

Working Paper 215

**How the Sphere Project Came into Being:
A Case Study of Policy-Making in the Humanitarian Aid
Sector and the Relative Influence of Research**

Margie Buchanan-Smith

July 2003

Overseas Development Institute
111 Westminster Bridge Road
London
SE1 7JD
UK

ISBN 0 85003 665 8

© Overseas Development Institute 2003

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the publishers.

Contents

Acknowledgements	iv
Acronyms	v
Executive Summary	vi
1 Introduction	1
1.1 Case study of the Sphere project	1
2 The Research Framework and Approach	3
2.1 An integrated framework	3
2.2 The approach	5
3 The Story of Sphere’s Origins	6
3.1 The overall context – the end of the Cold War and the changing face of humanitarianism	6
3.2 Early concerns about behaviour, standards and performance	6
3.3 The Rwanda crisis	8
3.4 The international humanitarian response to the Rwanda crisis, and the critique	8
3.5 Rwanda evaluation	9
3.6 The beginnings of Sphere	11
3.7 Sphere gets off the ground	12
4 Explanation	14
4.1 The policy context: politics and institutions	14
4.2 Evidence	16
4.3 Links: influence and legitimacy	18
5 Discussion	21
5.1 The context: politics and institutions	21
5.2 Evidence: credibility and communication	22
5.3 Links: influence and legitimacy	23
6 Conclusions	24
References	25
Appendix 1 Research Methodology and Timing	27
Appendix 2 Key Informants	28
Appendix 3 The Research Questions	29
Appendix 4 Sphere Project – Chronology	31

Acknowledgements

The author would like to thank all the people listed in Appendix 2 who gave their time so generously in interviews for this study. Their openness and insights were much appreciated. The author takes full responsibility for this interpretation of the story. Special thanks to Nan Buzard, the current Sphere project manager, for making this case study possible, and for facilitating access to key people and relevant documents in Geneva. The author is also grateful to members of the Sphere Management Committee, for their permission and encouragement in carrying out this case study. Thanks also to all those who commented on earlier drafts of this report: Nan Buzard, Susan Purdin, Niels Dabelstein, John Borton, Jim Bishop, Peter Walker, Tony Vaux, Frances Stevenson and Sean Lowrie. Finally, the author would like to acknowledge John Young's support throughout this project.

Margie Buchanan-Smith is a Research Associate with ODI, and was formerly Coordinator and a Research Fellow within ODI's Humanitarian Policy Group. Her experience within the humanitarian aid sector ranges from policy research to operational management. She was Head of ActionAid's Emergencies Unit between 1995 and 1998.

Email: Margiebsmith@aol.com

Acronyms

ACT	Action by Churches Together
AHA	African Humanitarian Action
ALNAP	Active Learning Network for Accountability and Performance in Humanitarian Action
DAC	Development Assistance Committee
EU	European Union
FAR	Force Armeé Rwandaise
HAP	Humanitarian Accountability Project
IBRD	International Bank for Reconstruction and Development
ICVA	International Council of Voluntary Agencies
ICRC	International Committee of the Red Cross
IFRC	International Federation of Red Cross and Red Crescent Societies
IOM	International Organisation for Migration
JEFF	Joint Evaluation Follow-up, Monitoring and Facilitation Network
LWF	Lutheran World Federation
MSF	Medecins Sans Frontieres
NGO	Non-governmental organisation
ODA	Overseas Development Administration (later became DFID, UK Government)
OECD	Organisation for Economic Cooperation and Development
OFDA	United States Office of Foreign Disaster Assistance (USAID)
PVO	Private voluntary organisation
RPF	Rwandese Patriotic Front
RRN	Relief and Rehabilitation Network
SCF-UK	Save the Children Fund – UK
SCHR	Steering Committee for Humanitarian Response
TOR	Terms of Reference
UK	United Kingdom
UN	United Nations
UN DHA	UN Department for Humanitarian Affairs
UNDP	United Nations Development Programme
UNHCHR	United Nations High Commissioner for Human Rights
UNHCR	United Nations High Commission for Refugees
UNICEF	United Nations Children’s Fund
US	United States
USAID	United States Agency for International Development
VOICE	Voluntary Organisations in Cooperation in Emergencies
WFP	World Food Programme
WHO	World Health Organisation

Executive Summary

One of the most significant policy shifts in the international humanitarian sector in the last decade has been the move to strengthen the accountability of humanitarian agencies and to find ways of improving performance in humanitarian response. The decision to launch the Sphere project in 1996 was one of the key policy initiatives associated with this shift. It resulted in the publication of a ‘Humanitarian Charter and Minimum Standards for Disaster Response’ in 2000 (Sphere, 2000). This case study explores the process that led up to this policy initiative, and how buy-in and ownership were achieved (or in some cases, not) during the first year of Sphere’s existence, in order to learn more about the linkages between research and policy.

In the first half of the 1990s, the policy context within the humanitarian aid sector was becoming increasingly conducive to an initiative such as Sphere. The days of unquestioning acceptance of the ‘good work’ of humanitarian agencies were over. As agencies entered more difficult conflict environments, they were subject to much more rigorous scrutiny and more sophisticated political analysis. There was a growing sense of unease amongst some agencies and individuals about the range of standards and performance to which different agencies operated. This gave rise to initiatives such as the ‘Code of Conduct for the International Red Cross and Red Crescent Movement and NGOs in Disaster Relief’, drafted in Europe (RRN, 1994), and the Providence principles, drafted in the US (Minear and Weiss, 1993). These were attempts to provide some codification of the basic principles that should guide agencies in a humanitarian emergency.

Although these concerns created an atmosphere conducive to the Sphere project, it was the scale and intensity of the humanitarian crisis in Rwanda in 1994, which determined the vigour, depth and direction of its study. The refugee crisis that followed and the unprecedented international humanitarian response, under the spotlight of the international media, publicly magnified the fault lines within the humanitarian system to an extraordinary degree. These were major factors in creating the momentum for change, particularly amongst NGOs, and their increasing openness to radical change. This was articulated in the oft-repeated mantra within the NGO community: ‘we have to get our act together’, and reinforced by the response of some donor representatives: ‘if you don’t do it, we will’. Thus, there were strong push factors to do something to improve NGO performance. This seems to be quite typical of the humanitarian system. It is most responsive to change under pressure, when the push factors are strong. A few individuals (and agencies) capitalised upon this unprecedented policy window of opportunity in order to get their ideas accepted, eventually resulting in the Sphere initiative.

These ideas were given real impetus by the ‘categoric, clear and well-documented’¹ report of Study 3 of the Joint Evaluation of Emergency Assistance to Rwanda. Critical of some NGO performance in the Rwanda crisis, it commended the embryonic NGO initiative to develop performance standards. (The Sphere project developed out of two separate initiatives that pre-dated the Rwanda evaluation: In the US, InterAction commenced a project to develop best-practice for disaster work. Meanwhile in Europe, two individuals from Oxfam and the IFRC were developing a project to improve quality and accountability standards in humanitarian relief). But Study 3 went further than any NGOs in also recommending some form of self-managed regulation or accreditation of NGOs to monitor compliance. NGOs have however consistently shied away from this since.

Study 3 of the Joint Evaluation had real impact on policy-makers. This well-resourced evaluation had high credibility, partly because of the composition of the team, who were seen to be objective, combining research and operational experience. Study 3 was also seen to be a very thorough piece of work, more akin to a research project than a conventional consultancy. Its findings and recommendations were clearly presented, and were often targeted at particular groups of actors.

¹ Personal communication, Peter Walker.

Also, crucially, the structure for managing the evaluation proved critical to safeguarding the independence of the sometimes unpopular findings of the Study 3 team. The Management Group, comprising objective evaluators from a selection of donor agencies, was able to shield the evaluation team from undue pressure from the agencies and governments who were being evaluated, and who were represented on the evaluation Steering Committee.

This case study supports the hypothesis that close links between researchers and policy-makers greatly increase the likelihood of research influencing policy. Those links were, unusually, institutionalised in the structure put in place for the Joint Rwanda Evaluation. The evaluation was effectively commissioned by the respective evaluation departments of different donor agencies, represented on the evaluation's Management Group. However this was done with the support and input of policy-makers who were represented on the evaluation's Steering Committee. Thus, a critical and cooperative link was established right at the beginning between those who commissioned the 'research', and the policy-makers at whom the findings were directed. Also, there was a clear synergy of interest between the evaluators and certain key policy-makers; they were working to a similar agenda, to bring about change to improve NGO performance. The evaluation's endorsement of these early initiatives gave them real momentum. This inclusive experience of the Rwanda evaluation influenced the process put in place for Sphere. All the main NGO umbrella organisations, plus individual agencies, were given membership of the Project Committee. This highly inclusive process, designed to get maximum buy-in, was put in place in order to draft Sphere documents in its first year.

There were also certain key actors who played catalytic roles in Sphere's birth and in its infancy. Particularly noteworthy is the role of two 'policy entrepreneurs' – Peter Walker and Nick Stockton – who are most closely associated with the launch of Sphere. In Gladwell's (2000) terms they were 'connectors' (i.e. networkers) and 'salesmen' (i.e. persuasive, trusted and listened to). They were also 'innovators' and they spotted the policy window to take forward their ideas. Other individuals, such as Jim Bishop of InterAction and Susan Purdin, the first Sphere project manager, also played critical 'connector' roles.

But whilst most humanitarian actors shared a common analysis of the varied and sometimes poor performance of NGOs in response to the Rwanda refugee crisis (and elsewhere), and that something had to be done, there was a divergence of views about an appropriate solution. Although most agencies were eventually in favour of the proposed performance standards, a group of mainly Francophone NGOs, including MSF, were not. Their concerns have been the focus of much debate since. They argued that Sphere was in danger of isolating technical standards from humanitarian principles and deflecting attention from the ruling authorities and international community who remain responsible for protecting humanitarian space. The inclusion of Sphere's Humanitarian Charter in the report, designed to alleviate some of these concerns, was regarded by these agencies as an uncomfortable institutionalisation of the Sphere process.

In conclusion, this case study shows that the policy impact of Study 3 of the Rwanda evaluation, on finding ways of improving NGO performance, was more or less immediate in terms of the impetus it gave to the embryonic Sphere project. However, the impact was also partial; the more challenging and radical recommendations around accreditation and regulation of NGOs were ducked. This case study supports the following three hypotheses that research is more likely to contribute to evidence-based policy if:

- it fits within the political and institutional limits and pressures of policy-makers, and resonates with their ideological assumptions, or sufficient pressure is exerted to challenge those limits;
- researchers and policy-makers share particular kinds of networks and develop chains of legitimacy for particular policy areas;
- outputs are based on and credible evidence and are communicated via the most appropriate communicators, channels, style, format and timing.

1 Introduction

This Working Paper forms part of the Overseas Development Institute's Bridging Research and Policy project which is seeking to learn more about linkages between development research, policy and practice and to develop simple tools for researchers and policy-makers to promote evidence-based international development policy. The research project includes a literature review, the development of a framework paper to guide the research and three case studies, with local partners, of specific policy changes where research may, or may not, have played a significant role. These are:

- **Poverty Reduction Strategies.** How, during 1999, the international discourse about the Common Development Framework became linked to the adoption of the Enhanced HIPC framework by the G8, and then translated into the process of preparing the first interim Poverty Reduction Strategy Papers. What happened in between? Who influenced whom, on what and how? What was the specific contribution of research-based knowledge, and what conditions enabled this influence to be exercised in such a striking way?
- **Humanitarian Aid.** One of the most significant policy shifts in the international humanitarian sector in the last decade has been the move to strengthen the accountability of humanitarian agencies and to find ways of improving performance in humanitarian response. One of the key policy initiatives, representative of this shift, was the decision to launch the Sphere project in 1996, in the wake of the much-criticised international humanitarian response to the Rwanda crisis which began in 1994. Sphere resulted in the publication of a 'Humanitarian Charter and Minimum Standards for Disaster Response' in 2000. This case study explores the process that led up to this policy initiative. For example, how significant was the Joint Evaluation of Emergency Assistance to Rwanda? What were the other key factors that triggered the launching of Sphere? How significant was the policy context, in which humanitarian agencies were subject to harsh and public criticism?
- **Livestock Services.** Livestock services have long been regarded as an easy target for reform and privatisation, first under structural adjustment programmes in the late 1980s and early 1990s, and more recently, as part of re-orientating agricultural services under poverty reduction strategies. Veterinarians and governments in most countries, however, have been very reluctant to liberalise the policy framework to allow private and especially para-professional services to flourish, despite good evidence that paravets can provide an effective, cost-efficient, and safe service. This research identifies the critical factors and the relevance of research in the evolving livestock service policies particularly in Eastern Africa, and the Horn of Africa.

