounts of undertaking nine evaluations of humanitarian actions (Wood, Apthorpe & Borton 2001b) were used as case studies to identify influences which appeared to limit the objectivity of the evaluation and learning process. This resulted in the identification of 14 subjective influences (see Chapter 6) which were increased by the inclusion of a 15th, aspects of real time evaluations, that emerged from the interviews conducted during this research. Consideration of these subjective influences identified them with some of the same causal elements as in the instrumentalisation of humanitarian action. They indicated elements of the structure in which humanitarian action takes place, acting in their own interests, to influence the evaluation of humanitarian action, and hence influencing humanitarian action itself. As a result, a broader literature review was undertaken on the instrumentalisation of humanitarian action in both the historical and current context

and the results analysed using a modified "agency structure" framework identified in Chapter 2.

Two of the 15 subjective influences on evaluation of humanitarian action were found to involve a contractual relationship and are analysed in a "principal agent" framework in the form of the principal agent problem defined by the Financial Times (2013) *"The problem of motivating one party (the agent) to act on behalf of another (the principal)".* These two relationships are:

- The relationship between the commissioner of the evaluation (principal) and the evaluation team (agent)
- The relationship between the commissioner of the evaluation when the commissioner is the donor (principal), and the project being evaluated (agent)

In the evaluation of humanitarian action, the task of the evaluation team is defined in the terms of reference (TORs) which are drafted by the commissioner and which the evaluation team often has little opportunity to influence. These terms of reference are criticised, in both the baseline case studies and the interviews conducted, as often being:

- Imprecise documents containing more objectives than can be accomplished in the time available
- Sometimes appearing threatening to the staff of the project being evaluated.

In the case of the independent external evaluation of humanitarian action, the principal agent problem sets up a particularly difficult dilemma. This is best illustrated by modifying the Financial Times definition of the principal agent problem, by the inclusion of the underlined text, as:

The problem of motivating one party (the agent), <u>while retaining their independence of judgement and action</u>, to act on behalf of another (the principal)

The difficulty is in understanding the dilemma of retaining independence of judgement and action while acting on behalf of another. When the commissioner of the evaluation agrees with the judgement and actions of the evaluators no conflict results, but when there is a disagreement, compromise becomes necessary and inevitably some measure of the independence of the evaluator is lost. The commissioner of the evaluation retains the power to compromise the independence

of the evaluation while still publishing the evaluation report as that of an external independent evaluator.

The incentives for the evaluation team to act on behalf of the commissioner are often included in the TORs or in contractual documents, in which, the withholding of a sizeable final payment until a satisfactory report is submitted is stipulated. The commissioner retains the right to define satisfactory in this context. The baseline case studies as well as participants in the research report that negotiations regarding redrafting to meet the commissioner's definition of satisfactory are common.

The sanctions, which remain undocumented, include the withholding of future commissions from evaluation teams that produce unsatisfactory drafts and are reluctant to make the changes demanded by the commissioner. In a clear case of this behaviour IC4 reported her organisation being blacklisted by a commissioner for a lengthy period of time.

The relationships between the commissioner of the evaluation, the project being evaluated and the evaluating team are, in practice, not as simple or distinct as may appear from the way they are presented above. Two variants are often encountered:

- The donor may commission the evaluation of a project being undertaken by a humanitarian organisation (Fig 8.1A)
- The donor may give funding to the organisation undertaking the project for the organisation to commission an external evaluation of the project. (Fig 8.1B)

Figure 8.1 Relationships between donors, humanitarian organisations, projects and evaluation teams
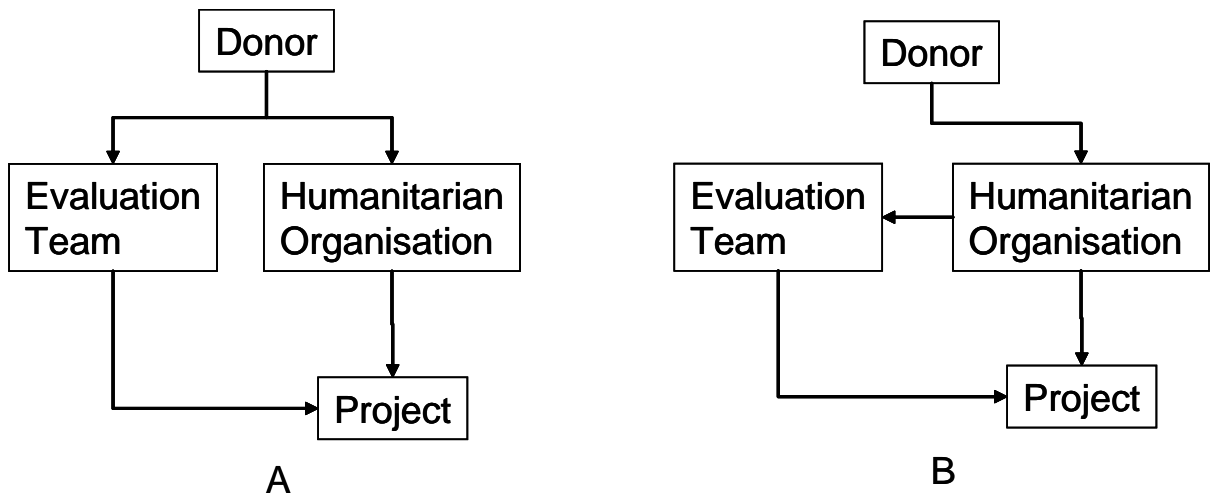
Source: Author

In the first of these variants (see fig 8.1 A) the principal, the commissioner of the evaluation, is the donor and the agent is the evaluation team. The evaluation team is working as the agent of the donor. The TORs are prepared by the donor with varying input sought from the humanitarian organisation. The humanitarian organisation holds very few incentives or sanctions to influence the report submitted by the evaluation team. The humanitarian organisation and the project being evaluated may both feel threatened by the evaluation process.

In the second variant (see fig 8.1 B) the principal, the commissioner of the evaluation, is the humanitarian organisation carrying out the project with funding received from the donor. The evaluation team is working, primarily, as the agent of the humanitarian organisation. The TORs are prepared by the humanitarian organisation, almost certainly with some consultation with the donor, of which the humanitarian organisation is, itself, an agent. The TORs may require the evaluation team to meet the donor at an early stage in the evaluation process. The humanitarian organisation holds the contractual incentives as well as the sanction of withholding future contracts, although the evaluation team may remain acutely aware of the influence of the donor in obtaining future commissions. Both the humanitarian organisation and the project being evaluated are likely to feel much less threatened in these circumstances and more likely to cooperate with the evaluation.

In the baseline case studies, the evaluations are largely commissioned by donors or, in the case of Telford (2001), rather disastrously, as a cooperation between the donor and the humanitarian organisation. The interviews with participants in this research revealed that commissioning of evaluations by humanitarian organisations is currently more common.

Principle-agency operational-agency analysis of subjective influences on the evaluation of humanitarian action

In this section the 14 subjective influences identified from the case studies and subsequently confirmed as being experienced in undertaking current evaluations, together with the 15th identified in the interviews, are analysed. The framework used in this analysis is that of the organisational trading of principle-agency for operational-agency and in the process being instrumentalised that was used in Chapter 2

The analysis presented in detail below and summarised in tabular form in Annex 1.

1. **Subjective influences are perceived as universal in the evaluation of humanitarian action.**

In both the baseline case studies and the interviews with participants there is wide acceptance that subjectivity is an inevitable component of the evaluation of humanitarian actions, with little suggestion that it should be recognised as such and resisted.

Principle-agency compromised.

- Evaluators derive their principle-agency from the supposition that they are impartial independent observers who objectively report their observations and recommendations to elements of the humanitarian structure (donors and humanitarian organisations) which are then able to act on these observations and recommendations.
- An acceptance that subjectivity is a universal element in the evaluation of humanitarian action lowers resistance to the subjective influences that are known to exist. As a result, evaluation of humanitarian action continues to be

perceived as more of an art than a science and any attempt to present it as a science is perceived to be misguided.

<u>Operational-agency gained</u>

The acceptance of subjectivity in the evaluation process constrains evaluators from being dogmatic in their criticism of projects, such criticism can be easily in turn criticised as simply the subjective opinion of the evaluator. It encourages positive feedback, veiled criticism and vague recommendations that are unlikely to be challenged by the commissioner or subject of the evaluation

<u>Resulting instrumentalisation</u>

Those involved in the evaluation of humanitarian action routinely accept subjective influences as part of the environment in which they work, and there is general recognition that evaluation is less than completely objective. Yet evaluation terms of reference and evaluation criteria (See Development Assistance Committee 1991) are presented as objective and evaluation reports are normally presented as an objective assessment. The permanent record of evaluations therefore masks the extent to which subjectivity is recognised and accepted by participants in the process.

<u>Comment</u>

Subjective influences should be acknowledged for what they are - pressures to be recognised and resisted – thus maximising the objectivity, the key principle-agency, of the evaluation process. It should not be assumed that professional judgement based on the evidence available is unavoidably subjective.

2. **Perceived influence of the sense of urgency surrounding humanitarian action.**

The influence of the sense of urgency that surrounds humanitarian action, especially in the immediate response to a humanitarian crisis, was highlighted in the case studies and identified in the research interviews as a continuing influence that limits the objectivity of evaluations.

<u>Principle-agency compromised</u>

An acceptance that everything in humanitarian action must be undertaken urgently allows compromise of professionalism and thoroughness to be excused or even encouraged by all who participate in the process.

- Humanitarian responders compromise planning and cooperation with local structures
- Commissioners of evaluations compromise selection procedures and refinement of TORs.
- Evaluators agree to short evaluation and reporting deadlines that do not allow a thorough undertaking of the evaluation.

<u>Operational-agency gained</u>

- Humanitarian responders gain operational-agency by rapidly issuing appeals and submissions for funding and gaining media exposure in the immediate aftermath of the disaster.
- Commissioners of evaluations are able quickly to field teams of trusted evaluators.
- Evaluators working with over ambitious TORs justify selectivity of topics to be covered, shallow coverage, veiled criticism and recommendations while retaining their status with the commissioner as trusted evaluators.

<u>Resulting instrumentalisation</u>

Commissioners of evaluations maintain almost complete control over the evaluation product, with little or no risk of severe criticism, while being able to issue a positive independent external evaluation.

3. Perceived influence of the attitude, personality and interests of those undertaking the evaluation.

From both the baseline case studies and the interviews conducted for this research, it is widely accepted that commissioners, team leaders and members of evaluation teams permit themselves and/or others to act in an unprofessional manner by allowing their personalities and interests to influence their professional judgement while undertaking the evaluations of humanitarian action.

Principle-agency compromised

The acceptance that influencing of the substance of an evaluation by the attitude, personality and interests of those undertaking the evaluation is normal allows the acceptance of compromise of the principles of professional detachment and the separation from individual bias.

Operational-agency gained

The evaluator is free to pursue their own interests while disseminating their own particular bias. The evaluator may expect repeat commissions from organisations which agree with and appreciate their world view.

Resulting instrumentalisation

The commissioner of the evaluation can rely on the evaluator to produce a uniformly acceptable product.

Comment

The clear distinction between special experience or knowledge of an evaluator necessary for them to conduct a particular evaluation and the contribution of attitude, personality and interests should be clearly understood and carefully maintained.

4. Inter-organisational and interpersonal relationships

This is a general topic which is broken down further into types of relationships in subsequent headings. (See sections 5 to 10 below)

Principle-agency compromised

Allowing such relationships to impinge on objectivity compromises the independence of the evaluator or the evaluation team.

Operational-agency gained

Minimising conflict or the potential for conflict avoids risk of personal criticism which may impinge on further contracts.

<u>Resulting instrumentalisation</u>

Any exploitation of an interpersonal or interagency relationship that compromises the objectivity of any party in an evaluation allows a level of instrumentalisation to take place.

<u>Comment</u>

The building of open positive working relationships between all participants that allow a free exchange of information and ideas is a fundamental tool in a good evaluation process. It is the politicisation and exploitation of interpersonal relationships that results in instrumentalisation.

