

DRAFT: NOT FOR PUBLICATION

INTRAC Conference, 31 March – 4 April, 2003

**Balancing Measurement, Management & Accountability:
Lessons learned from SC UK's impact assessment
framework**

**Simon Starling
Save the Children UK**

1. Introduction

International non-governmental organisations (INGOs) are increasingly under pressure to 'measure' their performance and the results of their development and relief work. This has led to the development by many INGOs, including Save the Children UK (SC UK), of new frameworks and measurement systems. This paper looks at SC UK's new impact assessment framework and draws on the lessons learned from implementing it in our country programmes to illustrate some of issues for the non-governmental sector in relation to balancing the potentially competing needs for measurement, management and accountability¹.

The drive for better performance and impact assessment

Pressure to assess the impact of agencies work, and to build upon this to look at organisational performance, comes from both within and outside development agencies. Some key factors (not in any order of priority) include the need:-

- To understand the implications of, and improve, development and relief work
- To combat scepticism about aid in general
- To demonstrate organisational performance in a competitive marketplace
- To become more accountable to relevant stakeholders (those we work with and those who support us)

This is within a context of a 'contract culture', where demands for increased performance measurement and accountability has permeated not only the private and public sector, but increasingly the non-profit sector too (Hailey, 2003). While this drive towards better assessment of impact and performance has provoked rhetorical statements from NGOs, it has proved harder to put into practice.

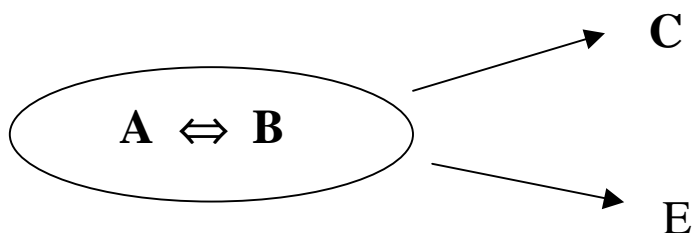
The complexities of assessing impact

INGOs struggle to define better performance and impact measurement systems is hampered by a range of factors to do with the context of development work as well as the dynamics of NGOs themselves (see Fowler, 1997: 160). Perhaps the most fundamental of these is the nature of development work itself. Any attempt to 'measure' development is hampered by the complexity of what must be assessed, inconsistencies about the assumptions of the aid system and how sustainable development actually occurs, and the tendency of development actors to over-simplify how change or 'development' occurs.

Traditional models of development have assumed a linear model of change; that there is a direct causal relationship between development project inputs, outputs and impact. However, recent years have seen more recognition among development practitioners that this model is an oversimplification, and that there is a complex relationship between project inputs, outputs and impact. Change is the result of a complicated interplay of events, people and conditions present in a given situation; thus a similar project with identical inputs in a different context is unlikely to produce identical impacts.

Roche (1999: 24) presents a more realistic model of change...

¹ This paper is a work in progress, and the author would welcome any comments on the draft to date.



...where the interaction of inputs (A) and outputs (B) lead to *intended* impacts (C) and *unintended* impacts (E). The recognition that contextual factors play a critical part in development processes, and that the results of the interplay of different influencing factors produces very different results to those predicted, both negative and positive, is an important one. It implies that INGOs, in order to make a more realistic attempt to assess their performance, must make an effort to understand the complex nature of the longer-term intended and unintended impacts of their work, and the complex processes that lead to these.

Moreover, development interventions affect people's lives in different ways; a positive impact for one person may not be so for another, or may even be negative. Assessing impact therefore involves making judgements about what changes have occurred and, critically, what change is considered significant by whom (*ibid.*). Impact assessment for INGOs, then, is not only about systematic analysis, but about drawing on the different perspectives and experiences of external and internal stakeholders in order to make as accurate a judgement as possible.

For larger development actors, the sheer scale and diversity of their operations can make making judgements about performance additionally difficult. SC UK, for example, works in nearly 70 countries world-wide with a huge range of development and relief programmes. Can data from such diverse programmes be aggregated to produce meaningful comments on SC UK's performance and impact as a whole²?

Moreover, SC UK works in collaboration with a countless number of other actors, including for example, community groups, children & young people, partner NGOs, local and national governments, and multi-lateral institutions. Where development impacts are identified, is it possible to attribute these to just one agency?