The literature review² has been carried out, and a framework has been developed to guide the research.³ This Working Paper presents the results of the case study from the humanitarian aid sector, tracing the origins of the Sphere project.

1.1 Case study of the Sphere project

The decision to launch the Sphere project was, essentially, a policy decision to strengthen the quality of work of humanitarian agencies, in particular non-governmental organisations (NGOs), and their accountability – especially their accountability to those affected by emergencies. Concerns about improving the professionalism of humanitarian aid work had been around since the late 1980s. However, the high-profile and publicly-criticised international humanitarian response to the Rwanda crisis, beginning in 1994, meant that these concerns became much more widespread, with a

² de Vibe et al. (2002).

³ Crewe and Young (2002).

greater sense of urgency that they needed to be addressed. The Sphere project was one of a number of initiatives that aimed to do just that.⁴

Sphere is probably best-known for developing minimum standards in five sectors that cover the basic and most common life-saving relief interventions required in many humanitarian emergencies:

- water supply and sanitation
- nutrition
- food aid
- shelter and site planning
- health services

It also developed the ‘Humanitarian Charter’, regarded as ‘the cornerstone’ of Sphere (Sphere, 2000: 1), although referred to less frequently by agencies using Sphere.

‘Based on the principles and provisions of international humanitarian law, international human rights law, refugee law, and the *Code of Conduct for the International Red Cross and Red Crescent Movement and NGOs in Disaster Relief*, the Charter describes the core principles that govern humanitarian action and asserts the right of populations to protection and assistance’ (ibid.: 1).

The Sphere project is intended to be relevant to the entire humanitarian sector, although it is predominantly an NGO initiative. The extent to which there has been widespread buy-in is impressive. However it has also provoked significant controversy, and there have been strong critics of Sphere, particularly from a number of francophone NGOs who have rejected Sphere (and in some cases begun to explore alternatives e.g. the Quality Platform developed by some French NGOs). It has been a largely ‘northern’-driven policy initiative, although there are many agencies and institutions in the ‘south’ who are now strong supporters of Sphere and who have played a role in the piloting of Sphere.

There are two stories to tell when exploring the origins of Sphere. The first story is about how Sphere was born. This is about understanding the different factors that led to the decision to launch Sphere, and in particular the relative influence of the ‘Joint Evaluation of Emergency Assistance to Rwanda’. Study 3 of this evaluation reviewed the performance of humanitarian agencies in responding to the Rwanda crisis (Borton et al., 1996). Its findings were critical, and highlighted the need for NGOs to improve their performance. It recommended that there be some form of self-regulation or accreditation, and that systems for accountability needed to be strengthened. This seminal evaluation is often described as a piece of policy research,⁵ and is considered as such in this project about bridging research and policy.

The second story, which of course is linked, is about how the Sphere project developed in its first year of existence, particularly in terms of how buy-in and ownership were achieved (or in some cases, not). This story throws up interesting and important lessons – for both researchers and policy-makers – about the policy process, especially when success depends upon buy-in from a range of different actors.

This paper covers both stories, reflecting on what can be learned, and especially on what has worked in both phases that is relevant to bridging research and policy.

⁴ Other initiatives include the Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP), and the Humanitarian Ombudsman Project which later became the Humanitarian Accountability Project (HAP).

⁵ For example, Borton describes it as ‘closer to a commissioned multi-disciplinary research study than a conventional evaluation’ (Borton, 2001: 74).

2 The Research Framework and Approach

Although many researchers have been driven by the aim to influence policy in the field of international development, research is often ignored, top-down, inaccurate or neglects the concerns of poor or marginalized people.

Research and policy defy neat separation but can be conceptually distinguished by their goals and methods; *research* produces knowledge (whether by, for example, action-research or academic study); *policy* aims for continuity or change of a practice (stretching from the statements on intent to the development of policy in practice).

Policy processes are influenced by a wide range of factors, personal and political, as well as, if not more so than knowledge and evidence, which may be practical and experiential or research-based. There may be a considerable time-lag between new knowledge acquisition and policy formulation, as is often the case with pure research, whereas with policy-driven action-research, the policy-impact may be more or less immediate.

For these case studies, ODI has taken an historical approach to map the key events and decisions leading to a clear policy shift, and applied a three-dimensional framework to analyse what influenced each decision.

2.1 An integrated framework

Ideas percolate into development discourse and people learn about how the world is ordered through practical experience, the media and intellectual activity. Political interests, formations of actors, and discourses, the role played by wider civil society and ‘street bureaucrats’, and concepts from psychology and marketing all influence how ideas and research based-evidence influence the policy process. Based on a wide-ranging review of the literature, ODI has developed a three-dimensional framework to assist the investigation into the impact of research on policy. The framework seeks to integrate an understanding of the political context, quality and legitimacy of the evidence, and the links between policy and research communities to explain why, how and when evidence influences policy.

The context: politics and institutions

The institutional, cultural and structural dimensions of international development are diverse, dynamic and overlapping. Policy-makers and researchers cut across categories but their position of power, and the aims of the organisations they work for, can be identified. They are not only limited by macro political and economic structures but also the assumptions underlying them. For example, it is arguable that belief in the need for economic growth both rationalises and serves the global economic system. The way ideology is gendered reflects power relations as well; the global division of labour – with men dominating most policy making structures – explains why men and women have different degrees of room for manoeuvre.

Shifts in worldview – where an explanatory model meets a crisis and is replaced by an alternative – may take place at different levels of development discourse as a consequence of intellectual revolutions or campaigns of citizen action. But while researchers in academia and NGOs often see themselves as championing radical causes, staff within bureaucracies tend to resist fundamental challenges to the status quo.

Successful strategies for influencing policy-makers and researchers have to take account of the various bureaucratic pressures limiting and enabling them, as well as those who commission or fund research:

- *the urge to simplify*: due to resource shortages
- *giantism*: the bigger the budget, the greater the status
- *inflexible long-term project planning*
- *fierce competition for funding*: discouraging collaboration

The complexity and diversity intensifies even further when researchers consider how policy is adapted, developed or distorted during practice. What influences policy practices will vary according to the priorities of street-level bureaucrats, local history, ideologies, and power relations.

Evidence: credibility and communication

The sources and conveyors of information may be as influential as the content; for example, people accept information more readily from those they trust. But it is the hypothesis of this paper that good quality research, local involvement, accurate messages *and* effective dissemination strategies are all important if aiming for more evidence-based policy-making. A lack of local involvement in drawing conclusions from research can lead to worrying impacts, such as children finding more abusive forms of work when child labour was banned in certain industries.

The temptation to keep messages simple can be strong but they are then more likely to be manipulated by policy-makers to gloss over the complexity of proposed solutions. There is no shortage of other ideas about why some information makes a mark. The element of surprise can usefully attract people's attention but not if the contents of the message are impossible for them to relate to their existing knowledge in any way.

Assessing the effectiveness of particular communication strategies on particular audiences would appear to be an important part of any attempt to change policy. Information technologies have huge possibilities for communication within networks and coalitions but are not sufficient for building trust; require investment in people as well as technology; can be resisted; can accentuate exclusion; and can lead to the erosion of public space.

Links: influence and legitimacy

The nature of the relationship between researchers and policy-makers shapes how much influence they have over each other. Policy networks usually share some common values and outlooks and consciously work together to take advantage of policy 'spaces' and 'windows'. Policy-makers themselves can be members of these networks, but if not directly involved in the influencing process, other kinds of support (such as citizen action) can become critical.

What influences policy making will depend on the policy area and geopolitical level. Scientific arguments are important in some areas, lobbying by professional groups is vital in others. Certain patterns about the perception of expertise prevail across all policy areas; the marginalisation or appropriation of indigenous knowledge is well documented. The reversals advocated in Participatory Rural Appraisal techniques – and its spin-offs – have challenged these power relations to some extent but have sometimes been merely another tool for 'developers' to control encounters more effectively.

It has been suggested that international NGOs will have to be more accountable to the citizens they claim to represent, more ‘south’-driven when international development goals are in mind, and extremely well-informed, to be taken seriously. They have been advised to map out their legitimacy chains through systems of accountability (building structures that are representative of local constituencies where necessary) or relevant experience. Researchers who ensure they build legitimacy chains to their informants may also be less easy to ignore.

2.2 The approach

The hypothesis here is that research is more likely to contribute to evidence-based policy making that aims to reduce poverty, alleviate suffering or save lives if:

- it fits within the political and institutional limits and pressures of policy-makers, and resonates with their ideological assumptions, or sufficient pressure is exerted to challenge those limits;
- researchers and policy-makers share particular kinds of networks and develop chains of legitimacy for particular policy areas;
- outputs are based on local involvement and credible evidence and are communicated via the most appropriate communicators, channels, style, format and timing.

To test these hypotheses ODI constructed an historical narrative leading up to the observed policy change in each case study. This involved creating a timeline of key policy decisions and practices, along with important documents and events, and identifying key actors. The next step was to explore why those policy decisions and practices took place and assess the role of research in that process. This was achieved through interviews with key actors and reviewing the literature, and cross-checking conflicting narratives. See Appendix 1 for more details on the research methodology and timing. The list of key informants is provided in Appendix 2, and a list of the key research questions in Appendix 3.

3 The Story of Sphere's Origins

3.1 The overall context – the end of the Cold War and the changing face of humanitarianism

The end of the Cold War marked a turning point in the way that humanitarian emergencies were perceived, and in how the international humanitarian response was organised. The political causes of many emergencies were more widely and openly acknowledged,⁶ as by proxy some wars came to an end, and as conflict became predominantly intra-state. The term 'complex political emergency' was coined, and has been used to describe many of the most serious humanitarian crises during the 1990s, from the Kurdish crisis in 1991, to the Rwanda crisis in the middle of the decade, and the Kosovo crisis at the end of the decade. There was a willingness to undertake military intervention in civil war situations, ostensibly for humanitarian objectives, for example, to create safe havens for the Kurdish population in northern Iraq; in Somalia in 1992; and at the end of the decade in Kosovo. But this has been a very selective strategy and a controversial one.⁷

Humanitarian aid flows rose sharply and have doubled or even tripled during the 1990s compared with the previous decade. The number of international NGOs engaged in humanitarian work has increased substantially during the 1990s, particularly evident in the high-profile Rwanda and Kosovo crises. The amount of donor funding channelled through NGOs has also risen significantly, to around 25% of their humanitarian aid flows by the end of the decade.⁸ Indeed, from the early 1990s onwards, humanitarian aid began to penetrate ever more difficult and contested environments, as humanitarian space⁹ 'opened up' after the end of the Cold War.

Associated with this change in political context, and with the growth of the 'humanitarian industry', has been much more rigorous scrutiny and analysis. This was evident in the growing number of evaluations of humanitarian operations, commissioned by NGOs, by donor governments and by UN agencies. It was also evident in the increasingly sophisticated analysis of the causes and consequences of humanitarian emergencies, and in particular of the political/humanitarian interface.¹⁰ This heralded a much more critical approach to emergency relief work, as it became apparent how the relief industry could become co-opted into the dynamics of violent conflict.¹¹

3.2 Early concerns about behaviour, standards and performance

The opening up of humanitarian space propelled agencies into ever more difficult conflict environments. As a result, they were faced with the day-to-day challenges and dilemmas of how to operate as an independent and impartial humanitarian actor – challenges that were familiar to agencies such as the International Committee of the Red Cross (ICRC), but less so to many NGOs, particularly those whose roots were in development rather than humanitarian work. Agencies began to grapple with the concept of humanitarian principles, *and* with the implications of inconsistency, where different agencies were working to different approaches and frameworks.

⁶ This is in contrast to the portrayal of some emergencies in the 1980s, such as the Ethiopia emergency in the mid 1980s when drought was emphasised rather than the political causes of the famine.

⁷ See, for example, Woodward (2001); Macrae and Leader (2000).

⁸ Macrae (2002).

⁹ Duffield (1994) describes humanitarian space as negotiated access, where the agreement of warring parties has been secured for the movement of neutral humanitarian aid.

¹⁰ See, for example, Keen (1994).

¹¹ See, for example, De Waal (1997).