5. Relationship between the organisation commissioning the evaluation and the evaluation team

This is a contractual relationship already analysed in the context of the principal agent problem, with a focus on the dilemma of retaining independence while working in the interests of another. As a result it is identified as the most influential relationship in the evaluation process.

<u>Principle-agency compromised</u>

It is almost inevitable that some level of independence will be compromised when the evaluation is commissioned by an organisation, some aspect of the work of which will be the subject of the evaluation.

<u>Operational-agency gained</u>

- External evaluators depend, for further contracts, on maintaining a good working relationship with their clients, the commissioning organisations.
- The commissioners may also withhold significant final payment until they receive an evaluation report that they consider to be satisfactory.

<u>Resulting instrumentalisation</u>

- Commissioners of evaluations maintain an unequal power relationship with external evaluators through which they can influence the content of the report while issuing it as both external and independent.

- Evaluators seeking to become established in the profession are in the weakest position to defend their independence. Established evaluators have gained knowledge of their clients and can adjust their level of independence to meet the needs of the clients they most want to retain. A few of the most senior evaluators have the prestige to retain a high degree of independence and will be commissioned for evaluations in which their name and reputation outweigh other considerations
- The level of compromise required may be greatly reduced when evaluations are commissioned jointly or by a consortium as was the case with both the JEEAR and the TEC evaluations.

5a Subjectivity introduced by the terms of reference for an evaluation

This is a sub-heading of 5 above which deserves separate analysis. The TORs are under the control of the commissioner of the evaluation and usually imposed on the evaluator. TORs are often written by the inclusion of contributions from various departments in the commissioning organisations with little control regarding excessive content.

Principle-agency compromised

- The commissioner of the evaluation compromises the thoroughness and integrity of the evaluation by issuing TORs too comprehensive to be accomplished in the time available for the evaluation.
- The evaluator colludes in this compromise by accepting the TORs and the timeframe allowed for the evaluation.

Operational-agency gained

- The commissioner of the evaluation maintains their internal working relationships with all departments of their own organisation by allowing a wide range of contributors' suggestions to be included and avoiding conflicts that may be encountered in prioritising out any contributions.
- The evaluator gains the contract, which they may lose if they try to renegotiate the TORs, as seeking another evaluator who will accept the TORs may be preferred by the commissioner to a renegotiation process.

Both the commissioner and the evaluator collude in the knowledge that the TORs are unreasonable and can justify a relatively shallow evaluation which avoids controversial issues that may otherwise have been the subject of criticism and recommendations.

6. Relationship between the commissioner of the evaluation and the project being evaluated.

Project staff may perceive the commissioner of the evaluation as being either focussed on accountability or more interested in lessons learned from the project.

Principle-agency compromised

Project staff will compromise their honesty with the evaluator, tailoring their information to suit what they believe the commissioner wants to hear from the evaluation.

Operational-agency gained

The commissioner of the evaluation usually has control over resources available to the project and the project seeks to ensure continued resourcing.

Resulting instrumentalisation

Project staff seek to control the evaluation through the information they make available, the field visits they facilitate and the stakeholders to which they facilitate the access of the evaluators.

7. Relationship between the evaluating team and the project being evaluated

This is judged to be the second most important relationship in the evaluation process, as the evaluation team is often highly dependent on the project being evaluated for accommodation, travel and access to informants. The effects of this relationship may be similar to those of 6 above, but control of the relationship is much more firmly under the influence of the evaluation team

<u>Principle-agency compromised</u>

A heavy reliance on the subjects of the evaluation (for the physical requirements necessary to undertake the evaluation) can result in compromise in the independence of the evaluation in reviewing critical areas of the project, determining the locations to be visited and the informants to be accessed.

<u>Operational-agency gained</u>

The evaluators complete the evaluation in the time allowed and are able to submit a report by the deadline, without risking negative feedback from the project being evaluated.

<u>Resulting instrumentalisation</u>

The subject of the evaluation is able to control and manipulate the evaluation process in order to gain a positive evaluation and ensure continued funding.

<u>Comment</u>

Building an open cooperative working relationship with the project being evaluated is an essential part of the evaluation team's work but this relationship must be based on trust, transparency and a willingness to be frank with each other. A good cooperative working relationship in the evaluation should form the basis for recommendations being agreed and implemented.

8. Relationship between the evaluating team and the other stakeholders in the project being evaluated.

The term other stakeholders here refers to all stakeholders other than the commissioner of the evaluation, the donor, the evaluation team and the project being evaluated. The beneficiaries of the project are included as important stakeholders, but local authorities, leaders and populations may be included even where they are not beneficiaries of the project.

<u>Principle-agency compromised</u>

Diverse opinions expressed in a widely distributed beneficiary population may communicate contradictory or difficult to reconcile information. Information gained from beneficiaries and other stakeholders may be considered politically motivated,

dominated by self-interest and unreliable. Difficulties in triangulation of information results in subjective selectivity in inclusion or rejection of data gathered.

<u>Operational-agency gained</u>

A shorter, simpler report, more acceptable to the project being evaluated and more easily understandable to the commissioner of the evaluation, is produced.

<u>Resulting instrumentalisation</u>

Simplification of complex and contradictory information results in the elimination of the more controversial and de-emphasising of the more critical information. The report tends to favour positive rather than negative feedback from the stakeholders.

9. Relationships within the evaluation team

Good working relationships should not be difficult to build in small teams of evaluators working together for a short period of time and reliant on each other for successful completion of the evaluation. However teams are often hurriedly put together and made up of individuals who have little previous knowledge of each other and are in competition for future contracts.

<u>Principle-agency compromised</u>

Hurriedly put together teams are given little or no time to work together with the TORs to agree methods, priorities etc. before deployment to the field.

Diverse, strongly held and voiced opinions on a project's successes and failures compromise a unified report.

<u>Operational-agency gained</u>

Where conflicting relationships or views have developed in an evaluation, the team leader can justify negotiating with the commissioner the relative emphasis to give the diverse opinions which appear in the report. In doing so the team leader can fend off criticism of his/her leadership role by producing a report more closely tailored to the wishes of the commissioner.

The commissioner influences the evaluation report to a greater degree than would otherwise be the case thus avoiding controversial, critical opinions.

10. Influence of interagency relationships in the field.

Interagency rivalries in the field may result in one or more parties trying to use the evaluation to denigrate another. These situations are usually obvious to the evaluators and so are relatively low risks.

Principle-agency compromised

Potential to compromise objectivity if the opinion of one organisation is uncritically accepted

Operational-agency gained

Relations with the project being evaluated can be eased by ignoring or strongly downplaying criticism made by other organisations.

Resulting instrumentalisation

The project being evaluated may be able to fend off criticism as unjustifiable interagency prejudice.

11. Subjective influences relating to learning from evaluations and the learning cultures of humanitarian organisations.

Institutional learning is identified in both the baseline case studies and the interviews conducted for this research as being a deficient, and largely dysfunctional, component of the evaluation/learning process. The effects of inadequacies in the institutional learning culture were encountered so frequently in this research that it is analysed both as a general topic here and under four sub-headings. (See sections 11a to 11d)

Principle-agency compromised

Innovation and constant challenging of established techniques in the search for better ways of delivering humanitarian action are compromised.

<u>Operational-agency gained</u>

Continued use of simple, low risk funding procedures for well-tried methodologies known to be accepted by large institutional donors.

<u>Resulting instrumentalisation</u>

Projects are designed around donor and/or organisational norms rather than being based on individual needs assessments that take into account beneficiary specific requirements.

<u>Comment</u>

Institutional learning is complicated by

- The highly fragmented nature of the humanitarian sector in which an organisation is only able to change its own methodology
- The lack of a widely accepted guideline formulation and updating system in the humanitarian sector
- Institutional learning is identified as a component separate from the evaluation process in a simple feedback loop analysis later in this research.

11a. Dissemination of the evaluation report and selectivity in putting the evaluation reports into the public domain

Best practice in the evaluation of humanitarian action is for the widest internal dissemination of evaluation reports within the commissioning and subject organisations coupled with the placing of evaluation reports in the public domain. The baseline case studies and the interviews with participants reveal that selectivity continues to be exercised.

<u>Principle-agency compromised</u>

Transparency and full accountability are compromised by selecting only those evaluations which reflect well on the organisation`s activities and suppressing dissemination and public availability of evaluations, or parts of evaluations, that are critical of the organisation's performance.

The organisation maintains an artificially high level of credibility for good performance and, to the extent that lessons are learned internally, they are not shared with others, providing a competitive edge for future projects proposed by the organisation.

Resulting instrumentalisation

Organisations selectively use the principle of transparency in evaluations to their own advantage, publicising success, hiding failure and treating improved methods of delivery of humanitarian action as trade secrets.

11b. Influence of the evaluation being viewed as an administrative requirement or "box ticking" exercise

As evaluation of humanitarian action has become a generally accepted part of the project cycle it has become a standard enshrined in institutional rules and in some funding contracts.

Principle-agency compromised

The value of evaluation as part of a feedback loop is compromised. The completion of the evaluation is perceived as an end in itself rather than a means to a greater end.

Operational-agency gained

The obligation to complete an external, independent evaluation is fulfilled with minimum effort on behalf of the organisation.

Resulting instrumentalisation

The evaluation process is instrumentalised to fulfil the requirements of organisational rules or the terms of a funding contract. Shallow evaluation and vague recommendations are encouraged while quantity may be substituted for quality with the encouragement of lengthy annexes. If evaluation is imposed on a project as a prerequisite for future funding then the project is likely to see the evaluation process as ritualistic and its relevance only that of creating the best impression possible, rather than being an honest reflection of the project.

11c. Perceptions of the level of "ownership" of the evaluation results taken by the organisation being evaluated

As evaluations become ritualised they become perceived by the subjects of the evaluation as an external exercise that will have little relevance to or impact on their work. An individual, group or organisation can participate in a ritual without having any sense of ownership of that ritual. Ownership of a ritual rests with those who preside over it rather than those who participate in it.

Principle-agency compromised

Ownership of the evaluation process and responsibility for the recommendations made by the subjects of the evaluation are compromised.

Operational-agency gained

The subjects continue working, safe in the knowledge that the evaluation will have minimum impact on them or the project.

Resulting instrumentalisation

With little or no sense of ownership on the part of the subjects of the evaluation ritualization of the evaluation process may be assumed to be complete and the evaluation marginalised in its role as a measuring instrument in an evaluation/learning feedback loop.

Comments

The sense of ownership of an evaluation process gives a sense of ownership of the recommendations. However, ownership is a concept that has to be built into the evaluation process by a conscious effort of the evaluation team, whatever the motivation for the evaluation itself may be. The sense of ownership by the project of the evaluation may be the ultimate measure of success of the building of a good relationship between the evaluation team and the staff of the project being evaluated.

11d. The influence of the readability of the report and the means of presentation

A short, less inclusive, report that is read will be more effective in bringing about change than a massive, much more inclusive, report that bores potential readers with detail and intimidates potential readers with hard hitting criticism.

Compromise for the sake of readability, however, impacts on the external independent nature of the evaluation process.

<u>Principle-agency compromised</u>

Depth of evaluation, solid explicit criticism of deficiencies observed and detailed recommendations may be compromised in order to achieve readability.

<u>Operational-agency gained</u>

Wider dissemination and learning are encouraged and the evaluation process gains credibility as a learning tool.

<u>Resulting instrumentalisation</u>

Brevity and readability implies selectivity in the information included and careful nuancing of criticisms made, in order to avoid those being criticised from defending themselves by attacking the credibility of the report.

<u>Comments</u>

- In a world where so much information is obtained in "screenfulls" from the web then the four pages, suggested by one participant, may not be too short for an executive summary and the 30 to 35 pages, suggested by another, for the whole report may not be unreasonable.

- The question of the strength of criticism presented is, however, a much more difficult issue to judge. Too strong and all that may be generated is defensive resistance and rejection, too weak and there will be little or no impact. It appears to remain an important part of the commissioner – evaluator relationship to find the space between rejection and irrelevance.