2. SC UK's Global Impact Monitoring Framework

Over the last 18 months SC UK has been developing a framework for its programmes world-wide known as Global Impact Monitoring (GIM). The main purpose of the GIM

² The complexities of assessing impact and making judgements about organisational performance for large development agencies were illustrated in a recent report commissioned by DFID which tried to assess its 'development effectiveness'. The report concluded that it was not possible, drawing on existing evidence and mechanisms, to make any sound global statements about the impact of DFID's work and therefore its effectiveness as an organisation (Flint et al., 2002).

process is to summarise and analyse, at different levels, the impact of Save the Children's work and the progress made towards achieving change objectives. It provides a consistent framework across SC UK's wide range and aims to:-

- better understand the implications of our work, and to be able to share this understanding with others;
- facilitate institutional learning and support decision making by feeding into strategic policy and planning processes in order to enhance the quality and impact of our work;
- improve our accountability to all relevant stakeholders, including our partners, children, young people and their communities as well as trustees, donors and management at all levels;

The key characteristics of the GIM approach are:

- A focus on impact, i.e. on changes as a result of our work and on the key processes leading to such changes;
- A system and process that is simple and built upon existing mechanisms, but flexible enough to accommodate the specific elements of diversity inherent in any development process, and which allows for revisions as a result of experience and emerging priorities;
- An inclusive process that involves key external as well as internal stakeholders at all levels.
- Identification of enough general dimensions of change to facilitate some comparability across country programmes and regions within an area of work.

SC UK has defined impact as the extent to which our work has produced lasting or significant changes - positive or negative, intended or not - in the lives of children, young people and their communities. Drawing on organisational values, we have identified a series of 'dimensions of change', that try to break down 'impact' into 5 key dimensions that could be applied to any/all of our work:

Common dimensions of change of SC UK work:

1. Changes in the lives of children and young people

Which **rights** are being better fulfilled? Which rights are no longer being violated?

2. Changes in policies and practice affecting children and young people's rights

Duty bearers are more **accountable** for the fulfilment, protection and respect of children's and young people's rights. Policies are developed and implemented and the attitudes of duty bearers take into account the best interests and rights of the child.

3. Changes in children's and young people's participation and active citizenship

Children and young people **claim their rights** or are supported to do so. Spaces and opportunities exist which allow participation and the exercise of citizenship by children's groups and others working for the fulfilment of child rights.

4. Changes in equity and non-discrimination of children and young people

In policies, programmes, services and communities, are the most marginalised children reached?

5. Changes in civil society and communities capacity to support children's rights

Do networks, coalitions and/or movements add value to the work of their participants? Do they mobilise greater forces for change in children and young people's lives?

The 5 dimensions of change provide a lens through which impact can be examined, though different programmes will contribute more in certain dimensions than others and some may not be relevant at all.

The GIM process occurs at three different levels:

- Country level: stakeholder Impact Review meetings, compilation and cross checking of evidence from different sources, analysis in Country Impact Report.
- Regional level: convening of regional workshop(s) to examine country reports, and identify regional gaps, trends. Production of Regional Impact Reports.
- Global level: synthesis and analysis of Regional Impact Reports to produce a Global Impact Report.

The key elements of the process, therefore, are the identification, collation and analysis of impact-related evidence and the specific involvement of a range of stakeholders in producing and analysing this information. It does not replace existing project/programme monitoring, evaluation and review mechanisms, rather it seeks to build on existing mechanisms and uses a variety of mechanisms and methodologies from which evidence can be drawn (project monitoring systems, external evaluations, interviews, case studies, etc.). It encourages a mix of qualitative and quantitative evidence and explicitly sets out to capture negative/unintended impacts as well as positive/intended ones. Information from different sources can then be cross-checked and jointly analysed using the 5 dimensions of change in Impact Review meetings with external stakeholders. The result should be as an accurate picture as possible of the actual impact of our work, drawn from a variety of mechanisms, sources, and contributors.

The GIM process was piloted in 16 country programmes in 2001-2. Many of the lessons learned from the pilot relate to the issues of measurement, management and accountability and are described below.

3. Lessons Learned: Measurement

Quality of evidence

The GIM process aims to draw out examples of impact, both positive and negative, intended or not. In this it was relatively successful. However, the quality of evidence used to back up claims about impact varied. In some cases comments were overly generic, unreferenced or in some cases unsubstantiated (i.e. “children are experiencing direct benefits”). On the other hand, the most interesting examples combined both quantitative and qualitative information from different sources to illustrate impact.