This also gave rise to concerns about agency performance. Staff within some of the more-established humanitarian agencies became increasingly uneasy about the range of practice – and bad practice – amongst some NGOs, and about breakdown in cooperation at field level. For some, this uneasiness began during the widespread food crises and famines in sub-Saharan Africa in the mid-1980s, when they witnessed a range of undesirable emergency aid practices. Also, there was now more competition for donor funding as a sub-contracting culture became more established, and as the number of NGOs increased. As NGOs began to ‘globalise’, differences in approach between NGOs originating in different countries became more apparent, for example in fund-raising practices and in their use of the media.¹²

As the humanitarian industry grew in size and prominence, a number of initiatives were triggered within different agencies and in different parts of the world to codify and to develop some kind of written rules about how humanitarian agencies should work, to ‘re-establish’ some of the basic principles. The French Red Cross Society proposed such an initiative to the International Federation of Red Cross and Red Crescent Societies (IFRC). At about the same time, the Emergencies Coordinator in Oxfam (Tony Vaux) decided that it was time to codify and put down in writing some of the basic, so far unwritten, principles to guide Oxfam staff in emergencies.¹³ These efforts were combined under the auspices of the Steering Committee for Humanitarian Response (SCHR).¹⁴ The Head of the IFRC’s Disaster Policy Department (Peter Walker) worked with Oxfam’s Emergencies Coordinator to produce the ‘Code of Conduct for the International Red Cross and Red Crescent Movement and NGOs in Disaster Relief’ (RRN, 1994). This sets out 10 basic principles of behaviour for agencies and their staff, and outlines what, in turn, agencies expect from the state. The Code was published in 1994 as a Network Paper by ODI’s Relief and Rehabilitation Network (RRN, 1994). However, there was no communications strategy to promote the code. Instead, it was ‘lodged’ with the IFRC, which kept a record of all the agencies that voluntarily signed up to the Code and reported on this in the annual World Disasters Report. Although the Code was widely welcomed, there was no means of verifying whether agencies were adhering to it or not, hampered by the very general nature of the principles outlined in the Code.

Meanwhile, on the other side of the Atlantic the relatively new Humanitarianism and War project was also grappling with the challenges of doing humanitarian work in conflict situations. It came up with the Providence principles and a code of conduct for practitioners, as ‘fundamental objectives towards which humanitarian action should be oriented ... benchmarks against which performance can be measured and (to) help prevent energetic pragmatism from degenerating into unprincipled opportunism’ (Miner and Weiss, 1993: 5–6). It built upon an earlier piece of work by Miner (1988), commissioned by InterAction, on professionalism in US voluntary humanitarian assistance.

These are all indicative of a trend within the sector to clarify how best to work in situations of conflict (and, to a lesser extent, natural disasters), and how to achieve some kind of consensus and consistency, and to professionalise. On this latter point, it is worth noting another initiative in the early 1990s: to establish minimum standards and best practice in the recruitment and support of emergency staff. This was proposed by four British-based agencies in 1994, and eventually resulted in the People in Aid Code of Best Practice, published in 1997.¹⁵ In many ways, the humanitarian system was ‘coming of age’.¹⁶

¹² For instance, one NGO reportedly hired a crew to take photos of the famine in Ethiopia which were then sold to the media. This was regarded as distasteful practice by many other NGOs.

¹³ Oxfam did have written policies, but lacked a statement of the fundamental principles.

¹⁴ SCHR is an alliance for voluntary action of CARE International, Caritas Internationalis, IFRC, International Save the Children Alliance, Lutheran World Federation, MSF International, Oxfam International and World Council of Churches. ICRC joined in 2001.

¹⁵ People in Aid, ‘Code of Best Practice in the Management and Support of Aid Personnel’, first published by RRN, ODI, in 1997.

¹⁶ See also Leader (1999) for a full review of these and other ‘code of conduct’ initiatives.

3.3 The Rwanda crisis¹⁷

Against this backdrop, the Rwanda crisis happened in 1994. For at least three years beforehand, there had been fighting between the Rwandese Patriotic Front (RPF) and the Government of Rwanda, causing internal displacement of up to 900,000 people. Large-scale relief operations were launched in response. But it was the shooting down of President Habyarimana's plane on April 6th 1994 that triggered one of the worst humanitarian crises of the twentieth century. At the heart of the crisis was a genocide of horrifying magnitude. 500,000 to 800,000 people were violently killed over a three to four month period.

The crisis was also associated with an unprecedented refugee movement. As the RPF gained control of more of the country, capturing the towns of Kigali, Butare and Ruhengeri by the middle of July, a massive outflow was triggered of around 850,000 people into Goma, Zaire, over a four-day period between 14th and 18th July 1994. Almost all of those fleeing Rwanda were Hutu – civilians and the Force Armee Rwandaise (FAR). This huge refugee population arrived in a terrain of waterless lava fields, precipitating a sudden and massive humanitarian crisis. Within a few days cholera had broken out, claiming the lives of around 30,000 people within barely two weeks – the highest mortality rate ever recorded in a humanitarian emergency. Overall, the combination of cholera, dysentery, dehydration and violence claimed around 80,000 refugee lives. A total of about 1.5 million refugees fled Rwanda to neighbouring countries as the RPF took over government in Kigali.

3.4 The international humanitarian response to the Rwanda crisis, and the critique

The largest-ever international humanitarian operation was launched in response to the Rwanda crisis, specifically to the refugee movements outside the country.¹⁸ It cost an estimated US\$1.4 billion between April and December 1994 alone, and involved about 250 NGOs (Borton et al., 1996). It was also one of the most high-profile humanitarian response operations. The Public Information Officer for the United Nations High Commission for Refugees (UNHCR) in Goma estimated that there were around 500 journalists and media technicians in the Goma area by the end of July 1994 (quoted in Borton et al., 1996). This high-profile media coverage undoubtedly fuelled the massive response by both governments and the public around the world.

Overall, according to Borton et al. (1996: 10):

‘the response contained many highly commendable efforts... [including] the courage and commitment shown by UN, ICRC and NGO personnel in extremely difficult and often dangerous situations.... Given the magnitude and scale of the population movements and the distance of the beneficiary populations from coastal ports, this was a substantial achievement’.

Despite this huge investment of resources, the failure to prevent the highest mortality rates ever recorded in a refugee population – in Goma in late July – haunted many organisations and their staff for years to come. The Rwanda evaluation concluded that:

‘there was frequently an imbalance in resource provision between preparedness and capacity increasing measures, on the one hand, and response measures in the face of a pressing humanitarian need on the other, particularly where such needs were well covered by the media’ (Borton et al., 1996: 157).

¹⁷ Most of this section and the following section on the international humanitarian response is based on Study 3 of the Rwanda evaluation (Borton et al., 1996), discussed at greater length in section 3.5.

¹⁸ It should be made clear, however, that failure to prevent the genocide, which claimed most lives, was a political, diplomatic and military failure, rather than a failure of the humanitarian operation.

After this initial catastrophe, when mortality rates were eventually brought under control, agencies were still faced with the massive task of servicing the huge refugee camps along the Rwanda/Zaire border in inhospitable volcanic terrain. Numerous agencies were involved, working under the media spotlight. Faultlines that already existed in the international humanitarian community were magnified dramatically in this hothouse of a relief operation. For example, the divergence of standards and experience between NGOs quickly became apparent, as much to the agencies themselves as to external observers and critics. According to the Rwanda evaluation:

‘Whilst many NGOs performed impressively, providing a high quality of care and services, 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’ (ibid.: 161).

The critique of humanitarian agencies, mainly focused on NGOs, was picked up vociferously by the media, particularly in Britain. The critique focused not only on poor performance of NGOs, but also, and more controversially, on the appropriate role and mandate of international NGOs.¹⁹ They were accused of indiscriminately feeding the killers, the ‘genocidaires’, and thus fuelling violence and conflict in the region. The debate was heated, widespread and often soul-searching. It took place in a number of different fora:

- within and amongst NGOs, where there was a lot of questioning and self-criticism
- in the media, as mentioned above
- within and amongst donor and other (e.g. UN) agencies, partly triggering the decision to launch the Joint Evaluation of Emergency Assistance to Rwanda.

Key issues raised in the debate, and relevant to this case study, included:

- the poor performance of some NGOs and general lack of professionalism
- the proliferation of NGOs and lack of regulation
- appropriate demarcation of responsibility and roles between humanitarian agencies and political actors, i.e. who was responsible for separating the genocidaires from the bona fide refugees
- the rights of those affected by the emergency (although this was articulated explicitly by only a few people, it was implicit in much of the ongoing debate)

3.5 Rwanda evaluation

In October 1994, six months after the genocide began, the head of Danida’s Evaluation Secretariat (Niels Dabelstein) presented a proposal to the Development Assistance Committee’s (DAC) Expert Group on Aid Evaluation to carry out an evaluation of the international response to the Rwanda crisis. Although a number of members of the group were enthusiastic, others were not. No consensus decision was reached. However, Dabelstein continued to pursue the idea, initially in an unofficial meeting after the DAC Expert Group meeting, and subsequently, a month later in Copenhagen, with a much broader group of agencies, including bilateral donors, UN agencies, international organisations and NGOs. All had a stake in the response to the Rwanda crisis and thus in the evaluation. An ambitious, multi-stakeholder process was the result.

A comprehensive structure for the evaluation was put in place, which proved to be very effective, as is discussed in section 4. A Steering Committee was established, representing 38 agencies and organisations.²⁰ NGOs were represented by umbrella organisations such as InterAction, SCHR

¹⁹ See for example African Rights (1994).

²⁰ Australia, Austria, Belgium, Canada, Denmark, France (until it left in December 1995), Finland, Germany, Ireland, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Spain, Sweden, Switzerland, UK, USA, Commission of the EU, OECD/DAC,

and Voice, and a Management Group was established, comprising members of the evaluation departments of five donor agencies (from Sweden, Norway, Denmark, UK and the US). The chair of the Steering Committee (Niels Dabelstein) was also the chair of the evaluation's Management Group.

The members of the evaluation Management Group played a key role in drafting the terms of reference for the evaluation, which were broadened out at an early stage from a narrow (and conventional) focus on efficiency and effectiveness to include, explicitly, political and conflict-management issues. In the words of Niels Dabelstein, 'The terms of reference were comprehensive, and the special interests of the agencies were largely accommodated' (Dabelstein, 1996: 4).

The speed at which this international multi-stakeholder evaluation got off the ground is remarkable. In December 1994, the Steering Committee of the Joint Evaluation met for the first time in Washington, and approved the final terms of reference. By January 1995 the evaluation teams had been appointed and started work.

ODI was awarded the contract for Study 3 of the evaluation, the largest study, on 'Humanitarian Aid and Effects'. A 20-person team was put together, led by John Borton. The 'usual' evaluation approach was used in gathering information, albeit on a huge scale. For example, the database of documentation collected for the evaluation eventually exceeded over 2,500 catalogued items, and a total of 620 individuals were interviewed. 1,122 days' work (or four person years) were put into the evaluation by the team members. Two field trips were made to the region: a 20-day reconnaissance mission, and then substantial fieldwork over a two-month period.

The first draft of Study 3's findings was submitted to the evaluation's Management Group, and subsequently to the Steering Committee, in mid October 1995. Over 100 pages of comments were received and responded to. The final report of the Rwanda evaluation was published in March 1996 (Borton et al., 1996).

Study 3 made two sets of recommendations specifically targeted at NGOs. The first set of recommendations was to do with improving NGO performance. Having noted the unprofessional, and at times dangerous, performance of some NGOs, Study 3 commended the NGO initiative to develop a set of standards (the embryonic Sphere project), and recommended that there be some form of regulation or enforcement, either through self-managed regulation, or through an international accreditation system. It also recommended that donor organisations give greater support to NGO emergency training and lesson-learning activities. The second, and related set of recommendations was to do with improving accountability. The availability and quality of performance data and reporting by NGOs (and other agencies) were found to be highly variable, and often incomparable. This gave further weight to the recommendation that there be some means of enforcing standards and codes of conduct. Alternatively it was recommended that a unit within the UN's Department for Humanitarian Affairs (UN DHA) should play a monitoring or even an ombudsman function, or an independent organisation be appointed as ombudsman (Borton et al., 1996).