12. Subjective influence of the inclusion of multiple objectives of "accountability and "lessons learned" into one evaluation

The inclusion of into the same evaluation is specified as by According to the Development Assistance Committee (1991, p. 5) accountability and lessons learned are the "main purposes of evaluation".

The baseline case studies and participants both agree that accountability always gains supremacy.

Principle-agency compromised

The emphasis, or perceived emphasis, on accountability rather than lessons learned compromises the use of evaluation as a primary source of information for institutional learning. Evaluations become perceived as an extension of audits and inspections.

Operational-agency gained

The institutional need for narrow accountability based on documented plans and proposals is satisfied.

Resulting instrumentalisation

An emphasis on accountability and a perception of evaluation as an extension of an audit or inspection results in evaluation being used as a tool to demonstrate compliance with documented plans and submissions. The successful outcome of a project is viewed as a result of such compliance. Lessons learned may be perceived as demonstrating that the project could have been better formulated in the planning phase and hence reveal a lack of accountability.

13. Influence of the culture of evaluators and the "Western" origins of evaluation

Evaluation as a tool has been developed within the cultural framework of the second half of the 20th century Western democracies and reflects the institutional needs of this environment. This tool is used to evaluate projects, largely funded by organisations which share this cultural base, but are implemented in very different political, economic, cultural and social environments. Currently evaluators are

predominantly, themselves, products of the culture that developed the concept of evaluation.

<u>Principle-agency compromised</u>

The ability to contextualise concepts of impact and success into alternative political, economic, cultural and social settings are compromised by the use of a 'one size fits all' evaluation methodology.

<u>Operational-agency gained</u>

Projects are evaluated in a framework understood and accepted by the commissioners of the evaluation and the funders of the project.

<u>Resulting instrumentalisation</u>

Evaluations approach the project from the viewpoint of the commissioner and donor rather than the social, cultural and economic context in which the project is implemented i.e. that of the beneficiary and other local stakeholders.

<u>Comments</u>

Currently no alternative evaluation framework or methodology exists and local evaluators are being trained in, possibly indoctrinated into, the established evaluation framework. As local evaluators develop and local organisations commission their own evaluations alternative approaches may develop.

14. **Subjective influences resulting from the use of the OECD-DAC criteria in the evaluation of humanitarian action.**
- The OECD-DAC criteria used for the evaluation of humanitarian action comprise seven topics: efficiency, effectiveness, impact, connectedness, relevance, coherence and coverage. TORs often require all seven to be covered in a single evaluation. This is considered a major contributing factor in over ambitious TORs as commented on above.
- The attribution of impact to a humanitarian project is particularly difficult in short term emergency situations.

<u>Principle-agency compromised</u>

- The inclusion of all the evaluation criteria in TORs compromises thoroughness and encourages selectivity when insufficient time is allowed for the evaluation.
- Forced attribution of impact to a project favours short term, planned, impact and compromises the reporting of longer term, unplanned, consequences.

<u>Operational-agency gained</u>

- The evaluation is perceived to be thorough as it touches on all the specified criteria.
- Attribution of successful impact is highlighted.

<u>Resulting instrumentalisation</u>

- The perceived need to include all the OECD-DAC criteria justifies expanded and unrealistic TORs which in turn encourage collusion between the commissioner and evaluator in selecting in what is observed and reported.
- The obligation to attribute impact to the project strongly favours the reporting of simplistic, positive outcomes.

15. Influence of the emergence of "Real time evaluations"

Cosgrove, Ramalingam and Beck (2009, p. 10) define an RTE as *"an evaluation in which the primary objective is to provide feedback in a participatory way in real time (i.e. during the evaluation fieldwork) to those executing and managing the humanitarian response."* The concept of real time evaluations has rapidly gained wide acceptance within the humanitarian community. By 2013, the ALNAP database of evaluations revealed that almost all evaluations of the response to the earthquake in Haiti were termed 'real time'.

<u>Principle-agency compromised</u>

- The description 'real time' has become devalued by being applied to evaluations which do not meet the definition.
- The definition of RTEs suggests that elements of the role of the evaluator as a dispassionate, external, independent observer of the project may be changed as a different role emerges.

<u>Operational-agency gained</u>

- An impression of immediacy and modernity without the substance.
- The ability of the evaluator to bring about real, immediate change to the implementation of a project and to see the results of this change.

<u>Resulting instrumentalisation</u>

- RTE demands the ability to field qualified, respected evaluators rapidly to projects operating in the immediate aftermath of a disaster. Using the term without this ability devalues the concept and renders the descriptor 'real time' meaningless.
- The direct involvement of the evaluator in the formulation of, and as the observer of, changes to project implementation suggests a changed role from external independent evaluator towards that of advisor or counsellor to the project.

<u>Comments</u>

It is not implied that the possibility of a changing role for evaluators involved in RTEs should be viewed in a negative light. Evaluators closely related to projects but outside line management may simply be acting in a different consultative role. Changes taking place should, however, be recognised and the development of appropriate new skills and techniques encouraged.

<u>Overall Caveat</u>

Although this analysis has demonstrated the potential consequences of each of the subjective influences identified in the research, it is not suggested that all evaluations are compromised in every manner identified above or compromised to an equal degree. The best evaluations will largely avoid compromise while the worst will succumb to a much greater degree. Considering the dynamics and pressures involved, it is unlikely, however, that any will be entirely independent.

Chapter 9 - Conclusions and Recommendations

In this chapter the researcher will bring together the overall conclusions reached by this research and make recommendations for further improvement of the process of evaluation and of learning for humanitarian action. The researcher recognises that throughout the analysis of the results a number of conclusions have been drawn related to individual headings under which the discussion was structured. These have been deliberately left in situ as, in the researcher's opinion, separating them and bringing them individually to this chapter would have fragmented the logical flow of the work beyond the point of easy comprehension.

Conclusions

The main conclusions of this work are presented by representing the evaluation of humanitarian action and institutional learning as components in a simple feedback loop, common in control mechanisms, illustrated in Diagram 9.1. Evaluation analyses the output of a system for the delivery of humanitarian action (typically either a humanitarian project or organisation). The evaluation process should feed information into an institutional learning function which then brings about changes in the inputs (planning and methodology) of the process of delivery of humanitarian action.

Diagram 9.1 Evaluation of humanitarian action and institutional learning as a simple feedback loop
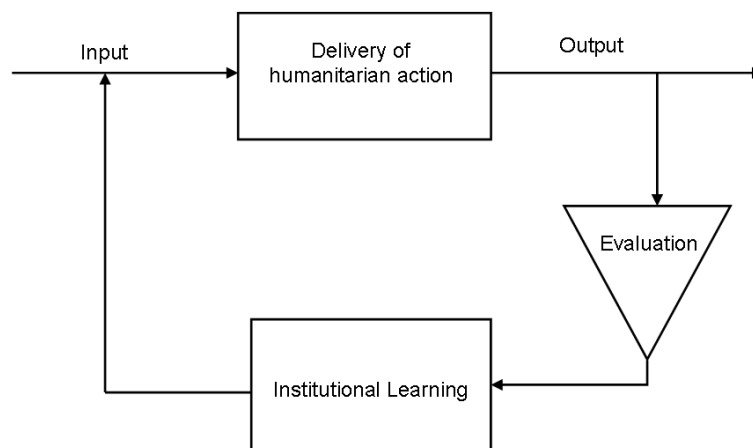
Source: Author

Real time evaluation (RTE) techniques, when properly applied, provide a mechanism for making this feedback loop functional at the organisation or project levels. However the highly fragmented nature of humanitarian delivery systems, whereby over a thousand independent organisations may respond to a single emergency, restricts the effectiveness of RTE to the organisation and project (micro) levels and renders any application to the wider system (macro) level impossible.

This research has demonstrated that, in situations other than RTE, the effectiveness of this feedback loop is compromised by serious defects that exist in both the evaluation and institutional learning components. The evaluation is compromised by subjective influences which usually reinforce each other to bias the evaluation towards positive inputs (i.e., inputs that reinforce the status quo) over negative inputs (i.e., inputs that point to problems the solution of which demands change).

Working within these constraints, evaluators find themselves trading their principle-agency of independent objectivity for their on-going operational-agency of contracts to undertake further evaluations. Apthorpe (2001) sums up the evaluator`s dilemma well *"Having a client is much better than not having a client if you want to make a difference to policy."* (p. 117)

An exception to this general principle exists when circumstances arise in which negative findings are considered politically essential in order to respond to a major failure or scandal. IC8 identified these circumstances referring to "*a scandal or some comprehensive failure that causes outrage or great disgust"… "It's got to strike an emotional chord, it can't just be rational or logical, a lot of people have to suffer or die or some iconic person or image or belief has to be shattered".*

Two deficiencies have been demonstrated regarding the institutional learning component

1. Many humanitarian organisations have weak systems for incorporating evaluation recommendations into their institutional learning.
2. The highly fragmented nature of the international humanitarian system results in minimal dissemination of learning between organisations.

The result of both components of the feedback loop being defective is that the deficiencies in one component act as a disincentive to improvement in the other.

- There is little incentive to make the efforts necessary to improve the quality of the critical nature of evaluations, and the strength of recommendations, as long as there is no effective institutional learning mechanism in operation.
- There is little incentive to put in place a robust mechanism for institutional learning as long as the quality of critical evaluation and recommendations necessary for its effective functioning remains unavailable.

This research has shown that weakness in the evaluation component can be attributed to a variety of political and subjective pressures, while the institutional learning mechanisms that do exist are largely as fragmented as the humanitarian system itself. Many humanitarian organisations independently produce handbooks, guidelines and tool kits which set standards and methodology for their own projects.

Recommendations

The recommendations made in this research to improve the functioning of the evaluation learning feedback loop for humanitarian action focus on:

1. The careful development of real time evaluation techniques and in particular the changing relationship between the evaluator and the project involved and the short timeframe which can be considered in these evaluations.
2. The establishment of a robust institutional learning system which could have a wide acceptance throughout the humanitarian system.

1. Real time evaluations (RTE)

The evolving phenomenon of RTE needs to be the subject of on-going study, particularly with regard to the implicit changes in the relationships between the evaluator, the evaluees and the commissioning organisation, as well as the consequent change in role and approach of the evaluator. The possibility of evaluators advising projects through RTE and conducting training of personnel working in projects that they could later evaluate may carry professional and ethical considerations new to the practice of evaluation of humanitarian action.

With its focus on short term results, impacts and changes to programmes, the use of the word evaluation may reasonably be challenged as the function more closely represents external monitoring. The researcher recognises, however, that the word evaluation has a

significance that transcends such a pedantic definition and that its use in the context of RTE is likely to continue.

Despite the criticisms that may be levelled, the RTE does however, have a potential to take evaluation, a tool originally designed for long term development projects, and transform it into a tool specially adapted to the more immediate needs of emergency humanitarian action.

It is important that precise criteria are defined for the term 'real time' and that it is given a protected status in the humanitarian community. The description used by Cosgrove, Ramalingam and Beck (2009, pp. 10-11) as evaluations carried out by small teams of *"one to four people"*, *"in the early stages of a response"*, designed to *"look at today to influence this week's programming"* is certainly a good starting point for the creation of a suitable definition. The use of the term for evaluations that do not meet the necessary criteria should be strongly criticised and databases of evaluations should refuse to include evaluation reports that use the term inappropriately.

The researcher perceives development of RTE techniques and skills as an ongoing process and it is interesting to note in this context that Cosgrove, Ramalingam and Beck (2009) consider their guide to be a pilot version.

2. Guidelines for humanitarian response

The researcher has already pointed out the fragmentation of the material produced to establish procedures and methodologies for humanitarian action, with many humanitarian organisations producing a wide variety of handbooks, manuals, guidelines and tool kits. Some of these publications have gained a wider acceptance and usage but many remain exclusively used by a single organisation. Such fragmentation makes the standardisation of any advances in practice as a result of lessons learned almost impossible. One document, the Sphere Standards, stands out as being compiled and revised by a representative group of humanitarian organisations and having a wider acceptance and use by the humanitarian community, including donors, governments and humanitarian organisations. The Sphere Standards form one model for this recommendation.