Box: Reunification of separated children in Democratic Republic of Congo

The protection of vulnerable children was improved through activities to reunify, reintegrate and promote community care of children, and through the participation of Community Child Protection Networks in questions relating to the rights and protection of children.

- 709 non-accompanied children were identified in Ituri province through networks of local partners. We were able to reunify 82% of these children through the networks.
- 230 Rwandan children who lived with Congolese families (mainly in North and South Kivu) were supported in their voluntary repatriation.
- 264 former child soldiers were referred to interim care centres managed by our partners. 68% of these children were reintegrated within their families and their communities.

In South Kivu, an unintended impact of SC UK’s work to demobilise and reintegrate child soldiers was that other child soldiers in areas where SC UK worked began leaving armed groups on their own initiative (i.e. without the formal demobilisation process).

Finding an appropriate balance in terms of presenting examples is also important. Some examples were very short - just a line or two - not therefore enabling the reader to understand the context of the example and SC UK's role in it. Longer examples were found to be more useful, but including more context would probably mean selecting fewer examples overall to maintain 'readable' reports.

As the GIM process builds largely upon evidence from existing project and programme monitoring and evaluation mechanisms, unless these mechanisms produce data of sufficient quality, the content of GIM reports will be general and weak.

Avoiding Bias

Integrating the perspectives of external stakeholders (partners, children & young people, communities, local government, etc.) is a key element of the GIM process. The 'human factor' – the intentions, commitment and approach of those involved - is therefore an important one in such assessment processes (Hailey, 2003). Ensuring that Impact Review processes and meetings do not become merely 'good news factories' is therefore critical. Given the complex nature of relationships between INGOs and those they work with, however, ensuring stakeholders give honest and frank accounts of their experiences, good and bad, is not easy. In Bangladesh for example, some participants involved in Impact Review meetings clearly had agendas of painting their own roles in a positive light. Certain participants had been selected by partner organisations to present a particular (positive) viewpoint. Creating spaces for honest and open exchanges is a subject which is given a lot of attention in the guidelines for the GIM process, but is dependent, to a degree, on existing relationships, and the ability of staff to create spaces where constructive criticism is welcomed.

While the GIM process elicits a range of views, it is ultimately a SC UK senior staff member who must score or explain progress towards objectives based on the evidence and perspectives gathered. Whilst perverse incentives to provide a positive picture may exist, the impact analysis presented under the 5 dimensions should steer the scoring process. The idea is that those responsible for programmes should, based on the impact analysis, be in a position to formulate a subjective but informed judgement on the progress made towards achieving the programme objectives.

Change oriented objectives and steps towards impact

One problem found in some cases was that programme objectives were not sufficiently change focused i.e. generic, unfocused, or focusing on programme actions rather than conditions to be changed. While this does not prevent an impact assessment process, it does make it more difficult. Fowler (1997: 169) notes that intended outcomes are infrequently laid out by INGO projects. This was also the case in some of the pilot countries: some programmes have not sufficiently thought through the steps needed to achieve change over time, and the milestones that would indicate these had been achieved. Whilst this is a handicap to the GIM process in the short-term, the process of engagement could encourage programmes to outline more clearly steps towards impact and milestones. Indeed, in Central America programmes found it useful to use a matrix tool to retrospectively examine changes over time in relation to the 5 dimensions and to help identify examples of impact.

Timeframe

GIM was conceived as an annual process to be applied across country programmes. However, experience from the pilot challenged whether this is the most appropriate timeframe for a genuine and rigorous assessment of impact. Some argued that an annual timeframe was too short – that significant changes in the 5 dimensions are unlikely to occur on an annual basis and processes leading to change generally take longer than one year. In addition, it was found to be difficult to limit analysis to activities and changes that have occurred in the previous year – it may make little sense in the timescale of the programme to look at changes in the period stipulated by the GIM process or SC UK's financial year. Different options are now being considered to address this issue (e.g. GIM process on different themes annually, or full GIM process on country programmes less frequently).

Assessing the Impact of Advocacy work

The 5 dimensions of change are explicitly designed to be used for both operational programming and advocacy activities. However, assessing the impact of advocacy initiatives can be particularly challenging, given their often long-term and flexible nature means they may evolve rapidly in response to political/environmental opportunities, and the question of attribution when working in a collaborative manner with other organisations.