After such a momentous evaluation of the response to such a major crisis, a formal process of follow-up seemed important to the team leaders of the different evaluation studies, and to the evaluation Management Group. It was agreed that the Steering Committee of the Rwanda evaluation should meet again, one year after publication, to review its impact. So a formal follow-up group, known as JEFF (Joint Evaluation Follow-up, Monitoring and Facilitation Network) was established in May 1996, comprising 11 members drawn from the Management Group, from the

study team leaders and from the Steering Committee. JEFF's principal objective was to monitor the discussion and follow-up action on the Rwanda evaluation, in as many fora as possible. It was also involved in disseminating the findings and recommendations from the Joint Evaluation. Its members made over 20 separate formal presentations, including at a three-day conference in Kigali. JEFF's findings were presented to the Steering Committee in early 1997. They focussed principally on the extent to which the evaluation's 64 recommendations, contained in the synthesis report (Eriksson, 1996), had been taken up. JEFF's final report was printed in June 1997, although it was not formally published (JEFF, 1997). The report noted the SCHR/InterAction initiative to establish 'a major project to develop a set of common minimum standards for humanitarian response, based on the rights of beneficiaries' (ibid.: 39). It awarded category 'D' to the response to the recommendation of 'Self-regulation through improved agency coordination and standards within NGO community'. Category D meant that the recommendation had been formally discussed and resolution reached and/or action taken.

3.6 The beginnings of Sphere

Sphere's origins date from before the publication of the Rwanda evaluation findings, stemming from two different initiatives on either side of the Atlantic.

In June 1995, InterAction – the umbrella body for US development and humanitarian NGOs – submitted a proposal to the US Office of Foreign Disaster Assistance (OFDA) for the 'Development of Training for PVOs in Complex Emergencies', specifically to improve their response capability in the wake of the Rwanda crisis. Three months later, InterAction held a PVO (Private Voluntary Organisation) conference on Disaster Response. Commitment to best practice and defining minimum standards for operational response were both seen as important ways of improving operational effectiveness. InterAction subsequently secured funding from the Ford Foundation to develop an NGO Field Cooperation Protocol, and to develop best practice for disaster work. This was seen as 'an integral part of our community's wider effort to enhance performance and professionalism in disaster response'.²¹

A few months later, on the European side of the Atlantic, Peter Walker of IFRC, and Nick Stockton, the new Oxfam Emergencies Coordinator, had come together in various meetings in the wake of the Rwanda crisis, and had agreed that something needed to be done to improve the quality and accountability of NGO emergency work. Both individuals were also members of the Steering Committee of the Joint Rwanda Evaluation, representing IFRC and SCHR respectively. In February 1996 they submitted a proposal to SCHR for a project entitled 'Towards Quality and Accountability Standards in Humanitarian Relief'. They proposed establishing a set of technical standards in four sectors, based on the rights of those affected by emergencies.

Shortly afterwards, in April 1996, discussions began between InterAction and the authors of the SCHR proposal. In early October 1996, SCHR hosted a meeting in Geneva to discuss all ongoing work on best practice. This was attended by representatives of all the main NGO umbrella organisations: InterAction; the International Council of Voluntary Agencies (ICVA); Voluntary Organisations in Cooperation in Emergencies (VOICE); Action by Churches Together (ACT); and by representatives of a number of individual NGOs such as Medecins Sans Frontieres (MSF) and CARE International. This was the critical meeting which brought together the two initiatives that had started on either side of the Atlantic. It resulted in an agreement to consolidate all work on standards and best practice within the framework of the SCHR project, which was now called 'Quality and Accountability Standards in Humanitarian Relief'.²² At this point the governance

²¹ Quote taken from a memo from InterAction to Peter Walker and Karen Donovan in February 1996.

²² After this meeting, InterAction canvassed the views of its members, who agreed to this collaborative action.

structure for the project was agreed upon, which incorporated two InterAction members into the Project Committee.

Also in October 1996 fund-raising for Sphere began in earnest. Recruitment for the first project manager also commenced. Susan Purdin was appointed and took up her post in Geneva in July 1997. When she joined, the title of the ‘Sphere project’ was coined, and the first one-year project officially began.

3.7 Sphere gets off the ground

The Sphere project came into being very quickly. It is useful to trace how this happened, and in particular how buy-in was achieved.

Having recently experienced the inclusive process of the Joint Rwanda Evaluation, Peter Walker and Nick Stockton proposed an inclusive process for managing (and for funding) Sphere. The Project Committee (as it was then called) initially comprised all SCHR member agencies. When this became a collaborative venture with InterAction, two InterAction members joined the committee. ICRC, ICVA and VOICE all became observer members (until 2001 when ICRC joined SCHR). Thus, all the main NGO umbrella organisations were involved. A number of UN agencies indicated their interest in, and desire to get involved in Sphere, but a policy decision was made in early 1997 to keep this as an NGO initiative, although UN agencies did contribute to the drafting of the minimum standards.

With the arrival of the project manager, Phase 1 of Sphere got off the ground. Four SCHR agencies and one InterAction agency were identified to lead each of the five sectors. A staff member from each of those agencies was seconded to Sphere for a period of six months, to lead a working group of sectoral experts and to put together a first draft of the proposed minimum standards. This was also a very inclusive process, involving more than 700 people and over 200 agencies. For most of the sectors there was a small core group of experts who provided most input. This was supplemented by e-mail discussion, open to anyone, which enabled large numbers to participate, albeit sometimes very briefly. There were regular meetings of all the sector managers, appropriately held in different cities in Europe or the US, and very regular communication and updates between the project manager and the respective sector managers. The short time-period allocated for drafting the sectoral standards put a lot of pressure on the sector managers, but also meant that Sphere took shape very quickly, and could not get bogged down in too much detail.

A few months later, in early 1998, a Working Group was established to draft the Humanitarian Charter, revising an earlier draft prepared by the Sphere project manager in August 1997.²³ This addressed the rights-based component of Sphere. In the wake of the Rwanda crisis, how could a stronger basis for the ‘right to life with dignity’ be established? Also mindful of the recent Rwanda experience, the distinction between combatants and non-combatants was re-stated, as well as the principle of non-refoulement. This was the more controversial part of Sphere, particularly with InterAction representing US NGOs, who feared the litigation implications of a rights-based charter, and therefore preferred to focus on the technical standards to improve quality.²⁴ Oxfam, IFRC, ICRC, InterAction and Save the Children Fund-UK (SCF-UK) were involved in drafting the charter. After much discussion, it was finally completed and approved by the Sphere Management Committee in May 1998. But this was too late to feed into, or to influence the minimum standards

²³ The concept of a charter seems to have been at least partially influenced by developments in the commercial and public sector at the time. In the UK, for example, charters were being drafted in different sectors to define the rights of consumers or service-users.

²⁴ It is also interesting to note that the US State Department initially opposed Sphere ‘as an attempt by NGOs to create international law’ (personal communication), until they were persuaded that this was an attempt to implement existing international law, at which point it became a strong supporter and funder of Sphere.

part of the project; editing had begun in April on the first draft of the minimum standards document. Also, conceptual understanding of what a rights-based approach meant in practice, and therefore its implications for setting technical standards was rather unclear. The debate about a rights-based approach was still in its infancy at this stage and five years later the sector is still grappling with many of these issues.

4 Explanation

4.1 The policy context: politics and institutions

Political, social and economic structures and interests

In the first half of the 1990s, the policy context was becoming ever more conducive to an initiative such as Sphere. As described in section 3.2, humanitarian agencies were faced with increasingly challenging operational issues – for example, how to operate when there was no state, how to operate in close conjunction with the military, how to ensure humanitarian aid did not fuel conflict. As the analysis and discourse became increasingly politicised and sophisticated, the days of unquestioning acceptance of the ‘good work’ of humanitarian agencies were over. The sense of discomfort felt by some agencies and individuals about the range of standards and performance to which different agencies operated, and the moral and political dilemmas of providing humanitarian assistance within war zones were becoming more evident and familiar as more agencies entered the fray.

Although impossible to prove, it seems unlikely that these concerns alone, in the early 1990s, would have been strong enough to trigger an initiative on the scale of the Sphere project. Instead, it was the unprecedented scale and intensity of the Rwanda refugee crisis, and the unprecedented international humanitarian response, under the international media spotlight, that magnified the faultlines within the humanitarian system to an extraordinary degree, and did so very publicly. That was a major factor in creating the momentum for change, and an openness to radical change. That momentum was particularly strong within the NGO community – the most visible part of the much-criticised international humanitarian system. Indeed, in the words of one of the members of the Joint Rwanda Evaluation team, ‘the donors got off lightly’. A widely-accepted and oft-repeated message reverberating within the NGO community was the mantra: ‘we have to get our act together’. This was reinforced by some donor representatives in strong messages such as: ‘if you don’t do it, we (i.e. the donors) will’. And this message was forcefully communicated in a number of meetings following publication of the Joint Rwanda Evaluation. The Rwanda evaluation was seen by many in the NGO sector to have opened up the possibility of external regulation.²⁵

In short, there were strong push factors to do something to improve NGO performance. And a few individuals (and agencies) capitalised upon that in order to get their ideas accepted and eventually to launch the Sphere initiative.

Assumptions and attitudes

Most humanitarian actors involved in the response to the Rwanda crisis shared a common analysis that performance by some NGOs had been poor, that lives had been lost as a result, and that something had to be done to raise the quality of humanitarian (in particular relief) assistance. Thereafter, the analysis started to diverge. For example, was it a problem of too many NGOs, of poor coordination, of lack of regulation? Hence, there was some divergence of view about an appropriate solution, and in the words of some commentators, there was a lack of vision about what to do. However, when performance standards were proposed, it transpired that the majority of NGOs (on both sides of the Atlantic) were in favour.

²⁵ See, for example, a note on the InterAction Annual Forum in April 1996.

Indeed, within SCHR, the two initiators of Sphere, who knew how critical the forthcoming Rwanda evaluation was going to be about NGO performance, used the threat of heavy-handed donor action to generate support within SCHR for their quality and accountability proposal. Similarly, within InterAction the threat of external regulation provided a driver to develop the 'Field Cooperation Protocol'.

There was some concern early on that setting standards would disadvantage or discriminate against small agencies, in favour of the larger more established ones. This concern was articulated on both sides of the Atlantic.

The more significant objections to the ideas being put forward in the Sphere initiative appear to fall into two broad camps. First, there were those who thought that what was being proposed was not do-able (e.g. SCF-UK). This debate happened early on, in 1996. It particularly revolved around the implementation of Sphere and regulation. There were also concerns about the potential cost of the initiative – a concern also of the Lutheran World Federation (LWF). However, both these sets of concerns abated – the former as it became clear that regulation would not be incorporated into Sphere, and the latter when it became clear that donors were prepared to fund the initiative.

The second and more fundamental opposition came from those agencies that were opposed ideologically to the way that Sphere developed. This form of opposition is most closely associated with MSF and with a number of other Francophone NGOs, and has been most evident at headquarters level.²⁶ The arguments have been articulated with different emphasis by different individuals, but a central cause for concern has been the apparent isolation of technical standards from humanitarian principles in what was regarded, uncomfortably by these agencies, as an institutionalisation of the Sphere process. When MSF initially perceived Sphere to be a relatively low-key exercise to share technical guidelines and policies between different operational agencies, they were prepared to participate (albeit with reservations about what such an exercise would achieve). However, as the process gathered momentum, and the rather short and minimalist (in MSF's eyes) Humanitarian Charter was added, MSF became increasingly concerned that a highly technical approach to humanitarian action was being promoted, in a high profile way, at the expense of the more difficult and pressing political issues.²⁷

'In attempting to achieve minimum standards for humanitarian action, there is a risk that humanitarian action may simply become a technical and purely professional pursuit' (Orbinski, 1998: 2).

MSF and some of the other agencies also had concerns about how Sphere would be used, for instance by donor governments in evaluations. If donors insisted that their implementing partners adhere to Sphere standards, would their partners be locked into ever-closer relationships with donor governments that may be pursuing political rather than humanitarian objectives, thus compromising the independence of their partner NGOs (Terry, 2000)? And by focusing on NGOs, Sphere was in danger of deflecting attention away from where it was really needed – with ruling authorities and the international community who are ultimately, and formally responsible for providing and protecting humanitarian space (Orbinski, 1998). A further criticism leveled at Sphere was its failure to address issues of protection.

This more fundamental opposition did not really surface until 1998, a few months before the Sphere project was launched, when a number of Francophone agencies expressed their disquiet and concerns in what became famously known as the 'French letter'. For a short while it threatened to derail the Sphere process. Should the launch go ahead as planned? How could MSF and other

²⁶ According to the Sphere project staff, opposition at field level has not been so strong, and sometimes retracted.