There is one discipline amongst humanitarians, medicine, which does have widely accepted and regularly updated procedures in the form of the WHO Guidelines and these form the

second model for this recommendation. In suggesting the WHO Guidelines as a model, the researcher recognises that medical guidelines have been the subject of considerable criticism (See Sniderman & Furberg 2009) and care would have to be taken that the proposed humanitarian guidelines avoided similar pitfalls.

In response to these criticisms a system for the evaluation of medical guidelines (AGREE Next Steps Consortium 2009) was established, with the acknowledgement that *"The potential benefits of guidelines are only as good as the quality of the guidelines themselves"* (p. 1) AGREE II has three purposes (p. 1) to:

1. *"Assess the quality of guidelines*
2. *Provide a methodological strategy for the development of guidelines; and*
3. *Inform what information and how information ought to be reported in guidelines."*

AGREE II provides a tool for the evaluation of medical guidelines scoring the guidelines on a seven point scale (p. 8) on each of 23 topics which are grouped under six headings (pp. 2-3)

1. *Scope and Purpose*
2. *Stakeholder Involvement*
3. *Rigour of Development*
4. *Clarity of Presentation*
5. *Applicability*
6. *Editorial Independence*

The researcher believes that AGREE II could be adapted as a similarly useful tool to provide a rating for humanitarian guidelines.

An important question to be addressed, if a more universally accepted set of guidelines for humanitarian action is to be produced and regularly revised is, within which institution or institutions should the guidelines be based? The first model for the proposed guidelines is the Sphere Standards which are based in a separate organisation, The Sphere Project, which *"is a voluntary initiative that brings a wide range* (The Sphere Project Board comprises 18 organisations) *of humanitarian agencies together around a common aim - to improve the quality of humanitarian assistance and the accountability of humanitarian actors to their constituents, donors and affected populations."* (The Sphere Project 2013) In May 2013 the Sphere Project together with People in Aid and the Humanitarian Accountability Partnership, were working towards a greater coherence in standards, having *"launched a process* (known as the Joint Standards Initiative JSI) *to seek greater coherence for users of standards, in*

order to ultimately improve humanitarian action to people affected by disasters." (Joint Standards Initiative 2013a) By December 2013, however the JSI was brought to and end (Joint Standards Initiative 2013b) following the disengagement of the Sphere Project which stated that "*the harmonisation process that combining operational and organisational standards is not easy and might not be desirable and that looking at complementarity of different approaches is a better way forward*". (People in Aid 2013b) quoting a statement made at a meeting of the JSI.

The researcher believes, however, that a similar consortium to the JSI could be an important structure in the development of broadly based humanitarian guidelines.

 Looking towards the second model, the WHO Guidelines, the researcher notes the strength of a base in the United Nations system. The United Nations Office for the Coordination of Humanitarian Affairs (OCHA) is mandated by the United Nations General Assembly (1991) to coordinate the humanitarian actions of the United Nations and, through the Emergency Relief Coordinator, holds the chairmanship of the Inter Agency Standing Committee (IASC). UN operational organisations are all members of the IASC, the ICRC and IFRC are observers and "*relevant non-governmental organisations are invited to attend on an ad hoc basis*". (Article 39) OCHA, through the IASC has influence with the UN organisations and recognition by the UN member states. OCHA has already produced some guidelines for humanitarian operations.

The recommendation is for JSI and OCHA to work together to form, fund and manage groups of qualified and experienced individuals who would produce and regularly update guidelines for aspects of humanitarian action that would be endorsed by both the JSI and the IASC. Such guidelines would take into account and incorporate, where appropriate, existing guidelines, successful innovations and best practice developed from the lessons learned from evaluations. The researcher does not in any way underestimate the difficulty in realising this cooperation. It would, however, be a powerful force in producing guidelines with a high degree of acceptance amongst humanitarian organisations.

An obvious difficulty is that many organisations have already developed their own guidelines specifically tailored to their particular mandates and modes of operation and these would certainly have to be considered in the production of the proposed guidelines. Even so, resistance can be expected from organisations to modify their guidelines or adopt new procedures to comply with more universally accepted guidelines.

The researcher can see two very distinct options for the way in which a set of humanitarian guidelines may be structured.

- The researcher's preference would be for individual guidelines to be developed, each for use in a certain type of emergency; floods, earthquakes, forced displacement, etc. setting priorities for response, identifying risks of secondary crises and identifying response capacities that either have to be available within the affected state or brought in from outside. Individual states could link these guidelines to their own disaster response plans and humanitarian organisations to their standby capacity and response planning.
- An alternative would use the existing structure of coordination through sector specific clusters such as water, sanitation and hygiene (WASH), shelter, and health. The use of such a structure would have the disadvantage of maintaining the current separation of the components necessary to mount an effective humanitarian action and is likely to result in guidelines which would be at best an extension of the Sphere Standards and at the worst a mere duplication.

The researcher would further recommend that the organisations responsible for the development of the proposed guidelines are also active in their dissemination and promotion as well as the production of training materials and the training of humanitarian workers in the application of the guidelines.

Annex 1 Principle-agency operational-agency analysis of subjective influences on the evaluation of humanitarian action

	Subjective influence	Principle-agency compromised	Operational-agency gained	Resulting instrumentalisation
1	Perception of subjectivity as universal in evaluation of humanitarian action	Objectivity	Encourages veiled criticism and vague recommendations.	Strong criticism and far reaching recommendations can be countered by describing them as subjective.
2	Sense of urgency surrounding humanitarian action	**a.** humanitarian actors compromise preparation of projects **b.** commissioner compromises selection of evaluators and preparation of the evaluation **c.** evaluators agree to short deadlines	**a.** Speed of action necessary for immediate presence and fundraising. **b.** Quickly able to field evaluation teams **c.** Justify selectivity and shallow evaluations	Commissioner maintains control of the evaluation process resulting in little risk of severe criticism for poorly planned and implemented projects.
3	The attitude, personality and interests of those undertaking the evaluation	Professional detachment, i.e. separation of individual bias from the evaluation	Dissemination and wider application of the evaluator`s own views and interests	Commissioners able to select evaluators whose bias is supportive to the work of the organisation. Such evaluators are ensured repeat contracts.
4	Inter-organisational and interpersonal relationships	Independence of the evaluator or evaluation team	Minimises potential for conflict and negative impressions for future contracts	In attempting to influence the evaluator, each party is trying to instrumentalise the evaluation to further their own objectives
5	Relationship between the organisation commissioning the evaluation and the evaluation team	Independence of the evaluation process. Well established consultant`s report being more easily able to resist this compromise.	Securing present and future contracts. Securing full payment for present contract.	Commissioners exploit the unequal relationship with evaluators to obtain a favourable evaluation which they can present as both external and independent.

	Subjective influence	Principle-agency compromised	Operational-agency gained	Resulting instrumentalisation
5a	Commissioner imposed terms of reference	Realistic terms of reference Evaluator input or negotiation	Avoidance of conflicts within the commissioning organisation Evaluator secures the contract	Collusion between the commissioner and the evaluator in selectivity in implementation of the terms of reference to give the most favourable report.
6	Relationship between the commissioner of the evaluation and the project being evaluated	Full and accurate disclosure by project staff to the evaluators	Continued resourcing and support for the project from the commissioner of the evaluation	Project controls the evaluation, marginalising the independence of the evaluator to the maximum extent
7	Relationship between the evaluating team and the project being evaluated	Independence of the evaluators (Due to heavy reliance on the subject of the evaluation for physical support).	Easy relationships with the project staff enable evaluation team to complete the evaluation and report in the limited time available to them. Positive feedback to the commissioner from the project staff.	Subjects of the evaluation gain control of the evaluation process.
8	Relationship between the evaluating team and other stakeholders in the project being evaluated	Thorough and meaningful beneficiary input to the evaluation process. Balance of diverse opinions received and analysis of the interests motivating the responses	Shorter simpler report, more easily understood by the commissioner and more acceptable to the project being evaluated	Input from the beneficiaries and other local stakeholders is marginalised as contradictory and more critical opinions are selectively discarded.

	Subjective influence	Principle-agency compromised	Operational-agency gained	Resulting instrumentalisation
9	Relationships within the evaluating team	Ability of the team to produce a coherent coordinated evaluation with agreed conclusions	Operational-agency is gained by the commissioner in the form of shorter, hence cheaper, consultancy contracts Control of the evaluation report by colluding with the team leader to selectively include and exclude diverse opinions.	Commissioner's influence over the evaluation report is increased
10	Interagency relationships in the field	Inclusivity of information and opinions presented to the evaluation team	Ease of relations with the project being evaluated by easily ignoring criticism of rivals.	Project able to fend off criticism as unjustified interagency prejudice
11	Learning from evaluations and the learning culture of humanitarian organisations	The value of evaluation as an important component in a "feedback loop" designed to improve delivery of humanitarian action.	Enables the continued presentation to donors of "well tried" methodology without the risk that new methodologies may not attract funding.	Repetition of projects based around donor norms rather than appropriately tailored to beneficiary needs
11a	Selectivity in dissemination and putting report in the public domain	Transparency and accountability	Artificially high level of credibility	Organisation able to present the most favourable public image
11b	Evaluation as a box ticking exercise	Value of evaluation as a learning tool	Requirement to evaluate met	Evaluation becomes ritualistic.
11c	Lack of ownership of the evaluation	Use of the evaluation as a learning tool	Evaluation report can be largely ignored	Ritualization of the evaluation becomes complete
11d	Readability of evaluation reports	Depth of criticism, analysis and recommendations	Easier wider dissemination and a larger readership	Brevity and nuanced criticism

	Subjective influence	Principle-agency compromised	Operational-agency gained	Resulting instrumentalisation
12	Dual objectives of accountability and lessons learned	Focus on evaluation as a learning tool	Satisfaction of a universal demand for accountability	Defines success in a framework of compliance rather than one of learning.
13	Culture of evaluators and the western origins of evaluation	Input of beneficiary communities and other local stakeholders who evaluate a project from a different cultural perspective	Evaluation criteria and methodology easily understood by western donors and western based humanitarian organisations.	Success and failure are defined in the context of the commissioner rather than that of the beneficiary
14	The use of the OECD-DAC criteria	- Realistic TORs - Depth of evaluation against moat relevant criteria - Difficulty in assigning impact and causality to short term projects.	- Allows the inputs of multiple departments within the commissioning organisation to be included without the necessity for prioritisation and difficult internal political negotiation. - Forces short term impact to be attributed .	- Encourages collusion between the evaluators and the commissioner on selectivity of objectives and the relative emphasis to be put on each. - Favours reporting of short term positive impacts.
15	Real time evaluations	The external, independent nature of the evaluator	Immediate feedback and change of project implementation.	Feedback and change is focussed at an individual organisation or project with little possibility for wider dissemination.

Annex 2 Consent form

<u>Liverpool School of Tropical Medicine</u>

Consent form for participants in academic research

Please read this form carefully and sign the declaration at the end if you agree to participate in this research project. This form will be collected from you by the researcher at the beginning of the interview or by prior e-mail in the case of a telephone interview.

If the interview is to be conducted by telephone please scan the signed form and e-mail it to A.Land@liverpool.ac.uk

This research is being undertaken by Anthony (Tony) Land who is writing a thesis titled "Enhancing responsibility and accountability in humanitarian action: understanding the subjective factors that influence evaluation of humanitarian actions and the implementation of the recommendations made" for the degree of Doctor of Philosophy at Liverpool University, England. Mr. Land has 30 years experience of working in humanitarian organisations, NGOs and United Nations. Over this period he has worked in field and headquarters situations including the direct planning and implementation of humanitarian projects as well as external relations where he has worked both with donors and the media. He now teaches on courses run by, or in association with, Fordham, Copenhagen, Manchester and Liverpool Universities, participates in CIMIC training with military forces and undertakes various consultancies.