Programmes can find it difficult to track their influence on policy-makers attitudes and positions. More daunting still is monitoring whether new policies are actually being implemented and whether sufficient budget allocations are taking place. This implies a particular kind of analysis - investigating the content of policies and budgets and monitoring legislative decisions and amendments – activities which development and relief agencies like SC UK may be less familiar with (SC UK, 2003).

Programmes sometimes found it difficult to connect their policy influencing successes (dimension 2) with concrete changes in children's lives (dimension 1). Clearly, advocacy activities should all ultimately have impact at this level, but as is widely acknowledged by the evaluation community, evaluating the ultimate outcomes of public policies on citizens is an immense challenge. It implies a long-term commitment in terms of impact assessment, and more sustained investment in establishing micro-macro linkages to determine how policies impact the populations we aim to serve (ibid.). In recognition of some of these specific challenges, we are now producing some more in depth guidelines on how the GIM framework can be used to assess advocacy work.

4. Lessons Learned: Management

Purpose of reporting

Some of the country programmes in the GIM pilot felt that the new reporting format did not sufficiently cover some areas they felt important in terms of programme management information e.g. staff development, security issues, internal and external context, and funding levels. Indeed, these are all potentially relevant to determining the impact (or not) of a programme.

In addition, the GIM process quite deliberately shifts the emphasis towards highlighting the perspective of others outside SC UK involved in or affected by our work. This to a degree

de-emphasises the manager's perspective (though an overall assessment of progress is still required from managers based on the multi-stakeholder impact analysis), and this was felt by some to be an important component of the existing reporting process.

Both these issues relate to a broader question highlighted by the pilot: what is the purpose of (annual) reporting? There appears to be an inherent tension between evaluation as a management and as a learning tool. In many senses an impact report cannot be a management report. SC UK does not have a Management Information System (MIS), or rolling computerised database of management information. Its existing annual reporting mechanism therefore aimed at providing an overview of what is being done, provide information needed to make management decisions and provide some basic accountability in terms of what has/has not occurred. More in-depth reviews of our programmes' impact then are likely to be at the expense of management information if the two purposes are combined into a single report. We are now looking at various options to address this issue.

Resource implications

Perhaps unsurprisingly given that this was the first round of an introduction of a new process, many involved found it to be a time and labour intensive process. The Democratic Republic of Congo reported that it was a "useful and interesting process [but] ...would dread having to do it every year". Some countries reported that it was difficult to bring stakeholders together for financial, geographical or political reasons (China, Congo, Zimbabwe).

While it is evident that impact assessment and analysis takes time and energy, it is expected that as SC UK programmes begin to include GIM principles and processes in planning and budgeting it should become a natural and ongoing part of our work rather than 'additional'. Indeed, as it replaces an existing reporting mechanism, the only extra work is in bringing together different stakeholders to elicit and analyse impact information (where this is not already happening as a matter of course within programmes).

Another observation is around staff skills. The GIM process involves being able to create 'spaces' where people feel encouraged and able to talk about the effects of our work, despite the complex nature of relationships between people and organisations (see Avoiding Bias section). It is therefore a process that requires skills in eliciting sensitive information, in careful facilitation of meetings, and in analysis of data (quantitative and qualitative). Not all staff, or programmes may have sufficient capacity in these areas (particularly where programmes have traditionally been service delivery oriented). Adopting impact assessment processes like GIM has implications in terms of skills required from NGO staff, and this may require revisiting staff development priorities, job specifications and performance appraisal processes.

5. Lessons Learned: Accountability

It is hoped, and an aim of the GIM process, that greater involvement of those we work with in assessing our work will to a degree increase the organisation's accountability to those people. While it may be too early to comment on whether GIM has had a significant impact in this respect there are some early experiences worth mentioning.

In Bangladesh in February this year, a GIM Impact Review meeting was held in the Kurigram area in the north of the country. The meeting brought together around 70 local people involved in, or affected by SC UK's work, to discuss its impact. Participants included parents, teachers, working children, local government representatives, NGO partners, and community leaders. While such large multi-stakeholder meetings have distinct disadvantages as well as advantages³, what was clear was that participants, including the children, relished the opportunity to have their say about the work and its effects upon them. Levels of participation and engagement were high, and participants asked for a similar event on a monthly basis! The programme must now grapple with how best to build upon the enthusiasm and 'buy-in' generated, and how to ensure these perspectives are considered in decision-making processes. Plans are already being made to feedback to the participants in different ways.