²⁷ Personal communication, Jean-Marie Kindermans.

significant humanitarian agencies be kept on board? But eventually the decision was taken to go ahead. Although there were attempts to keep these agencies on board, eventually the differences became too great, and MSF withdrew from the Sphere Management Committee. Thereafter, positions became polarised and the debate has, at times, been acrimonious.

Although Study 3 of the Rwanda evaluation commended and supported the development of standards, it went further than any NGOs did, at the time or since, in recommending some form of self-managed regulation or accreditation of NGOs to monitor compliance. NGOs have consistently shied away from this, whether on the Sphere Management Committee or outside.²⁸ There has been an inherent reluctance to go down this route, mainly through fear of the implications. To the frustration of some, the push factors that resulted in Sphere did not seem to have been strong enough for this extra step to have been taken. However, during the period of research and writing of this case study, there has been an important development which may signal a shift. The Humanitarian Accountability Project (HAP), which evolved out of the Humanitarian Ombudsman project, has just been established as a self-regulatory body with the specific objective of accrediting its members, particularly in relation to accountability to beneficiaries. The full impact and implications remain to be seen.

4.2 Evidence

The researchers and their methodology

The team that carried out Study 3 of the Joint Rwanda Evaluation was generally well-respected. A number of the team members had a research background and came from research institutes, and were thus seen as more objective. Others had strong operational experience, or a combination of both. Many of them were well-established in their respective fields. This gave the team, and thus their findings, credibility.

As detailed in section 3.5, the team adopted a fairly thorough approach to the evaluation, combining an extensive documentation review with numerous interviews in the field and at agency headquarters. Indeed, the Study 3 team leader observed that:

‘Study 3 was a veritable Rolls Royce of humanitarian evaluation with unprecedented scope and unprecedented resources available to it, and this made it a privileged and professionally extremely rewarding process... [It] meant we were able to afford what other evaluation teams would regard as luxuries: a large multi-disciplinary team of very high-calibre individuals; an approach that at times bore a closer resemblance to research than to conventional consultancy... and a funded follow-up process that enabled some of those involved in the team to ‘sell’ the report and its key messages within the humanitarian system as well as to monitor and comment upon the reactions to the report and its recommendations by the principal humanitarian agencies’ (Borton, 2001: 99).

The fact that this evaluation was more akin to a research project than to conventional consultancy contributed to the credibility of the final product. The combination of thoroughness and pragmatism was commended by the Chair of the evaluation Steering Committee. In his view, the team did a good job in presenting their findings in a practical and accessible form for policy-makers, using research methods but without a lot of theorising.²⁹ This was attributed to a combination of guidance from the Rwanda Evaluation Management Group about how best to present the findings, and because many of the evaluators had a field background and therefore a more practical orientation.

²⁸ Any efforts to develop accreditation have been very specific, for example, related to child sponsorship within InterAction.

²⁹ Personal communication.

However, the team members were mostly European, along with two North Americans. There were no African members – ‘the cause of subsequent criticism of the team and the source of some embarrassment to members of the team’ (Borton, 2001: 85). At short notice, they had difficulty identifying African researchers who were familiar with the Great Lakes region, yet could also be rigorously objective in that highly polarised context. This was hampered by the fact that neither the Management Group nor the evaluation team leader had an existing network of African emergency evaluators. Evaluation had tended to be somewhat ‘northern’-dominated.

Credibility of the evidence

Generally, the findings of the Study 3 evaluation team were regarded as highly credible, for a number of reasons. First, the report clearly laid out and analysed what most humanitarian agencies already knew to be the case, but it did so in ‘a cooler-headed (way) than was happening elsewhere’.³⁰ Second, it was regarded as ‘categorical, clear and well-documented’,³¹ partly related to the research approach that had been adopted. Crucially, at a time of heated debate and public criticism of aid agencies, the evaluation provided an independent and objective analysis.

Nevertheless, there were organisations and agencies who disputed the findings of Study 3 (and other reports) of the Evaluation, but they were mostly agencies that were heavily criticised in the report. For example, the French government pulled out of the Evaluation’s Steering Committee in protest at the criticism of the French government. And UNHCR vociferously disputed criticism of their preparedness for the refugee influx to Goma. This was one occasion where the objectivity of the evaluation team member who had prepared most of the evidence against UNHCR came into question. A lengthy process of dialogue and consultation between the evaluation team leader and UNHCR ensued, not entirely satisfactorily in the view of the team leader (Borton, 2001). It was a small minority of agencies who questioned the credibility of the evidence.

Despite this, the structure for managing and directing the evaluation proved critical to safeguarding the independence of the findings of the Study 3 team. This was because of the role of the evaluation Management Group; most of the agencies being evaluated were represented on the evaluation Steering Committee and therefore the Management Group was a very important intermediary and buffer, ‘in several instances shield[ing] the teams from undue pressure from the agencies and governments’ (ibid.: 100). For example, some members of the Steering Committee (the UN Secretariat, UNHCR and the Government of France) wanted parts of the reports that were highly critical to be removed or edited. ‘The Management Group resisted this pressure by expressing full confidence in the quality of the work and continued to ensure the teams’ independence’ (Dabelstein, 1996: 291). The Management Group also played an important quality control function, demanding rigorous analysis from the evaluation teams, and supporting evidence. Borton comments that:

‘The fact that the members of the Management Group were all heads of evaluation departments in bilateral donor organisations was helpful in this regard, for they combined the status of being within donor organisations with a strong tradition of objectivity’ (Borton, 2001: 100).

Impact

The Joint Rwanda Evaluation had a very big impact on the humanitarian aid sector, immediately and in the following years, for a number of reasons. First, the failure to prevent a crisis of such scale and intensity shocked the many actors involved and prompted major debate. As a result, there was an openness amongst most actors to learn the lessons (although also some defensiveness,

³⁰ Personal communication, Nick Stockton.

³¹ Personal communication, Peter Walker.

alluded to above). The scale and process of the Joint Rwanda Evaluation was somehow proportionate to the crisis. In the view of the Study 3 team leader, the evaluators were being encouraged to ‘think big’ because of the widespread view that things had to change. This was an opportunity to think radically when making recommendations.

The Joint Rwanda Evaluation was also very high profile – ‘it was almost front page news’ in the words of one interviewee. It was taken seriously and read by many, and so agencies had to be seen to be responding.

Second, the process of the Joint Rwanda Evaluation seems to have played a key role in influencing policy-makers, partly because they were heavily involved throughout, and also because discussions were ongoing – with the evaluators and amongst themselves – especially once the draft evaluation reports were available. Involving so many different actors in the Steering Committee – UN agencies, donor government agencies, NGOs etc – meant that system-wide discussion was going on, probably to a greater extent than ever before. This seems to have added momentum to the taking forward of recommendations – for example, some donor representatives threatening to take action if the NGOs did not ‘get their act together’.

The findings and recommendations in the evaluation were clearly presented, and were often targeted specifically at particular groups of actors – for example, donor agencies or NGOs. This was backed up by the JEFF process, whereby members of the evaluation were able to ‘sell’ the recommendations to different audiences.

However, this does not mean that all the recommendations were taken on board. On the issue of improved performance and accountability, one of the proposed options that the UN DHA serve as ombudsman to improve the accountability of humanitarian agencies was ‘laughed out of court’ because it was regarded as not do-able. Although Study 3 gave real impetus to the early Sphere initiative, its recommendation of some form of regulation or enforcement to ensure improvement in NGO performance has never yet seen the light of day. The Sphere project document refers to ‘the Humanitarian Charter and the Minimum Standards contributing to an operational framework for accountability in humanitarian assistance efforts’ (Sphere, 2000: 6). In reality however this has been very limited, partly because of concerns about its do-ability, but mainly because of fear of going this far. This is also the rock on which the original concept of the Ombudsman project foundered.

4.3 Links: influence and legitimacy

Key actors and individuals

There are a number of key actors in this story about how Sphere came into being. First, it is worth drawing attention to the way that the Joint Rwanda Evaluation was carried out and to some key features, as these contributed to the status and legitimacy of the final findings and recommendations.

- The fact that the DAC Expert Group on Aid Evaluation was unable to agree, by consensus, to take on the evaluation, fortuitously meant that it became a truly international exercise, with buy-in from a much broader range of agencies. ‘Once the critical mass was established no one could stay out – it could be embarrassing for any one agency not to participate’ (Dabelstein, 1996: 4).
- The evaluation was effectively commissioned by the respective evaluation departments of different donor agencies, represented on the Rwanda Evaluation Management Group, but this was done with the support and input of policy-makers who were represented on the evaluation’s Steering Committee. Thus, a critical and cooperative link was established right at the beginning between those who commissioned the ‘research’, and the policy-makers at whom the findings were directed.

- Finally, the vision and drive of the Head of Danida's Evaluation Secretariat was pivotal in taking forward such an ambitious project. The lack of a system-wide evaluation since has been remarked upon a number of times in ALNAP. It is partly attributed to the enormous management task which no individual (or donor department) is prepared to shoulder. It may also be due to a wariness of what such a major evaluation entails for individual agencies, in terms of the very public scrutiny to which they are likely to be subjected.

Second, within the NGO community, umbrella organisations played a critical role in turning concerns about NGO performance into action – in particular, InterAction in the US, and SCHR in Europe. They provided a forum for inter-agency debate and for inter-agency initiatives, at a time when a system-wide response was clearly required. In some ways, this marked a 'coming of age' of these umbrella organisations. InterAction in the US was probably more established than equivalent organisations in Europe, and took the first steps towards an inter-agency initiative to identify best practice. By coming together, SCHR and InterAction provided the opportunity (and legitimacy) to launch a global policy initiative, with which other umbrella organisations quickly became associated, for example ICVA and VOICE.

Third, a constructive axis existed between IFRC and Oxfam. Although this depended very much on the work of particular individuals (both in the case of the Code of Conduct in 1994, and in the case of Sphere), the fact that these individuals represented two very different but respected agencies was significant. Whilst Oxfam was a UK-based NGO with a strong record in emergency work, IFRC was more international, with a legitimacy based upon the network of Red Cross and Red Crescent societies around the world.

As far as individuals are concerned, the two key individuals who drove Sphere in its early days were Peter Walker (of IFRC) and Nick Stockton (of OXFAM). They both played a key role in the development and implementation of Sphere, and in promoting ownership and buy-in within the NGO community. According to many interviewees, Sphere would not have happened without them. What seems particularly important was the complementarity of their skills and approach. Whilst one had a strong vision of what Sphere could be and was able to put a strong intellectual case for the centrality of rights and the Humanitarian Charter, the other, whilst sharing the vision, was particularly adept at understanding the politics of the international humanitarian system and therefore piloting the idea of Sphere through the rapids of organisational politics. Both were articulate and good presenters: they were able to make a convincing case to colleagues from other agencies. It is also significant that these two individuals were given the 'space' by their respective agencies to develop and run with this idea.

The role of others, such as Jim Bishop of InterAction, were also important in terms of facilitating and encouraging buy-in, in this case amongst US NGOs.

Other key individuals were members of the Study 3 evaluation team, and in particular John Borton, the team leader. They are credited with having done a very thorough job, making clear and in some cases quite radical recommendations for what needed to change in the international humanitarian system.

When the Sphere project finally did get underway, the role of the first Sphere project manager (Susan Purdin) was critical in guiding and managing the process of putting the standards and Charter together. Many of the terms and parameters had still to be defined – for example, what is an emergency, how to differentiate between standards and indicators. Susan is credited by many for holding together this inclusive but challenging and time-pressured process.

Roles and Relationships

It is not just the role of the individuals that is important, it is the way in which they interacted that resulted in this system-wide initiative. As previously mentioned, the process of the Joint Rwanda Evaluation was very important in bringing together the evaluators/researchers, policy-makers and practitioners. Also, the key individuals mentioned in the previous section were all working to a very similar agenda. Thus, the idea for the Sphere project (although not the name) was already being discussed before the evaluation report was published. The evaluators were aware of this, and made a point of supporting this initiative as a step towards improving NGO performance; this is reflected in their final text. Meanwhile, both Nick Stockton and Peter Walker, as members of the evaluation Steering Committee, knew from draft reports how critical the evaluation would be of NGO performance, and were able to use this knowledge to persuade NGOs to sign up to the ‘Quality and Accountability Standards’ project. There was similar pressure within InterAction. Thus, there was a clear synergy of interest and of drivers.