1. This research is aimed at obtaining a better understanding of the dynamics at work between the findings of researchers, the recommendations of evaluators and the introduction of changes to the way in which humanitarian assistance is delivered. The research is particularly interested in identifying tools for facilitation of, areas of resistance and obstacles to, the implementation of research findings and evaluation recommendations.

2. You have been selected as a potential participant in this research as an experienced practitioner of humanitarian action, as a representative of a member of the donor community or as an established external observer of humanitarian actions.

3. If you agree to participate you will be asked to:
   - Complete a self-definition questionnaire in which you will define, within a set of categories, your present role and previous experience.
   - Participate in an interview with the researcher that is expected to last about one hour

   The interview will be arranged at a time and place convenient to you.

4. Your permission will be sought to make a digital audio recording of the interview so as to assist the researcher to make an accurate written record of the proceedings. The digital recording will be transferred to a password protected computer within 24 hours of the interview and at the same time stored on a remote server in an account to which only the researcher has access. The interview will be transcribed into a word processor file and all copies of the digital audio recording erased. Parts of the interview may be quoted verbatim in the thesis being prepared from this research, however, unless specific agreement has been obtained to attribute any quotation, all quotations will be anonymous or attributed to a generic category e.g. senior UN official.

5. If you wish any part of the interview, if used in the final thesis, to be attributed to you personally then this may be agreed during the interview in which case the exact wording to be quoted will be transcribed, returned to you by e-mail and will only be personally attributed if an e-mail reply authorising this is received by the researcher.

6. At the completion of the research a summary of the results and conclusions will be provided to you.

7. Your participation in this research is completely voluntary and you may cease to participate at any time and withdraw information you have provided up until the point when the thesis is submitted to the University for academic evaluation

8. The researcher may be contacted by e-mail at A.Land@liverpool.ac.uk in case you have any further queries or difficulties related to this study.

I have read and understood the information given above and I agree to participate in this research project in the manner described.

Name (Block capitals) ……………………………………………………….

Date (dd/mmm/yyyy) ………/………………/………………..

Place (Town or city) …………………………………………

Signature ………………………………………………………

FOR TELEPHONE INERVIEWS ONLY

If the form is being returned by e-mail in preparation for a telephone interview please enter the telephone number on which you would like to be contacted and any 3 digit number below. You will be asked to quote this number at the beginning of the interview in order to verify that the signatory of the form is the person participating in the interview.

Telephone number on which I would like to be contacted:

Consent verification digits   _   _   _

Participant Number …………………….

Liverpool School of Tropical Medicine

This questionnaire is designed for use by participants in conjunction with the research currently being undertaken by Anthony Land titled "Enhancing responsibility and accountability in humanitarian action: understanding the subjective factors that influence evaluation of humanitarian actions and the implementation of the recommendations made."

Completing this questionnaire will assist in the correct attribution of quotations to appropriate anonymous categories and in the final analysis of the information collected. **Please complete all three pages – Only 6 Questions**

Part 1

1.1 Employer

Please tick the appropriate boxes that describe your **current employer.**

Government Department or Agency	
Inter-governmental Agency	
United Nations	
Red Cross / Red Crescent Movement	
Non-Governmental Organisation	
Media Company or Organisation	
Consultancy Firm	
Academic Body or Organisation	
Political Body or Organisation	
Self Employed	
Military	
Other (Specify Below)	

How long have you been working with your current employer?

Underline the relevant period

< 3 years   3-5 years   5-10 years >10 years

1.2. Type of work being undertaken

Tick the **box** that most closely represents you **current job**

Global Policy Development & Management	
Regional Policy Development & Management	
National Policy Development & Management	
Field Team Management	
Administration and Finance	
Technical Specialist	
Evaluation / Inspection	
Fund Raising	
Media Relations	
Training	
Research	
Advocacy & Lobbying	
Donor	
Politics	
Journalism / Author	
Civil Military Cooperation	
Other (Specify below)	

How long have you been working in your current job

<u>Underline</u> the relevant period

< 3 years   3-5 years   5-10 years >10 years

…………………………………………

1.3. Level in your organisation

Tick the box that most closely describes your **current level**

Head of Organisation or Director	
Senior Official (equivalent UN P4 – P5)	
Junior Official (equivalent UN P2 –P3)	
Temporary Staff Ungraded / Intern	
Consultant	
Self employed	
Other (Specify Below)	

How long have you been working at your current level?

<u>Underline</u> the relevant period

< 3 years   3-5 years   5-10 years >10 years

………………………………………………….

301

Part 2

Please tick all the boxes that describe your **previous experience** in, or related to, humanitarian action

2.1 Employers

Please indicate the length of time spent with **previous** types of employers

Type of employer	Number of years with this type of employer
Government Department or Agency	
Inter-governmental Agency	
United Nations	
Red Cross / Red Crescent Movement	
Non-Governmental Organisation	
Media Company or Organisation	
Consultancy Firm	
Academic Body or Organisation	
Political Body or Organisation	
Self Employed	
Military	
Other (Specify Below)	

…………………………………………………………………

2.2. Types of work undertaken

Indicate the length of time spent in **previous jobs**

Type of job	Number of years in this type of job
Global Policy Development & Management	
Regional Policy Development & Management	
National Policy Development & Management	
Field Team Management	
Administration and Finance	
Technical Specialist	
Evaluation / Inspection	
Fund Raising	
Media Relations	
Training	
Research	
Advocacy & Lobbying	
Donor	
Politics	
Journalism / Author	
Civil Military Cooperation	
Other (Specify below)	

………………………………………………………

2.3 Length of humanitarian experience

How long have you been working in / involved with humanitarian action

Underline the relevant period

> < 3 years   3-5 years   5-10 years >10 years

Annex 4 Semi structured interview outline.

1	You were asked to think of situations where you have been involved in the evaluation process where subjective factors have had an influence. Do you have some examples you would like to talk about?	
2	How many?	If participant is forthcoming then the interview proceeds in a relatively unstructured manner.
3	What experiences have you had in participating in the evaluation process?	Commissioner Planner Recruitment of evaluators Trainer of evaluators Writer of TOR Evaluator Working on a project being evaluated Managing a project being evaluated Providing information on a project being evaluated Editing or approving evaluation reports Using information from evaluations
4	How involved in the evaluation process did you feel?	Ownership Relationships Perceptions Reactions Transparent Spin Pressures (what and from where)
5	What is the influence of external factors on evaluation?	Donors Auditors/Inspectors Senior Management Designers/managers of projects Beneficiaries
6	How are evaluators perceived ?	Internal External Trusted Encourage honesty Independent Influential in org. change Use of reports
7	Have you experienced others roles that look into humanitarian projects? How is each perceived?	Academics Journalists Inspectors Investigators (Misconduct

8	Dissemination and implementation of evaluation findings/recommendations	Experience How much used Accessibility of evaluations Accessibility in evaluations Report structure facilitating? Improvements
9	How did organisational systems facilitate the use of evaluations?	
10	What factors influence use of evaluations?	Awareness Accessibility Time Attitude Experience of user
11	Why are some of the most repeated criticisms of humanitarian action not addressed?	Coordination Speed of response Cost effectiveness Out of context Overhead/administrative costs
12	Why do humanitarian organisations undertake evaluations?	Drivers Reasons Utility
13	Do you believe that humanitarian organisations are learning organisations?	
14	Where do you believe that the greatest improvements could be made in the process of evaluation and learning?	
15	What else would you like to tell me that would be helpful to me in my research?	

Bibliography

Active Learning Network for Accountability and Performance (2013a) *About ALNAP* [Online], Available from: http://www.alnap.org/about.aspx (Accessed: 9 May 2013).

Active Learning Network for Accountability and Performance (2013b) *Annual Report 2012/13*, London

Adelman, H. & Suhrke, A. (1996) *Study 2 Early Warning and Conflict Management*, Adelman, H. & Suhrke, A., vol. 2, Steering Committee of the Joint Evaluation of Emergency Assistance to Rwanda [Online], Available from: http://www.grandslacs.net/doc/0742.pdf (Accessed: 12 January 2012).

AFP (2005) EU alert to corruption risk to Asia aid, *AFP*, 2 Oct 2012.

AGREE Next Steps Consortium (2009) *The AGREE II Instrument*, AGREE Next Steps Consortium, [Online], Available from: http://www.agreetrust.org (Accessed: 20 Feb 2013).

Anderson, M. (1999) *Do no Harm*, Lynne Rienner Publishers.

Annan, K.A. (2000) *We the peoples: The role of the United Nations in the 21st century*, United Nations, New York

Apthorpe, R. (2001) Mission Possible: Six Years of WFP Emergency Food Aid in West Africa, in A. Wood, R. Apthorpe & J. Borton (eds), *Evaluating International Humanitarian Action: Reflections from Practitioners*, Zed Books in association with ALNAP, London & New York, pp. 102-121.

Apthorpe, R. (2011) With Alice in Aidland: A seriously Satirical Allegory, in D. Mosse (ed.), *Adventures in Aidland: The Anthropology of Professionals in International Development*, vol. 6, Berghahn Books, New York, Oxford, pp. 199-219.

Argyris, C. & Schön, D.A. (1978) *Organizational learning : a theory of action perspective*, Addison-Wesley, Reading, Mass. ; London.

Barnett, M. (2002) *Eyewitness to Genocide: The united Nations and Rwanda*, Cornell University Press, Ithaca & London.

Barnett, M. (2005) Humanitarianism Transformed, *Perspectives on Politics*, vol., no. 4, pp. 723-740 [Online]. Available from: http://www.jstor.org/stable/3688176 (Accessed: 10 August 2013).

Barnett, M. (2011) *Empire of Humanity*, Cornell University Press, Ithaca & London.

Barnett, M. & Weiss, T.G. (2011) *Humanitarianism Contested: Where Angels Fear to Tread*, vol. 51, Routledge, London & New York.

Bartsch, D. & Belgacem, N. (2004) *Real time evaluation of UNHCR's response to the emergency in Chad*, UNHCR - Evaluation and Policy Analysis Unit, Geneva

Beck, T. (2006) *Evaluating humanitarian action using the OECD-DAC criteria*, Beck, T., ODI [Online], Available from: http://www.alnap.org/pool/files/eha_2006.pdf (Accessed: 31 January 2013).

Bennett, J., Bertrand, W., Harkin, C., Samarasinghe, S. & Wickramatillake, H. (2006) *Coordination of international humanitarian assistance in tsunami-affected countries*, Tsunani Evaluation Coalition, London

Bessler, M. & Seki, K. (2006) 'Civil military relations in armed conflicts', *Liaison- A Journal of Civil-Military Humanitarian Relief Collaborations*, vol. 3, no. 3.

Binder, A. & Grünewald, F. (2010) *IASC Cluster Approach Evaluation, 2nd Phase Country Study*, IASC,

Borton, J. (1994) The Upsurge on Interest in the 'Relief-Development Continuum': What does it mean?, *Humanitarian Exchange Magazine*, vol., no. 2, [Online]. Available from: http://www.odihpn.org/humanitarian-exchange-magazine/issue-02/the-upsurge-of-interest-in-the-relief-development-continuum-what-does-it-mean (Accessed: 8 May 2013).

Borton, J. (2001) Doing Study 3 of the Joint Evaluation of Emergency Assistance to Rwanda: The Team Leader's Perspective, in A. Wood, R. Apthorpe & J. Borton (eds), *Evaluating International humanitarian Action: Reflections from Practitioners*, Zed Books in association with ALNAP, London & New York, pp. 74-101.

Borton, J. (2014) *20 years after the Rwandan genocide: The unfinished accountability revolution* [Online], Sphere Project, Available from: http://www.sphereproject.org/news/20-years-after-the-rwandan-genocide-the-unfinished-accountability-revolution-john-borton/ (Accessed: 13 May 2014).

Borton, J., Brusset, E. & Hallam, A. (1996) *Study 3: Humanitarian Aid and Effects*, Borton, J., Brusset, E. & Hallam, A., Steering Committee of the Joint Evaluation of Emergency Assistance to Rwanda [Online], Available from: http://www.grandslacs.net/doc/0743.pdf (Accessed: 12 January 2012).