It is interesting to note, however, that efforts to build on the positive experience of the Rwanda evaluation, by establishing an ‘international standing capacity’ to ‘provide objective and critical analysis, monitoring, review, reporting and dialogue on the international community’s progress in preventing, mitigating and resolving violent conflict’³² did not come to fruition. Although some of those who had been involved in the evaluation developed the idea, there was not sufficient backing for it to get off the ground.

The experience of the evaluation process had a direct influence on the process that was put in place to manage and take forward the Sphere project. Achieving agency buy-in was a very specific objective. In the words of the first Sphere project manager: ‘we always said the process was as important as the product’. The formation and regular meeting of the Sphere Management Committee, with its observer members, was one part of this, to some extent modelled on the evaluation process. A Public Relations/advocacy plan for the ‘performance standards project’ was drafted early on, initially targeted at as many NGOs as possible to encourage their participation and support for the project. Also important was the highly consultative process of putting the Sphere standards together, described above. The sector managers played a key role in raising awareness of Sphere in developing countries by running a series of consultative meetings in Africa, Asia and the Balkans. The funding strategy that was put in place for Sphere was again designed to achieve maximum buy-in. All agencies on the Sphere Management Committee took responsibility for following up with particular donors, thus consolidating the sense of ownership, and also presenting an international image of Sphere. The decision to request funding from many different donors was a strategic one, so that no single donor ‘owned’ Sphere, but many had bought into it. The NGOs also contributed funds themselves – literally buying-in.

Legitimacy

The legitimacy and credibility of the evaluators/researchers is described in section 4.2. The key architects and advocates of Sphere – Peter Walker and Nick Stockton – established their legitimacy partly as a result of their senior positions within two large and respected agencies. Also, both had field experience and were respected in the policy arena. When it came to putting the Sphere standards together, it was important that each of the sector managers were ‘experts’ in their field, already tapped into existing sectoral networks,³³ thus reinforcing their legitimacy. However, not all the sector managers saw their role in the same way. Some played more of a facilitator role, some more of an ‘expert’ role.

³² Proposal for an ‘International Standing Capacity to Monitor and Report on International Progress in Addressing the Phenomenon of Complex Emergencies’, by Katherine Blakeslee (USAID), Niels Dabelstein (Danida) and Peter Walker (IFRC), September 1997.

³³ However, the strength of the different networks varies between sectors. For example, it was strongest in the nutrition sector and weakest in the health sector. Shelter also suffered from a weak network, and from a lack of technical research and experience.

5 Discussion

5.1 The context: politics and institutions

In this particular case study, the key policy-makers came from the NGO sector which is not unusual in the humanitarian aid sector. A lot of innovation at the operational end of the spectrum has taken place within NGOs, not within donor organisations (which are rarely operational in humanitarian emergencies), although donor agencies have played a critical funding role. For example, many of the policy initiatives associated with the drive to improve the quality of humanitarian operations have come out of the NGO sector, including Sphere, People in Aid, and the Humanitarian Ombudsman project. Interestingly in the Sphere case, the role of officials within donor agencies was to advocate for change, if necessary using implied threats. However the change they were advocating was mainly external to their own agencies/departments. The inertia commonly associated with large bureaucracies is more evident in relation to change *within* those donor organisations (Clay and Schaffer, 1984). For instance, the concept of ‘good donorship’ has taken another six years to come onto the agenda since these discussions began in the NGO sector.

The researchers/evaluators in this case study defy the ivory-tower image. Most of them also had operational experience, and this seems to have been important in terms of their credibility with NGO policy-makers, and in terms of their ability to communicate with these operationally-oriented policy-makers.

In the immediate aftermath of the Rwanda crisis there was a lot of room for manoeuvre within the humanitarian aid system. Discomfort and disaffection with NGO performance had been around for a few years beforehand, but became much more urgent and widespread as a consequence of the crisis. This seems to be quite typical of the humanitarian system which is most responsive to change under pressure when the push factors are strong. This was never more so than after the Rwanda crisis, with its high-profile media coverage, and criticism. There were simultaneous initiatives for change, on both sides of the Atlantic. At this stage, the discussions and initiatives were mainly ‘northern’-driven, which is again typical of how the humanitarian aid system is structured. Most international humanitarian NGOs have their roots and headquarters in the north – in Europe or North America, and this is where most policy discussion takes place. Southern-based NGOs are represented by umbrella organisations, such as ICVA, LWF, Caritas Internationalis or IFRC. Southern-based international humanitarian agencies, such as African Humanitarian Action (AHA), argue that they do not have the luxury of engaging in this kind of policy discourse, because of lack of resources and therefore being tightly locked into a contract culture with donors and others.³⁴

The strength of the push factors meant that the Sphere policy decision was of the ‘fundamental and emergent’ type (Lindquist, 2001). It goes beyond a questioning of underlying assumptions, to proposing a broad new vision based on rights and professionalism. A number of the NGO policy initiatives to improve quality and accountability fall into this category. This appears to be in contrast with the predominant nature of humanitarian policy change within donor organisations, which is more ‘incremental’ in nature (*ibid.*),³⁵ but this requires further research to be conclusive. However, the story of how Sphere came into being does seem to follow Kuhn’s thesis of a paradigm ‘bumb(ing) along quite happily making sense of the world until the puzzles it cannot solve become so numerous that it causes a crisis’ (as quoted in Crewe and Young, 2002: 9). Crudely, in this case study the paradigm that caused the crisis was of unregulated humanitarian agencies ‘doing good’ and meeting the needs of those affected. In the goldfish bowl of Goma in 1994/95, it became starkly

³⁴ Statement made at an ALNAP meeting by a representative of AHA.

³⁵ See for example, Macrae and Leader’s analysis of the search for coherence between political and humanitarian responses to complex emergencies (Macrae and Leader, 2000).

apparent that this was not the case. Three new paradigms emerged. The first has developed into the rights-based approach, where the starting point is the right of the affected person rather than an agency's assessment of that person's need. This is at the heart of the Humanitarian Charter. The second paradigm is to do with professionalising and improving the quality of performance of NGOs – represented in Sphere as the minimum standards. The third paradigm involves strengthening the accountability of humanitarian agencies, specifically NGOs. In the Rwanda evaluation, this was captured in the recommendation that there be some form of accreditation or regulation of NGOs. Although it has been discussed many times, this is the paradigm that has resulted in least action. This is evidence of the 'street bureaucrats' (i.e. the operational NGOs) ultimately making the decision about which of the researchers' recommendations for policy change they were prepared to accept: designing aspirational standards to improve quality, but not going as far as regulation or accreditation (although there may be some shift in this direction through HAP).

5.2 Evidence: credibility and communication

The Rwanda evaluation fulfilled two criteria regarded as important if research is to influence policy. It was good quality research and widely regarded as such, and it was effectively disseminated through the JEFF process. Many of the reasons why it was good quality research have been described previously in this paper, including the fact that it was a 'rolls-royce' of an evaluation. Also important was the role of the evaluation Management Group as a kind of peer review/quality control mechanism. Another feature of Study 3 that is worth mentioning was its ability to make the connections from the detailed evidence of what worked and what did not work to system-wide analysis and recommendations. Much has been written about the temptation for researchers to simplify their messages in order to have maximum impact. Study 3 put across its recommendations in a concise and simplified way, but they were backed up by a great deal of detail. Nevertheless, it is worth mentioning how one of the evaluation's key messages has been distorted and misinterpreted by policy-makers, in particular donor policy-makers. The message was '...that humanitarian aid cannot be a substitute for political action', making the case for increased coherence between political and humanitarian efforts, so that the mistakes of the response to the Rwanda crisis would not be repeated. It was the absence and ineffectiveness of the political response that had compromised protection of and assistance to the victims. Macrae and Leader (2000) have demonstrated how this evolved into the 'new humanitarianism' where humanitarian and political objectives became integrated, and relief was not only a substitute for political action, but the primary form of political engagement at the geopolitical periphery, in countries such as Afghanistan before September 11th. This is an important reminder of how research results can be used and distorted to support a particular policy position.

The way that the evaluation influenced policy can partly be explained using marketing theory. As described above, in the analysis and in particular in the recommendations there was an effective segmentation of the audience (for example, by UN agencies, NGOs and donors) within this piece of work about the whole system. Particular messages were directed at particular types of agencies. And the 'stickiness' factor (in other words, the findings having an impact) is partly due to Study 3 saying what everyone knew already, but in a much clearer, and cooler-headed way (Kotler et al, 1999).

The process that led to the Sphere project was more of a cumulative 'snowball' effect than the 'whisper' effect identified by Stephens (cited by Berkout and Scoones (1999)). However the snowball effect was less to do with research, and more to do with growing concern in the NGO sector, at an operational level, about the discrepancy of standards, performance and expertise between agencies. This started as snowflakes in the mid-1980s, snowballing in the early 1990s, and becoming a veritable avalanche as a result of the Rwanda crisis. The Rwanda evaluation gave clear

direction about how these concerns could be addressed, supporting and giving impetus to the embryonic Sphere initiative.

As described in section 5.3, the way in which the Rwanda evaluation influenced NGO policy-makers seems to be more to do with the ‘insider model’ described by Saywell and Cotton (1999), due to the close links between evaluators and policy-makers throughout the evaluation process. The fact that the ‘research’ was commissioned and the researchers chosen by the policy-makers, contributed to some degree to their credibility and the trust bestowed upon them. This is different from the way most research is conducted.

5.3 Links: influence and legitimacy

This case study of how Sphere was born, and in particular of the relative influence of the Rwanda evaluation, supports the hypothesis that close links (whether institutional, in terms of regular contact or common interest) between researchers and policy-makers greatly increases the likelihood of research influencing policy. Those links were institutionalised in the structure put in place for the Joint Rwanda Evaluation. Although it may not always be possible to replicate this inclusive structure involving so many stakeholders, there may be elements of it that can be replicated. The buffer role of the evaluation Management Group, to protect the objectivity and independence of the researchers, is an important one to bear in mind when research, which may be controversial, is commissioned.

The researchers/evaluators were really part of the same policy network as the policy-makers, not least because some of the evaluators had actually worked for the same agencies. So what kind of network was this? Using Stone et al.’s typology (2001), it has elements of a number of different networks. To some extent it was a ‘policy community’ and, at least temporarily, this was created through the structure of the Rwanda evaluation, in particular through the Steering Committee. To some extent it was a discourse coalition, whereby both evaluators and key NGO staff shared the same beliefs about how the system – and in particular, NGOs – needed to change. This is part of what has been called the ‘accountability revolution’ in the humanitarian aid sector, spearheaded by NGOs. To some extent it has elements of an advocacy coalition, whereby different people in different parts of the sector shared a similar vision and were prepared to support each other to bring about the change they thought necessary. However, as described previously, it was also a network dominated by expatriates – a source of some embarrassment to the evaluation team, but not untypical of policy discussion within the humanitarian sector, and as Nindi asserts, within the development sector more generally.³⁶

There was another, smaller policy network that came together in the Sphere Management Committee that was responsible for getting the project off the ground. Again, this had elements of a number of different types of policy network. To some extent it was an issue network, bringing together members concerned about improving NGO performance, but differing in their views about how. (These differences really came to the fore towards the end of Phase 1 of Sphere). To some extent it was a discourse coalition and there were also elements of an advocacy coalition.

Particularly noteworthy is the role of the ‘policy entrepreneurs’, especially Peter Walker and Nick Stockton. In Gladwell’s (2000) terms they were both connectors (i.e. networkers) and salesmen (i.e. persuasive, listened to and trusted). They were also innovators and they spotted the policy window to take forward their ideas. Other individuals, such as Jim Bishop of InterAction and Susan Purdin, the first Sphere project manager, also played critical connector roles.

³⁶ As quoted by Crewe and Harrison (1998).

6 Conclusions

This case study shows that the policy impact of Study 3 of the Rwanda evaluation, on finding ways of improving NGO performance, was more or less immediate in terms of the impetus it gave to the embryonic Sphere project. But, the impact was also partial. The more challenging and radical recommendations around accreditation and regulation of NGOs were ducked and have been consistently evaded.

This case study of how Sphere came into being supports all three hypotheses set out in section 2, on which this project is based.