Borton, J. & Eriksson, J. (2004) *Lessons from Rwanda - Lessons for Today*, Danish Ministry of Foreign Affairs, Copenhagen.

Bostridge, M. (2009) *Florence Nightingale : the woman and her legend*, Penguin, London.

Brahimi, L. (2000) *Report of the Panel on United Nations Peace Operations*, United Nations, New York

Brauman, R. (1996) *Humanitaire – le dilemme*, Les Editions Textuel, Paris.

Brauman, R. (1997) Introduction, in M.S.F. (MSF) (ed.), *World in Crisis: The Politics of Survival at the End of the Twentieth Century*

Routledge, London, pp. xix-xxvi.

Brauman, R. (1998) Refugee Camps, Population Transfers and NGOs, in J. Moore (ed.), *Hard Choices: Moral Dilemmas in Humanitarian Intervention*, Rowman and Littlefield, Lanham, Maryland, USA., pp. 177-194.

British Broadcasting Corporation (2010) *Q&A: Sudan's Darfur conflict* [Online], Available from: http://news.bbc.co.uk/1/hi/world/africa/3496731.stm (Accessed: 11 October 2013).

Britten, N. (1999) Qualitative interviews in health care research, in C. Pope & N. Mays (eds), *Qualitative Research in Health Care*, BJM Books, London, pp. 11-19.

CARMA (2006) *Western media coverage of Humanitarian Disasters*, CARMA Europe,

Centlivres, P. & Centlivres-Demont, M. (1988) 'Afghan Refugee in Pakistan: An Ambiguous Identity, The', *J. Refugee Stud.*, vol. 1, p. 141.

Christoplos, I. (2006) *Links between Relief, Rehabilitation and Development in the Tsunami Response*, Tsunami Evaluation Coalition, London

Churchill, W. (1946) *Iron Curtain Speech* [Online], Great Neck Pub., Available from: http://courses.kvasaheim.com/common/docs/ironcurtain.pdf (Accessed: 22 Aug 2013).

Cockayne, J. (2006) *Commercial Security in Humanitarian and Post-Conflict Settings: An Exploratory Study*, International Peace Academy, New York.

Collinson, S. & Elhawary, S. (2012) *Humanitarian space: a review of trends and issues*, Humanitarian Policy Group, London

Corbin, J. & Strauss, A. (2008) *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 3rd edn, Sage Publications, Los Angeles, London, New Delhi, Singaport.

Cosgrave, J. (2007) *Synthesis Report: Expanded Summary*, Tsunami Evaluation Coalition,

Cosgrove, J., Ramalingam, B. & Beck, T. (2009) *Real-time evaluations of humanitarian action: An ALNAP Guide - Pilot Version*, Overseas Development Institute, London.

Cutts, M. (1999) *The humanitarian operation in Bosnia, 1992-95: dilemmas of negotiating humanitarian access*, Discussion Paper, *New Issues in Refugee Research*, no 8, Policy Research Unit, UNHCR, Geneva.

Dallaire, R.A. & Beardsley, B. (2004) *Shake hands with the devil: The failure of humanity in Rwanda*, Arrow.

DARA (2010) *The Humnaitarian Response Index 2010: The problems of politicisation*, DARA, Madrid

DARA (2011) *The Humanitarian Response Index 2011: Addressing the gender challenge*, DARA, Madrid

de Vattel, E. (1757) *The Law of Nations*, in B. Kapossy & R. Whatmore (eds), *The Law of Nations: A republication of the 1797 translation of Vattel's work, along with new English translations of 3 early essays.*, Liberty Fund, Indianapolis.

de Ville de Goyet, C. & Morinière, L.C. (2006) *The role of needs assessment in the tsunami response*, Tsunami Evaluation Coalition, London

Denzin, N. (1970) *The Research Act in Sociology*, Butterworth, London.

Department for International Development (2011) *Dfid's Approach to Value for Money*, Department for International Development, DFID [Online], Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/49551/DFID-approach-value-money.pdf (Accessed: 28 April 2013).

Development Assistance Committee (1991) *Principles for Evaluation of Development Assistance*, OECD, Paris.

Development Assistance Committee (1999) *Guidance for Evaluating Humanitarian Assistance in Complex Emergencies*, OECD.

Development Assistance Committee (2010) *Quality Standards for Development Evaluation*, Development Assistance Committee, OECD [Online], Available from: www.oecd.org/dataoecd/55/0/44798177.pdf (Accessed: 16 October 2013).

Drifmeyer, J. & Llewellyn, C. (2004) 'Toward more effective humanitarian assistance', *Military Medicine*, vol. 169, no. 3, pp. 161-168.

Dufour, C., Geoffroy, V., Maury, H. & Grünewald, F. (2004) 'Rights, standards and quality in a complex humanitarian space: is SPHERE the right tool?', *Disasters*, vol. 28, no. 2, pp. 124-141.

Dunant, J.H. (1862) *Un Souvenir de Solferino. Deuxième Édition. [With a map.]*, Fick, Genève.

Encyclopædia Britannica (2013) Battle of Tsushima,, *Encyclopædia Britannica Online*, vol., [Online]. Available from: http://www.britannica.com/EBchecked/topic/607927/Battle-of-Tsushima (Accessed: 16 Aug 2013).

Enhancing Learning and Research for Humanitarian Assistance (2010) Forum Report, *International Working Forum on Humanitarian Professionalisation*, London.

Eriksson, J. (1996) *Synthesis Report* Eriksson, J., Steering Committee of the Joint Evaluation of Emergency Assistance to Rwanda [Online], Available from: http://www.grandslacs.net/doc/0745.pdf (Accessed).

Evalsed (2009) *A Short History of Evaluation* [Online], European Commission, Available from: http://ec.europa.eu/regional_policy/sources/docgener/evaluation/evalsed/guide/development/history_purpose/history_en.htm (Accessed: 12 Sep 2012).

Evans, G. & Sahnoun, M. (2001) *The responsibility to protect*, The International Commission on Intervention and State Sovereignty, Ottawa

Feinstein Center (2004) *Ambiguity and Change: humanitarian NGOs prepare for the future*, Tufts University, Medford, Mass

Fetterman, D. (2001) *Foundations of Empowerment Evaluation*, Sage Publications, Thousand Oaks, London & New Delhi.

Financial Times (2013) *Definitioni of the principal agent problem* [Online], Financial Times, Available from: http://lexicon.ft.com/term?term=principal/agent-problem (Accessed: 30 November 2013).

Flint, M. & Goyder, H. (2006) *Funding the tsunami response*, Tsunami Evaluation Coalition, London

Forsythe, D.P. & Rieffer-Flanagan, B.A.J. (2007) *The International Committee of the Red Cross : a neutral humanitarian actor*, Routledge, London.

Fox, F. (2002) 'New Humanitarianism: Does it provide a moral banner for the 21st century?', *Disasters*, vol. 25, no. 4, pp. 275-289.

Fuller, M. (2013) *Benefits Britain 1949 Episode 1* [TV]. Channel 4.

Geneva Conventions (2013), *Encyclopædia Britannica Online*, vol., [Online]. Available from: http://www.britannica.com/EBchecked/topic/229047/Geneva-Conventions (Accessed: 16 Aug 2013).

Giddens, A. (1984) *The Consitiution of Society*, Polity Press, Cambridge.

Gillham, B. (2005) *Research Interviewing: The range of techniques*, Gillham, B., Open University Press - McGraw Hill Education [Online], Available from: http://site.ebrary.com.ezproxy.liv.ac.uk/lib/liverpool/docDetail.action?docID=10161349 (Accessed: 11 Jan 1012).

Glenny, M. (1996) 'The fall of Yugoslavia: The third Balkan war'.

Good Humanitarian Donorship (2003) Principles and Good Practice of Humanitarian Donorship, vol., [Online]. Available from: http://www.goodhumanitariandonorship.org/gns/principles-good-practice-ghd/overview.aspx (Accessed: 4 October 2011).

Grünewald, F., Binder, A. & Georges, Y. (2010) *Inter-agency real-time evaluation in Haiti: 3 months after the earthquake*, Urgence Réhabilitaion Développement and Global Public Policy Institute,

Grünewald, F., Pirotte, C. & de Geoffroy, M. (2001) An Experimental and Inclusive Approach to Evaluation as a Lesson-learning Tool: Groupe URD's Work on the Post Hurricane Mitch Emergency, in A. Wood, R. Apthorpe & J. Borton (eds), *Evaluating International Humanitarian Action: Reflections from Practitioners*, Zed Books in association with ALNAP, London & New York, pp. 153-170.

Hallam, A. (1998) *Evaluating humanitarina assistance programmes in complex emergencies*, Hallam, A., vol. 7, Overseas Development Institute [Online], Available from: http://www.livestock-emergency.net/userfiles/file/common-standards/Hallam-1998.pdf (Accessed: 13 Sep 2012).

Hallam, A. (2011) *Harnessing the Power of Evaluation in Humanitarian Action*, Hallam, A., ALNAP [Online], Available from: http://www.alnap.org/pool/files/evaluation-alnap-working-paper.pdf (Accessed).

Hammersley, M. & Atkinson, P. (1983) *Ehtnography Principles in Practice*, Tavistock, London & New York.

Harrell-Bond, B. (1986) *Imposing Aid*, Oxford University Press, Oxford.

Hogan, R.L. (2007) The Historical Development of Program Evaluation: Esploring the past and present, *Online Journal of Workforce Education and Development*, vol., no. 4, pp. 1 - 14 [Online]. Available from: http://wed.siu.edu/Journal/VolIInum4/Article_4.pdf (Accessed: 12 Sep 2012).

Holbrooke, R. (2011) *To end a war*, Random House Digital, Inc.

Hollis, P.S. (2004) Operation Iraq Freedom Hallmarks: Integrated Joint and Coalition Operations with Adaptable Commanders and Agile Planning and Execution, *Field Artillery Journal*, March-June 2004, pp. 4-9

Hope, C. (2013) 30 Charity chiefs paid more than £100,000, *Daily Telegraph*, 6 August 2013.

Horst, C. (2005) *Transnational Nomads*, Berghahn Books.

Humanitarian Accountability Partnership (2010) *Standard in Accountability and Quality Management*, Humanitarian Accountability Project.

Humanitarian Accountability Partnership (2013a) *About us* [Online], Available from: http://www.hapinternational.org/about.aspx (Accessed: 3 Feb 2013).

Humanitarian Accountability Partnership (2013b) *Current Members and HAP Certified Organisations* [Online], Available from: http://www.hapinternational.org/members.aspx and http://www.hapinternational.org/projects/certification/certified-agency-register.aspx

 (Accessed: 30 August 2013).

Humanitarian Accountability Partnership (2013c) *HAP 2013 Annual Report*,

Humanitarian Accountability Partnership (2013d) *History* [Online], Available from: http://www.hapinternational.org/about/history.aspx (Accessed: 22 April 2013).

Interagency Standing Committee (2010) *Response to the Humanitarian Crisis in Haiti Following the 12 January 2010 Earthquake: Achievements, Challenges and Lessons To Be Learned*, IASC, New York, Geneva

International Committee of the Red Cross (2012) *Annual Report*, Geneva

International Council of Voluntary Agencies (2012) *Annual Report 2012*, Geneva

International Federation of Red Cross and Red Crescent Societies (1994) *Code of conduct*.

International Federation of Red Cross and Red Crescent Societies (2007) *The Global Aliance for Disaster Risk Reduction*, IFRC, Geneva

International Federation of Red Cross and Red Crescent Societies (2011a) *IFRC Framework for Evaluation*, International Federation of Red Cross and Red Crescent Societies, Planning and Evaluation Department (PED) IFRC Secretariat [Online], Available from: http://www.ifrc.org/Global/Publications/monitoring/IFRC-Framework-for-Evaluation.pdf (Accessed).

International Federation of Red Cross and Red Crescent Societies (2011b) *Introduction to the Guidelines for the domestic facilitation and regulation of international disaster relief and initial recovery assistance*, IFRC, Geneva.