- The **first hypothesis** concerns the political and institutional limits and pressures of policy-makers. As clearly demonstrated, the high profile failure to respond adequately to the Rwanda crisis (in particular, to the refugee crisis), really expanded the boundaries of what policy-makers would accept as necessary change. An unprecedented policy window of opportunity was created, which the evaluation exploited, and which was seized upon by certain policy-makers within the NGO sector.
- The **second hypothesis** is about researchers and policy-makers sharing networks. This was formalised in the structure of the Rwanda evaluation, although the networks also went beyond this temporary structure. In this case study, research was not of the conventional type. Instead, it was conceived of as an evaluation, and was commissioned by the policy-makers, which all contributed to the results being taken seriously and having a major impact.
- The **third hypothesis** is to do with local involvement, credible evidence and effective communication. The credibility of the evidence was high, which was a very important factor for it to have influence. Local involvement is harder to comment upon because it was limited, but effective communication of the findings, through a deliberate strategy of dissemination does seem to have been important.

References

- African Rights (1994) 'Humanitarianism Unbound? Current Dilemmas Facing Multi-Mandate Relief Operations in Political Emergencies' *African Rights Discussion Paper 4*, London: African Rights.
- Berkout, F. and Scoones, I. (1999) 'Knowing how to change: Environmental policy learning and transfer', *Development Research Insights* 30: 1–2, June 1999.
- Borton, J., Brusset, E., Hallam, A., Collins, S., Pottier, J., de Lame, D., Chalinder, A., Shoham, J., Lee, L., Connaughton, R., Ishøyet, T., Telford, J., Goetz, F., Shepherd, G., Hilsum, L., Jones, B., Turton, D., Jackson, L., Shellard, N. and Walsh, A. (1996) *The International Response to Conflict and Genocide: Lessons from the Rwanda Experience. Joint Evaluation of Emergency Assistance to Rwanda. Study 3. Humanitarian Aid and Effects*. Copenhagen: Steering Committee of the Joint Evaluation of Emergency Assistance to Rwanda.
- Borton, J. (2001) 'Doing Study 3 of the Joint Evaluation of Emergency Assistance to Rwanda: The Team Leader's Perspective', in Wood, A., Apthorpe, R. and Borton, J. (eds.) *Evaluating International Humanitarian Action. Reflections from Practitioners*. London, New York: Zed Books.
- Clay, E.J. and Schaffer, B.B. (1984) *Room for Manoeuvre: An Exploration of Public Policy in Agricultural and Rural Development*. London: Heinemann.
- Crewe, E. and Harrison, E. (1998) *Whose Development? An Ethnography of Aid*. London: Zed Press.
- Crewe, E. and Young, J. (2002) 'Bridging Research and Policy. Context, Evidence and Links', *ODI Working Paper 173*. London: Overseas Development Institute.
- Dabelstein, N. (1996) 'Evaluating the International Humanitarian System: Rationale, Process and Management of the Joint Evaluation of the International Response to the Rwanda Genocide', *Disasters Journal*, 20(4): 286–294.
- De Vibe, M., Hovland, I. and Young, J. (2002) 'Bridging Research and Policy: An Annotated Bibliography', *ODI Working Paper 174*. London: Overseas Development Institute.
- De Waal, A. (1997) *Famine Crimes. Politics and the Disaster Relief Industry in Africa*. Indiana: James Currey in association with African Rights.
- Duffield (1994) 'Complex Emergencies and the Crisis of Developmentalism', *IDS Bulletin*, 25(4). Sussex: Institute of Development Studies.
- Eriksson, J. (1996) *The International Response to Conflict and Genocide: Lessons from the Rwanda Experience. Synthesis Report*. Copenhagen: Steering Committee of the Joint Evaluation of Emergency Assistance to Rwanda.
- Freih, L. (2000) *Increasing Accountability Among Humanitarian NGOs In Disaster Response. A Study for the Sphere Project Management Committee*. Geneva: Sphere Project.
- Gladwell, M. (2000) *The Tipping Point: How Little Things Can Make a Big Difference*. London: Little, Brown & Company.
- JEFF (1997) *The Joint Evaluation of Emergency Assistance to Rwanda: A Review of Follow-up and Impact Fifteen Months after Publication*. Report by JEFF in conformity with a decision by the Joint Evaluation Steering Committee at its final meeting on 14 February 1997 in Copenhagen, London: Overseas Development Institute.
- Keen, D. (1994) *The Benefits of Famine: A Political Economy of Famine in South-Western Sudan, 1983–1989*. Princeton: Princeton University Press.
- Kotler, P., Armstrong, G., Saunders, J. and Wong, V. (1999) *Principles of Marketing* (second edition). Europe: Prentice Hall.
- Kuhn, T.S. (1970) *The Structure of Scientific Revolutions*. Chicago: Chicago Press.
- Leader, N. (1999) 'Codes of Conduct: who needs them?', *Relief and Rehabilitation Network Newsletter* 13. London: Overseas Development Institute.

- Lindquist, E.A. (2001) *Discerning Policy Influence: Framework for a Strategic Evaluation of IDRC-Supported Research*. Canada: International Development Research Centre (IDRC). Available on their website (www.idrc.ca).
- Macrae, J. (ed.) (2002) 'The New Humanitarianisms: a Review of Trends in Global Humanitarian Action', *HPG Report* 11. London: Overseas Development Institute.
- Macrae, J. and Leader, N. (2000) 'Shifting Sands: The search for 'coherence' between political and humanitarian responses to complex emergencies', *HPG Report* 8. London: Overseas Development Institute.
- Minear, L. (1988) *Helping People in an Age of Conflict: Toward a New Professionalism in US Voluntary Humanitarian Assistance*, New York and Washington DC: InterAction.
- Minear, L. and Weiss, T. (1993) *Humanitarian Action in Times of War. A Handbook for Practitioners*. Boulder and London: Lynne Rienner Publishers.
- Orbinski, J. (1998) 'On the Meaning of the Sphere Standards to States and Other Humanitarian Actors', speech delivered to Sphere launch, London, December 3.
- People in Aid (1997) 'Code of Best Practice in the Management and Support of Aid Personnel', first published by *Relief and Rehabilitation Network*. London: Overseas Development Institute.
- RRN (1994) 'Code of Conduct for the International Red Cross and Red Crescent Movement and NGOs in Disaster Relief', *Relief and Rehabilitation Network Paper* 4. London: Overseas Development Institute.
- Saywell, D. and Cotton, A. (1999) *Spreading the Word. Practical guidelines for research dissemination strategies*. Loughborough University: Water, Engineering and Development Centre.
- Sphere (2000) *The Sphere Project. Humanitarian Charter and Minimum Standards in Disaster Response*. Oxford: Oxfam Publishing.
- Terry, F. (2000) 'The Limits and Risks of Regulation Mechanisms for Humanitarian Action', *Humanitarian Exchange* 17. London: Overseas Development Institute.
- Stone, D. with Maxwell, S. and Keating, M. (2001) 'Bridging Research and Policy'. An International Workshop funded by DFID, Warwick University, 16–17 July.
- Woodward, S. (2001) 'Humanitarian War: A New Consensus?', *Disasters* 25(4): 331–344.

Appendix 1 Research Methodology and Timing

Activity	Timing
1. Consultation and agreement re appropriate case study for the humanitarian aid sector.	July 2002
2. Agreement from the Sphere Management Committee that this case study would complement the Sphere evaluation. Access to files etc approved.	August/ September 2002
3. Preliminary literature review and key informant interviews.	September 2002
4. Visit to Geneva. Key informant interviews, file search and documentation review.	October 2002
5. Preparation of draft chronology and 'preliminary answers'.	October 2002
6. Completion of key informant interviews and review of documentation.	November/ December 2002
7. Preparation of Working Paper.	December 2002/ January 2003
8. Circulation of draft for comments and revision of draft.	February to April 2003

Appendix 2 Key Informants

Niels Dabelstein	Head of Danida Evaluation Secretariat
John Borton	Team Leader for Study 3 of Joint Rwanda Evaluation
David Bryer	Ex-Chief Executive, Oxfam GB and Ireland (1992–2000)
Tony Vaux	Ex-Emergencies Director, Oxfam GB and Ireland
Nick Stockton	Ex-Emergencies Director, Oxfam GB and Ireland
James Darcy	Ex-Emergencies Manager, Oxfam GB and Ireland
Paul Smith Lomas	Emergencies Director, Oxfam
Hisham Khogali	Ex-Oxfam GB and Ireland
Peter Walker	Ex-Head of Disaster Policy Department, IFRC
Eva Von Oelreich	Head of Disaster Preparedness and Response, IFRC
Jim Bishop	Director for Humanitarian Response, InterAction
Susan Purdin	Sphere Project Manager, July 1997 to June 1998
Nan Buzard	Sphere Project Manager October 1998 to present
Sean Lowrie	Sphere training manager
Nancy Lindborg	Mercy Corps representative on Sphere Management Committee (and chair of Sphere Management Committee since 2001)
Rebecca Larson	LWF representative on Sphere Management Committee to 2002
Ed Schenkenburg	ICVA representative on Sphere Management Committee
Joel McClellan	SCHR representative on Sphere Management Committee
Pierre Perrin	ICRC representative on Sphere Management Committee
Agnes Callamard	Director, HAP
Jean-Marie Kindermans	President, MSF-Belgium
Craig Sanders	UNHCR
Dominic Bartsch	UNHCR
Lola Gostelow	Nutrition Sector Manager for Sphere 1997/98, SCF-UK
Peter Hawkins	Sphere Management Committee from 1996, SCF-UK

Appendix 3 The Research Questions

Overall research question

How can policy-makers and researchers make better use of research to contribute to more evidence-based policies that reduce poverty, alleviate suffering and save lives?

The policy context: politics and institutions

To what extent is the impact of research on policy-making shaped by political and institutional structures and ideological assumptions?

1. How did the global, national and community-level political, social and economic structures and interests affect the room for manoeuvre of male and female decision-makers in particular policy areas?

Any political, social or economic factors which might have influenced the key decision makers. These might include:

- financial interests of international banks in relation to debt cancellation
- socio-economic interests of vets in blocking the liberalisation of regulations

2. Impact of local political interests of warring factions on food distribution. How did assumptions influence policy-making, to what extent were decisions routine, incremental, fundamental or emergent, and who supported or resisted change?

Information about:

- existing knowledge and values and the various actors
- the old and new ideas underlying decisions
- how much new policies might change the status quo
- who supported and resisted the changes and how

3. How did applied and academic research influence the development of policy when being put into practice?

How were the policies were developed, adapted or distorted as they were put into practice by, for example, getting information about actions taken, and research used, by project partners, ‘street level bureaucrats’, and communities.

Evidence: credibility and communication

To what extent did local involvement, the quality of research, and communications strategies affect the impact that research had on policy-making in particular areas?

1. How was information gathered and by whom?

Who did the research – beneficiaries, NGO practitioners, activists, academics, consultants, government researchers; was it academic/applied and before/during/after/not part of a ‘project’ with non-research aims.

2. What was perceived as credible evidence by different actors and why?

Information about what parts of the knowledge produced was valued, by whom, and how did they assess research quality.

3. Did researchers segment their audience and if so, how, and did this affect impact on policy?

Information about whether information was tailored for particular policy-makers (according to organisation, gender, discipline, policy area, etc.) and what impact it had.

4. How and why was information ignored, reinterpreted and distorted and by whom?

Information about what happened to research findings within policy communities, e.g. snowballs, whispers, trickles etc.

5. Did the communicator, channel, format, style or timing of the communication play a role in influencing policy-makers?

How were the research findings communicated, at what stages, and by whom, and assess what worked, what failed, when and why.

Links: influence and legitimacy

To what extent is research used more effectively in policy processes if researchers and policy-makers share particular kinds of networks, common goals and chains of legitimacy for particular policy areas?

1. What roles were played by which kind of groups and male/female individuals and what were the links between them?

Information about the relationships and networks between actors. They might include: epistemic communities, global public policy networks, issue groups, advocacy coalitions and citizen action groups etc.

2. Which women or men had significant influence over the policy?

Who were regarded as experts in the particular policy area? Who were the effective and ineffective policy entrepreneurs, connectors, mavens and salesmen? Were there any differences in impact according to gender, nationality, race and/or class?

3. How did women and men researchers and advocates establish legitimacy; did it make any difference to the policy outcomes?

What were the chains of legitimacy between various stakeholders (e.g. whether southern partners controlled monitoring of action research, or degree to which informants were consulted about policy recommendations), and assess the influence of these chains on the quality of policy decisions.

Appendix 4 Sphere Project – Chronology

Date	Event/Activity	Location
1994		
	Publication of the ‘Code of Conduct for the International Red Cross and Red Crescent Movement and NGOs in Disaster Relief’. A voluntary code of conduct, seeking to safeguard high standards of behaviour and maintain independence and effectiveness in disaster relief. Presents ten principles of behaviour to be applied by the Red Cross and Red Crescent Movement and NGOs in disaster relief, and outlines five recommendations to governments of disaster-affected countries, donor governments and to intergovernmental organisations.	Geneva
	Four British-based agencies submit a proposal to ODA for a project to establish minimum standards and best practice in recruitment and support, and develop means of best practice implementation. A report commissioned into the working experience of expatriate support and managers employed in humanitarian assistance programmes. (This is the origins of People in Aid)	UK
April to July	Genocide in Rwanda, after the shooting down of President Habyarimana’s plane on April 6, resulting in the death of 500,000 to 800,000 people.	Rwanda
14–18 July	Approximately 850,000 refugees cross into Goma, Zaire, from Rwanda. Around 50,000 die in the first month as a result of cholera, dysentery, dehydration and violence. Over 200 NGOs involved in the humanitarian response.	Rwanda, Zaire
October	Danida presents concept of a joint evaluation of the international response to the Rwanda crisis to the DAC Expert Group on Aid Evaluation. No consensus agreement.	Paris
November	Danida calls a consultative meeting of bilateral donors, UN agencies and NGOs and achieves consensus on the outline terms of reference for a joint evaluation of the international response to the Rwanda crisis.	Copenhagen
December	Joint Evaluation Steering Committee meets for the first time, and approves final TOR for the evaluation.	Washington
1995		
January	Joint Evaluation begins. Four evaluation teams meet.	Geneva
April to July	Field work for Study 3 of the Joint Rwanda Evaluation carried out.	Great Lakes Region
June	InterAction submits a proposal to OFDA for the ‘Development of Training for PVOs in Complex Emergencies’, specifically to improve the response capability of PVOs.	US
September	‘Room for Improvement’ published, concluding that agencies need to agree a Code of Best Practice on human resource management, and a professional body for aid personnel be established, responsible for implementation of Code of Best Practice.	UK
17–18 September	InterAction PVO Conference on Disaster Response, as a result of which Disaster Response Committee establishes two task forces: one on NGO Field Cooperation Protocol and one on best practice for disaster work. Funded by Ford Foundation grant of \$75,000, most of which was used to fund InterAction participation in SCHR project (supplemented by some money from a 2 nd FF grant in 1996).	US
November to January 1996	Joint Rwanda Evaluation Steering Committee discusses and provides feedback on the draft evaluation reports.	Meetings in Copenhagen
December	InterAction’s best practice task force decides to focus on water and sanitation, and on food security. Two working groups created.	US
December	First Humanitarian Summit, hosted by European Commission.	Madrid
1996		
February	People in Aid project established.	UK
February	‘Towards Quality and Accountability Standards in Humanitarian Relief’ proposal drafted for SCHR members by NS and PW (and circulated for comment). Proposes establishing a set of technical standards in four sectors, based on the rights of those affected. Considers buy-in from donors, and collaboration with UN and non-SCHR agencies. Proposes in-house secondment for developing the standards. Budget of CHF467,000 (rises to \$500,000 by end of Phase 1).	Geneva

12 March	Publication of multi-donor evaluation of emergency assistance to Rwanda. Critical of the performance of some NGOs, supportive of current initiative to set standards, and proposing some form of regulation or enforcement. An international accreditation system is particularly recommended.	Copenhagen
April	OXFAM prepares a position paper on 'Setting Standards for European Humanitarian Aid', referring heavily to multi-donor Rwanda evaluation, and supporting the recommendation that the Code of Conduct and associated set of standards be widely disseminated and promoted within the international humanitarian system. At this early stage in the development of minimum standards, agencies should ensure compliance through self-regulation.	UK
29 April	InterAction DRC meeting. Progress on task forces reported upon. PW and NS present SCHR proposed project to identify practical standards for the provision of emergency aid, based on the rights of beneficiaries. Two tracks envisaged: a) identification of performance standards based upon best practice; b) developing a methodology for setting standards and ensuring compliance. ICVA commit to supporting SCHR/InterAction initiative rather than launching a parallel initiative.	Alexandria
End April	Danida asks Danish Red Cross and two other Danish NGOs to think through how to operationalise the Code of Conduct.	Denmark
7-8 May	Meeting to discuss follow-up to Rwanda evaluation, attended by donor representatives from Denmark, Finland, Germany, Ireland, Netherlands, Norway, Sweden and UK. Consensus re. the need to set and ensure attainment of standards of performance in the provision of humanitarian assistance, not just for NGOs. Current initiatives to do this were acknowledged. ODA and Danida to 'map' the various initiatives and to take the necessary steps to coordinate them. Donors also need to harmonise the monitoring and reporting required of implementing partners, e.g. with a core set of indicators, which should be linked to technically-based standards of performance.	Copenhagen
May	JEFF established, to disseminate and monitor follow-up to the Joint Rwanda Evaluation.	
11 June	'Informal Technical Working Group on Meeting Standards, Performance, Assessment and Monitoring in Humanitarian Aid Programmes' meets in Stakis Hotel, London, organised by ODA, attended by BRCS, Danida, DHA, IFRC, ODI, OXFAM, People in Aid and UNHCR. Part of follow-up to inter-agency Rwanda evaluation: taking stock of related initiatives (DAC, UN, US agencies, SCHR etc); looking at ways of encouraging implementation and monitoring of the RC/NGO Code of Conduct by developing operational performance criteria; developing and promoting greater consistency in funding requests to donors and subsequent reporting (against agreed package of core indicators); developing mechanisms by which donor agencies can share information on key M&E reports. Agreement re. the importance of developing universal standards (i.e. consolidating current initiatives). Ombudsman discussed, but concern that it would be impossible to implement, and would need a framework of standards and codes first. Agreed that ODA to explore the possibility of establishing an informal network on Standards, Performance Assessment and Reporting.	London
July	InterAction agencies adopt 'NGO Field Cooperation Protocol'.	US
September	SCHR project proposal prepared on 'Quality and accountability standards in Humanitarian Relief': a universal set of best practice to be developed, based on recognising and elaborating beneficiary rights.	Europe
September	Recruitment of SCHR Standards Project Manager commences.	Geneva
1 October	SCHR Management Group on Performance Standards meets. Role of the Management Group discussed, including taking on 'significant advocacy responsibility'. Nick Stockton appointed as Chair. Project Manager to be recruited by IFRC (PW). SCHR members identified to lead each of four sectors: food and nutrition, shelter, health and medical services, water and sanitation. Inclusion of southern NGO voices discussed.	Geneva
2 October	SCHR hosted a meeting to discuss all ongoing work on best practices: attended by InterAction, ICVA, VOICE and ACT, plus about 12 individual NGOs. Tentative agreement to consolidate all work on best practices and on standards within the framework of the SCHR project, with the possibility of InterAction coming onto the management team. (InterAction subsequently consulted with its membership re.	Geneva

	DRC participation). Issues raised included: clarifying whether the principal focus is best practice or minimum standards; how to handle compliance. Noted that the Claimants' charter is the heart of the document.	
October	SCHR members begin to commit funding to Performance Standards project.	
November	PR/advocacy plan prepared for SCHR Performance Standards project.	
December	ECHO/VOICE Forum on 'Ethics of Humanitarian Aid'.	Dublin
1997		
8 January	SCHR Performance Standards Management Team Meeting. Key points: widest possible professional and technical participation to be sought in each sector; funding strategy proposed, whereby SCHR/InterAction affiliates approach donors, who are asked to give proportionately, according to their relative significance as a humanitarian donor. Possibility of UN participation rejected. To remain an NGO initiative.	London
June	JEFF report published, reviewing the follow-up and impact of the Joint Rwanda Evaluation, fifteen months after publication.	
July	SCHR and InterAction officially launch the one-year Sphere project. Susan Purdin takes up post of project manager of Standards Project.	Geneva
July	DHA indicates it wants 'to participate actively' in the standards project. UNICEF also expresses interest in contributing, especially around the issue of rights. UNHCR to recommend technical experts to contribute to the standards work. WFP also wants to be involved.	Geneva
July	Standards project given the title of 'The Sphere Project'.	Geneva
July	VOICE invited to participate in Sphere Management Committee as an observer.	Geneva
July	Meeting of a small ad hoc committee to consider establishing an 'International Standing Capacity to develop, monitor and report on international progress in addressing the phenomenon of complex emergencies'. Paper proposing how this might work written in August.	Geneva
August	Sphere project manager meets with International Standards Organisation to explore setting standards for service organisations.	Geneva
August	Sphere website established.	Geneva
August	Sphere project manager sends her first draft of the humanitarian charter to the management committee and sector managers.	
18–19 August	First meeting of Sphere sector managers.	Geneva
25 August	SCHR receives a letter from Kofi Annan: 'At a time when the United Nations and its partners are working to better the needs of the victims of humanitarian crises, I am gratified to learn that your organisation is working closely with InterAction to develop a humanitarian charter and a set of standards for humanitarian response. The contribution from non-governmental organisations will indeed enhance the capacity of the international community towards an effective and coordinated response to emergencies'.	
August	UNHCR establishes an Inter-Agency Advisory Group for Refugee Health. The TOR include reviewing existing guidelines and developing common evaluation criteria.	Geneva
15–16 September	Second meeting of Sphere sector managers.	London
September	ICRC becomes a member of the Sphere Management Committee with observer status.	Geneva
26 October	Third meeting of Sphere sector managers.	Washington D.C.
October	Follow-up meeting on establishing 'International Standing Capacity', although this never happened in the end.	Geneva
4 December	Fourth meeting of Sphere sector managers. Outline of final Sphere document agreed.	Amsterdam
December	Formation of ALNAP to provide an interagency forum to discuss and link the various accountability initiatives, and undertake complementary activities on learning and accountability.	London
December	Group of UK organisations form to develop the concept of a Humanitarian Ombudsman, as recommended in the Rwanda evaluation. (In 2000 this becomes an international initiative, and is renamed the Humanitarian Accountability Project).	London

1998		
3 February	Humanitarian Charter Working Group meets to revise charter. Various drafts of humanitarian charter circulated amongst Sphere Management Committee members between February and May.	Brussels
14 February	Steering Committee of the Joint Rwanda Evaluation is re-convened to consider the JEFF findings.	Copenhagen
February	First drafts completed of the Sphere chapters on minimum technical standards.	
February	Sphere regional meeting held in Monrovia, led by sector manager.	Africa
March	Resuscitated ICVA invited to participate in Sphere Management Committee as an observer.	Geneva
March	Sphere regional meetings held in Bangkok, Dar Es Salaam, Harare, Dacca, Nairobi, Sarajevo, Abidjan and Conakry. Each led by two sector managers, and by the Sphere project manager in the case of the Sarajevo meeting.	Asia, Africa, Balkans
April	Sphere regional meeting held in Beijing, led by sector manager.	China
April	Sector managers complete their secondments to Sphere (except the health sector which continues until June). Editing on standards document commences.	
11 May	Humanitarian charter approved by Sphere Management Committee.	
May	Draft Sphere document printed and posted on web for review.	
May	Discussions between Sphere project manager and RedR re. incorporating standards into training programmes.	London
June	Discussion at British Red Cross World Disaster Forum about possible links between Sphere and the Ombudsman project, noting that there is a need to continue the dialogue, but there is not yet sufficient buy-in to the Ombudsman idea.	London
June	Two articles appear in two Geneva newspapers on Sphere.	Geneva
30 June	Susan Purdin leaves as Sphere project manager. (PW to provide cover in the interim, supported by Isobal Macconnan). Sphere Phase 1 due to end, but decision taken in May to extend Phase 1 to end of October 1998.	Geneva
July	'French letter' – a number of French NGOs express their disquiet and concerns with the Sphere initiative. (This was not responded to until November)	France
15 October	Nan Buzard commences as new Sphere project manager.	Geneva
1 November	Sphere Phase 2 begins, to focus on publication, institutionalisation and dissemination of standards.	
3 December	Sphere project launch, London and Washington. First trial edition published.	London, Washington
1999		
	Field-testing of Sphere handbook and revision of first trial edition.	Global
2000		
	Publication of first final edition of Sphere Project.	