Joint Standards Initiative (2013a) *About the initiative* [Online], Available from: http://www.jointstandards.org/about (Accessed: 1 May 2013).

Joint Standards Initiative (2013b) *Feedback on the first draft of the Core Humanitarian Standard* [Online], Available from: http://www.jointstandards.org/news/CHS (Accessed: 4 March 2014).

Karajkov, R. (2007) The Power of NGOs: They're Big, But How Big, vol., [Online]. Available from: http://www.worldpress.org/Americas/2864.cfm (Accessed: 15 August 2013).

Kennicot, P. (2010) Media offers full, blunt and hideous picture of trauma caused by Haiti earthquake, *Washington Post*, 16 Jan 2010.

Kirk, J. & Miller, M. (1986) *Relaibility and Validity in Qualitative Research*, vol. 1, Sage, London.

Kirkby, J., Howorth, P., O'Keefe, P. & Collins, A. (2001) 'A survey of evaluation experience in complex emergencies', *The international journal of human rights*, vol. 5, no. 2, pp. 114-129.

Kumar, K., Tardif-Douglin, D., Maynard, K., Manikas, P., Sheckler, A. & Knapp, C. (1996) *Study 4 Rebuilding Post-War Rwanda*, Kumar, K., Tardif-Douglin, D., Maynard, K., Manikas, P., Sheckler, A. & Knapp, C., Steering Committee of the Joint Evaluation of

Emergency Assistance to Rwanda [Online], Available from: http://www.grandslacs.net/doc/0744.pdf (Accessed: 12 January 2012).

Labonte, M. (2011) *Lecture on "Responsibility to Protect"*, [unpublished], Fordham University.

Lea, D.A.M. (2001) A Self-evaluation of my Experience Reviewing Australia's Official Assistance in Response to the 1997-98 Papua new Guinea Drought, in A. Wood, R. Apthorpe & J. Borton (eds), *Evaluating International Humnaitarian Action: Reflections of Practitioners*, Zed books in association with ALNAP, London & new York, pp. 133-152.

Lee, R.M. & Fielding, N.G. (eds) (1995) *'Users' experiences of qualitative data analysis software*, Sage, London.

Lindahl, C. (2001) Evaluating Sida's Complex Emergency Assistance in Cambodia: Conflicting Perceptions, in A. Wood, R. Apthorpe & J. Borton (eds), *Evaluating International Humanitarian Action: Reflections from Practitioners*, Zed Books in association with ALNAP, London & New York, pp. 59-73.

Livingston, S. & Eachus, T. (1995) 'Humanitarian crises and US foreign policy: Somalia and the CNN effect reconsidered', *Political Communication*, vol. 12, no. 4, pp. 413-429.

MacDonald, W. (2001) Forward, in A. Wood, R. Apthorpe & J. Borton (eds), *Evaluating International Humanitarian Action: Reflections from Practitioners*, Zed Books, London and New York.

Macionis, J.J. & Plummer, K. (2005) *Sociology : a global introduction*, 3rd ed. edn, Pearson Prentice Hall, Harlow.

Macrae, J., Jaspars, S., Duffield, M., Bradbury, M. & Johnson, D. (1997) 'Conflict, the Continuum and Chronic Emergencies: A Critical Analysis of the Scope for Linking Relief, Rehabilitation and Development Planning in Sudan', *Disasters*, vol. 21, no. 3, pp. 223-243.

Matthews, J. (2009) Media and message: communicating crises, *Humanitarian Exchange Magazine*, vol., no. 44, [Online]. Available from: http://www.odihpn.org/humanitarian-exchange-magazine/issue-44/media-and-message-communicating-crises (Accessed: 1 May 2013).

McLoy (1957) *Humanitarian Movement in Eighteenth-Century France*, University of Kentucky Press, Lexington.

Médecins Sans Frontières (2013) *About MSF, Our Finances* [Online], Available from: http://www.msf.org/our-finances (Accessed: 22 August 2013).

Middleton, N. & O'Keefe, P. (1998) *Disaster and Development: The Politics of Humanitarian Aid*, Pluto Press, London & Chicago.

Minear, L. (1994) International relief system: a critical review, *Parallel National Intelligence Estimate on Global Humanitarian Emergencies*, Washington, D.C., 22 September 1994.

Montreux Process (2000) Donors Common Observations, *Improving Coordination and Donor's Response in Complex Emergencies 9 – 10 March 2000*, Montreux.

Moore, J. (1999) The humanitarian-development gap, *International Review of the Red Cross*, vol., no. 833, [Online]. Available from: http://www.icrc.org/eng/resources/documents/misc/57jpt2.htm (Accessed: 8 May 2013).

Munslow, B. & Brown, C. (1999) 'Complex emergencies: the institutional impasse', *Third world Quarterly*, vol. 20, no. 1, pp. 207-221.

Munslow, B. & O'Dempsey, T.J. (2010) From War on Terror to War onWeather? Rethinking humanitarianism in a new era of chronic emergencies, *Third World Quarterly*, vol., no. 8, pp. 1223-1235 [Online]. Available from: http://dx.doi.org/10.1080/01436597.2010.542965 (Accessed: 4 October 2013).

North Atlantic Treaty Organisation (2003) *AJP-9 NATO Civil-military co-operation (CIMIC) doctrine (Unclassified publication)*.

Norwegian Refugee Council (2004) *Camp Management Toolkit*.

Norwegian Refugee Council (2008) *Camp Mangement Tool Kit*.

O'Dempsey, T.J. & Munslow, B. (2008) Complex Political Emergencies in the War on Terror Era in V. Desai & R.B. Potter (eds), *The Companion to Development Studies*, 2nd edn, Hodder Education, London, pp. 464-468.

O'Dempsey, T.J. & Munslow, B. (2009) 'Mind the gap!' rethinking the role of health in the emergency and development divide, *The International Journal of Health Planning and Management*, vol., no. S1, pp. S21-S29 [Online]. Available from: http://onlinelibrary.wiley.com/doi/10.1002/hpm.1020/abstract (Accessed: 16 October 2013).

O'Keefe, P., Kliest, T., Kirkby, J. & Wiert, F. (2001) Somalia: Towards Evaluating the Netherlands' Humanitarian Assistance, in A. Wood, R. Apthorpe & J. Borton (eds), *Evaluating International Humanitarian Action: Reflections from Practitioners*, Zed Books in association with ALNAP, London & New York, pp. 19-38.

O'Keefe, P. & Rose, J. (2008) Humanitarian Aid, in V. Desai & R.B. Potter (eds), *The Companion to Development Studies*, 2nd edn, Hodder Education, London, pp. 458-464.

Office of National Stastics (2013) *Labour Market Stastics, March 2013 Release*, UK Government,, London.

Olsen, G.R., Carstensen, N. & Høyen, K. (2003) 'Humanitarian Crises: What Determines the Level of Em ergency Assistance? Media Coverage, Donor Interests and the Aid Business', *Disasters*, vol. 27, no. 2, pp. 109-129.

Overseas Development Institute (2009) *Planning tools: Problem Tree Analysis* [Online], Available from: (Accessed: 13 August 2013).

Patrick, J. (2011) *Haiti Earthquake Response*, Network on Development Evaluation (OECD-DAC), London

Patton, M.Q. (2002) *Qualitative Research & Evaluation Tools*, 3rd edn, Sage Publications, Thousand Oaks, London, New Delhi.

Paulson, R. (1988) *Hogarth's graphic works*, 3rd rev. ed. edn, Print Room, 1989, London.

People in Aid (2013a) *About People in Aid* [Online], Available from: http://www.peopleinaid.org/about/ (Accessed: 7 Jan 2013).

People in Aid (2013b) *The harmonisation of humanitarian standards* [Online], Available from: http://www.peopleinaid.org/news/469.aspx (Accessed: 4 March 2014).

People in Aid (2013c) *Report and Financial Statements*,

Pickrell, J. (2005) Facts and Figures: Asian Tsunami Disaster, *New Scientist*, vol., [Online]. Available from: http://www.newscientist.com.ezproxy.liv.ac.uk/article/dn9931-facts-and-figures-asian-tsunami-disaster.html (Accessed: 2 Oct 2012).

Plummer, H.E. (2005) *Participant observation* [Online], Pearson Education, Available from: http://wps.pearsoned.co.uk/ema_uk_he_plummer_sociology_3/40/10342/2647687.cw/content/ (Accessed: 12 May 2014).

Porter, A. (2001) Trusteeship, Anti Slavery and Humanitarianism, in A. Porter (ed.), *The Oxford History of the British Empire*, vol. 3: The Nineteenth Century, Oxford University Press, Oxford, pp. 198-221.

Powell, C. (2001) *Remarks to the National Foreign Policy Conference for Leaders of Nongovernmental Organisations*.

Prunier, G. (2005) *Darfur : the ambiguous genocide*, Hurst & Company, London.

Rabbitt, E. (2003) *Insider research : the implications of conducting research in your home locale*, *Proceedings of the joint Australian Association for Research in Education /*

New Zealand Association for Research in Education Conference, Auckland, New Zealand.

Rack, H.D. (2004) Wesley [Westley], John (1703–1791), Church of England clergyman and a founder of Methodism *Oxford Dictionary of National Biography*, vol., [Online]. Available from: http://www.oxforddnb.com/view/article/29069 (Accessed: 3 Aug 2013).

Ramalingam, B., Scriven, K. & Foley, C. (2009) *Innovations in international humanitarian action*, Active Learning Network for Accountability and Performance,

Rieff, D. (2000) 'Kosovo's humanitarian circus', *World Policy Journal*, vol. 17, no. 3, pp. 25-32.

Rieff, D. (2003) *A Bed for the Night: Humanitarianism in Crisis*, Simon & Schuster, New York.

Robinson, P. (2002) *The CNN Effect: The Myth of News, Foreign Policy and Intervention*, Routledge, London & New York.

Rousseau, J.-J. (1968) *The social contract*, Penguin, London.

Runciman, S. (1947) *The Medieval Manichee*, Cambridge University Press, Cambridge.

Scheper, E., Parakrama, A. & Patel, S. (2006) *Impact of the Tsunami Response on Local and National Capacities*, Tsunami Evaluation Coalition, London

Schwartz, L., Sinding, C., Hunt, M., Elit, L., Redwood-Campbell, L., Adelson, N., Luther, L., Ranford, J. & DeLaat, S. (2010) 'Ethics in Humanitarian Aid Work: Learning From the Narratives of Humanitarian Health Workers', *AJOB Primary Research*, vol. 1, no. 3, pp. 45 - 54.

Sciba, D. (2011) Haiti Earthquake Facts and Figures, vol., [Online]. Available from: http://reliefweb.int/sites/reliefweb.int/files/resources/18548CFADA9D4B9D49257 81900243403-Full_Report.pdf (Accessed: 2 october 2012).

Scriven, M. (1996) The Theory behind Practical Evaluation, *Evaluation*, vol., no. 4, pp. 363 - 404 [Online]. Available from: http://evi.sagepub.com.ezproxy.liv.ac.uk/content/2/4/393.full.pdf+html (Accessed: 15 June 2011).

Sellstrom, T. & Wohlgemuth, L. (1996) *Study 1 Historical Perspectives: Some Explanatory Factors*, Sellstrom, T. & Wohlgemuth, L., vol. 1, Steering Committee of the Joint Evaluation of Emergency Assistance

to Rwanda [Online], Available from: http://www.grandslacs.net/doc/0741.pdf (Accessed: 13 January 2010).

Sharma, S. (2001) *Famine, Philanthropy and the Colonial State: North India in the Early Nineteenth Century*, Sharma, S., Oxford University Press [Online], Available from: http://www.oxfordscholarship.com.ezproxy.liv.ac.uk/view/10.1093/acprof:oso/978 0195653861.001.0001/acprof-9780195653861-chapter-2 (Accessed: 14 Aug 2013).

Shawcross, W. (1984) *The Quality of Mercy: Cambodia, holocaust and modern conscience*, Andre Deutch, London.

Sibeon, R. (1999) 'Agency, Structure and Social Chance as Cross Disciplinary Concepts', *Politics*, vol. 19, no. 3, pp. 139-144.

Silverman, D. (1993) *Interpreting Qualitative Data*, Sage Publications, London, Thousand Oaks, New Delhi.

Silverman, D. (2005) *Doing Qualitative Research*, 2nd edn, Sage Publications Ltd., London.

Slim, H. (2002) Claiming the humanitarian imperative *Refugee Survey Quarterly*, vol., no. 3, pp. 113-125 [Online]. Available from: http://rsq.oxfordjournals.org/content/21/3/113.full.pdf (Accessed: 23 January 2012).

Slim, H. (2003) Is humanitarianism being politicised? A reply to David Rieff, *paper delivered at The Dutch Red Cross Symposium on Ethics in Aid, The Hague, October*, vol. 8.

Slim, H. (2004) 'With or Against? Humanitarian Agencies and Coalition Counter-Insurgency', *Refugee Survey Quarterly*, vol. 23, no. 4, pp. 34-47.

Smith, J.A., Larkin, M.H. & Flowers, P. (2009) *Interpretative phenomenological analysis : theory, method and research*, SAGE, Los Angeles ; London.

Sniderman, A.D. & Furberg, C.D. (2009) 'Why Guideline-Making Requires Reform', *JAMA The Journal of the American Medical Association*, vol. 301, no. 4, pp. 429-431.

Stoddard, A. (2003) *Humanitarian NGOs : challenges and trends*, HPG, London.

Telford, J. (2001) UNICEF-DFID Joint Evaluation of UNICEF's Kosovo Emergency Preparedness and Response, in A. Wood, R. Apthorpe & J. Borton (eds), *Evaluating International Humanitarian Action: Reflections from Practitioners*, Zed Books in association with ALNAP, London & New York, pp. 171-189.

Telford, J., Cosgrave, J. & Houghton, R. (2006) *Joint evaluation of the international response to the Indian Ocean tsunami: Synthesis Report*, Tsunami Evaluation Coalition (TEC), London

Terry, F. (2002) *Condemned to Repeat: The Paradox of Humanitarian Action*, Cornell University Press, Ithaca & London.

The Bible: New International Version (1978), New York International Bible Society, New York.

The Nobel Foundation (1922) *Fridtjof Nansen - Biographical* [Online], Available from: http://www.nobelprize.org/nobel_prizes/peace/laureates/1922/nansen-bio.html (Accessed: 22 Aug 2013).

The Quran English Translation (1930), Govt. of the Nizam of Hyderabad, Hyderabad, Deccan.

The Sphere Project (2011) *Humanitarian Charter and Minimum Standards in Humanitarian Response*, 3 edn, The Sphere Project.

The Sphere Project (2013) *The Sphere Project in Brief* [Online], Available from: http://www.sphereproject.org/about/ (Accessed: 1 May 3013).

Thurer, D. (2007) 'Dunant's pyramid: thoughts on the" humanitarian space"', *INTERNATIONAL REVIEW-RED CROSS-NEW SERIES-*, vol. 89, no. 865, p. 47.

Tsunami Evaluaton Coalition (2007) *Consolidated lessons and recommendations from the TEC Synthesis Report and TEC Thematic Evaluations*, Tsunami Evaluation Coalition (TEC), London

United Kingdom Parliament (1999) *Conflict Prevention and Post-Conflict Reconstruction: Humanitarian Assistance* in Select Committee on International Development Sixth Report (ed.)London.

United Nations (2006) *Secretary-General's bulletin: Acceptance of pro bono goods and services*, United Nations, New York.

United Nations Childrens Fund (2005) *Emergency Field Handbook* New York.

United Nations Evaluation Group (2005) *Standards for evaluation in the UN system*, United Nations Evaluation Group, [Online], Available from: http://esa.un.org/techcoop/fellowships/SFOMeeting/ParticipantArea/Background Documents/4_UNEGStandardsEvaluation.pdf (Accessed: 2 February 2013).

United Nations General Assembly (1946) Resolution 57 (1) Establishment of an International Childrens' Emergency Fund, *General Assembly*, New York.

United Nations General Assembly (1950) Resolution 428 (V) Statute of the Office of the United Nations high Commissioner for Refugees, *General Assembly*, New York.

United Nations General Assembly (1961) Resolution 1714 (XVI) World Food Program, *General Assembly*, New York.

United Nations General Assembly (1971) Resolution 2861 (XXVI) Assistance in cases of natural disaster and other disaster situations, *General Assembly*, New York.

United Nations General Assembly (1991) Resolution 46/182 Strengthening of the coordination of humanitarian emergency assistance of the United Nations, *General Assembly*, New york.

United Nations General Assembly (1997) *Renewing the United Nations: A Programme for Reform*, United Nations, New York

United Nations General Assembly (2005) Resolution 60/1, World Summit Outcomes, *General Assembly*.

United Nations General Assembly (2009 (1)) Resolution A/63/677 Implementing the responsibility to protect: Report of the Secretary-General, *General Assembly*.

United Nations General Assembly (2009 (2)) Resolution GA/63/308, The responsibility to protect, *General Assembly*.

United Nations High Commissioner for Refugees (2003) *Framework for Durable Solutions for Refugees and Persons of Concern*, United Nations High Commissioner for Refugees, UNHCR [Online], Available from: http://www.unhcr.org/3f1408764.pdf (Accessed: 8 May 2013).

United Nations High Commissioner for Refugees (2005) *Environmental Guidelines*, UNHCR, Geneva.

United Nations High Commissioner for Refugees (2006) *Catalogue of Emergency Response Resouirces*, United Nations High Commissioner for Refugees, Geneva.

United Nations High Commissioner for Refugees (2007) *Handbook for Emergencies*, 3 edn.

United Nations High Commissioner for Refugees (2010) *UNHCR's evaluation policy*, United Nations High Commissioner for Refugees, United Nations High Commissioner for Refugees [Online], Available from: http://www.unhcr.org/3d99a0f74.html (Accessed: 23 June 2012).

United Nations High Commissioner for Refugees (2013) *Financial Figures* [Online], UNHCR, Available from: http://www.unhcr.org/pages/49c3646c1a.html (Accessed: 22 Aug 2013).

United Nations High Commissioner for Refugees & World Food Programme (1997) *Cooperation Memorandum of Understanding on the Joint Working Arrangements for Refugee, Returnee and Internally Displaced Persons*, United Nations High Commissioner for Refugees & World Food Programme, [Online], Available from: http://www.unhcr.org/refworld/docid/3ae6b31a10.html (Accessed: 1 October 2011).

United Nations High Commissioner for Refugees & World Food Programme (1999) *Guidelines for Selective Feeding Programmes in Emergencies*, United Nations High Commissioner for Refugees & World Food Programme, [Online], Available from: http://www.who.int/nutrition/publications/en/selective_feeding_emergencies.pdf (Accessed: 1 October 2011).

United Nations High Commissioner for Refugees & World Food Programme (2008) *Joint Assessment Mission (JAM) Guidelines*, 2 edn.

United Nations Inter-Agency Standing Committee (2004) *Civil-military relationship in complex emergencies - an IASC reference paper*.

United Nations Inter-Agency Standing Committee (2008) *Civil-military guidelines & reference for complex Emergencies*, New York.

United Nations Office for the Coordination of Humanitarian Affairs (2003) *Glossary of Humanitarian Terms: In relation to the Protection of Civilians in Armed Conflict*, United Nations, New York.

United Nations Office for the Coordination of Humanitarian Affairs (2007) *Oslo Guidelines: Guidelines for the use of Foreign Military and Civil Defence Assets in Disaster Relief*, Revision 1.1 edn, United Nations, New York.

United Nations Office for the Coordination of Humanitarian Affairs (2011) *OCHA on Message*, *Humanitarian Principles*, United Nations, New York.

United Nations Office for the Coordination of Humanitarian Affairs (2013a) *Financial Tracking Service: Global Summary for 2012: Global Humanitarian Funding:Total Contributions per Donor in 2012* [Online], Available from: http://fts.unocha.org/pageloader.aspx?page=emerg-globalOverview&Year=2012 (Accessed: 22 Aug 2013).

United Nations Office for the Coordination of Humanitarian Affairs (2013b) *History of OCHA* [Online], Available from: http://www.unocha.org/about-us/who-we-are/history (Accessed: 9 May 2013).

United Nations Publications (2000) *The predecessor: The League of Nations* [Online], United Nations, Available from: http://www.un.org/cyberschoolbus/unintro/unintro3.htm (Accessed: 22 Aug 2013).

United Nations Security Council (2004) *Resolution 1542*, United Nations, New York.

United Nations Security Council (2010) *Resolution 1908*, United nations, New York.

United Nations Security Council (2011) Resolution SC/1973,, *Security Council*.

United States Army (2009) *Commander's Guide to Money as a Weapons System: Tactics, Techniques and Procedures*, United States Army, Center for Army Lessons Learned (CALL) [Online], Available from: http://usacac.army.mil/cac2/call/docs/09-27/09-27.pdf (Accessed: 22 August 2013).

von Planta, C. & Atkins, S. (2004) *Dispatches - Living with Refugees* Insight News TV.

Waldron, S. (1987) 'Blaming the Refugees', *Refugee Issues*, vol. 3, no. 3.

Walker, P. & Maxwell, D. (2009) *Shaping the Humanitarian World*, Routledge, New York.

Walkup, M. (1997) 'Policy dysfunction in humanitarian organisations: The role of coping strategies, institutions and organisational culture', *Journal of Refugee Studies*, vol. 10, no. 1, pp. 37-60.

Weir, E.A. (2006) *Conflict and compromise: UN integrated missions and the humanitarian imperative*, Kofi Annan International Peacekeeping Training Centre.

Weisman, S.R. (2004) Powell Declares Genocide in Sudan in Bid to Raise Pressure, *New York Times*, 9 September 2004.

Western, J. (2002) 'Sources of Humanitarian Intervention: Beliefs, Information, and Advocacy in the U.S. Decisions on Somalia and Bosnia ', *International Security*, vol. 26, no. 4, pp. 112-142.

Wiles, P. (2001) Review of the International Federation of Red cross and Red Crescent Societies' Tajikistan Programme, in A. Wood, R. Apthorpe & J. Borton (eds), *Evaluating International Humanitarian Action: Reflections of Practitioners*, Zed Books, London & New York, pp. 122-132.

Wood, A. (2001) Exploring the Swedish Emergency Relief Experience in the Horn of Africa, in A. Wood, R. Apthorpe & J. Borton (eds), *Evaluating International Humanitarian Action: Reflections from Practitioners*, Zed Books in association with ALNAP

London & New York, pp. 39-58.

Wood, A., Apthorpe, R. & Borton, J. (2001a) Conclusions, in A. Wood, R. Apthorpe & J. Borton (eds), *Evaluating International Humanitarian Action: Reflections from Practitioners*, Zed Books in association with ALNAP, London & New York, pp. 190-211.

Wood, A., Apthorpe, R. & Borton, J. (2001b) *Evaluating International Humanitarian Action: Reflections from Practitioners*, Zed Books in association with ALNAP, London & New York.

Wood, A., Apthorpe, R. & Borton, J. (2001c) Introduction, in A. Wood, R. Apthorpe & J. Borton (eds), *Evaluating International Humanitarian Action: Reflections from Practitioners*, Zed Books in association with ALNAP, London & New York, pp. 1-18.

Wulf, H. & Debiel, T. (2009) *Conflict Early Warning and Response Mechanisms: Tools for enhancing the effectivness of regional organisations*, in Crisis States Research Centre Development Studies Institute LSE (ed.)London.

Zenco, M., Fellow, D. & Friedman, R. (2011) 'UN Early Warning for Preventing Conflict', *International Peacekeeping*, vol. 18, no. 1.

Zolberg, A.R., Suhrke, A. & Aguayo, S. (1989) *Escape from violence : conflict and the refugee crisis in the developing world*, Oxford University Press, New York.