This trend was repeated in other countries that undertook GIM and is a key lesson – that opportunities to comment freely on our work and its negative or positive impacts is valued by those we work with. In some country programmes these types of processes are already commonplace and ongoing; in others the GIM process was the first time that engagement outside the 'project box' had been attempted.

The GIM process has also been used to report to SC UK's institutional donors. An important step has come from DFID, who have accepted the GIM framework as the basis for reporting against the Programme Partnership Agreement between the two organisations. Though the framework will not replace specific donor reporting requirements, it is hoped that it can be used to ensure consistency of reporting internally and externally. This a key concern of programme staff who do not wish to have to operate to contrasting demands from within and outside SC UK.

6. Implications for INGOs: Towards Balancing Measurement, Management & Accountability

Interpretative vs. scientific approaches

Given the complex realities of development work, the nature of relationships between people and organisations, and the multiple contributions of different actors, it becomes almost impossible to 'prove' an impact can be solely attributed to the actions of one agency. The GIM impact assessment framework therefore draws upon evidence from both existing 'scientific' methods and more 'interpretative' approaches to try to build an accurate as picture as possible about impact, using triangulated evidence from key stakeholders to make an informed judgement.

However, the prevailing trend within the sector remains a Results-Based Management

³ Experience from Laos, Bangladesh and other countries indicates that larger multi-stakeholder meetings have significant advantages in that they provide an opportunity to cross-check impact information, generate commitment and motivation from different actors, increase transparency and understanding, and generate new opportunities for collaboration. However, it may be more difficult to elicit genuine examples of impact in such an environment, and participants will have varied levels of engagement with different aspects of the work, making detailed discussion more difficult.

approach, based largely on the use of the 'logical framework'. The limitations of this tool have been well documented, and perhaps the time has come for INGOs to explain the limitations of this approach, and the impact it has upon them (and their southern partners), and to be more confident about their own mechanisms for assessing impact and organisational performance. This has not happened to any great extent yet, perhaps because INGOs themselves are fearful of a backlash in an increasingly competitive environment. However, institutional donors are often struggling with the same methodological and practical issues and more dialogue between INGOs and institutional donors is needed⁴. Consistency across INGOs and across institutional donors can only help southern organisations struggling to manage several different accounting and reporting requirements for several different donor for every one of their projects.

Towards a culture of self-criticism and learning

Evidence from SC UK's experience is that understanding impact and stimulating learning will not result from any particular mechanism alone. It requires the engendering of a *culture* of constructive criticism and learning. This means not only creating an environment where external stakeholders can speak frankly, but also about INGO staff members becoming much more self-critical and reflective. Creating spaces for such learning to happen is in itself a resource challenge, and given that development workers are often already over-stretched, it is often seen as a luxury that can not be afforded. The onus is on senior INGO managers to make the case that reflection and learning processes are not costs, but an investment and an essential part of the development process, and to match this rhetoric with the structures, resources and support needed to put it into practice.

Towards genuine partnerships?

A more meaningful understanding of impact requires more inclusive processes and the frank participation of external and internal stakeholders. However, the Results Based Management tools that have been thrust upon NGOs in the south by their northern 'partners' may actually undermine this. Elements of the 'crisis of trust' in the UK public sector described by O'Neill (2002) could be equally valid for the development sector i.e. the more that regulation, reporting and control mechanisms are forced upon Southern partners the less they feel respected and trusted. Control, accounting and assessment may thus become the enemies of trust (Marsden, 2002), creating a vicious circle where the lack of trust engendered by control and reporting mechanisms makes it impossible to build a genuine picture of development impact and therefore assess an organisation's performance.

This brings us back, as many debates on development processed do, to the 'real politique' issues of power relations and funds. Can southern organisations dependent on INGOs really afford to highlight their own or their partners failings? Can communities really dare to be critical when the services they are receiving, if not perfect, are the only thing that they are getting? These issues are not new to the development sector, and indeed the dangers of dependency ("He Who Pays the Piper Calls the Tune") have been well documented and discussed over the last several decades ('Piper' ref.; Starling 1996). Despite this, the findings

⁴ The BOAG Evaluators Group, consisting of representatives of ActionAid, CAFOD, Christian Aid, Oxfam and SC UK, has been working to open this debate and recently presented a paper at the Development Assistant Committee – made up of government development agencies - conference in Paris (see BOAG, 2003).

of a large and comprehensive study of relationships between UK donors, UK INGOs and their partners in Uganda and South Africa indicates little has changed (see Section 4 in BOAG, 2002). Southern NGOs view their INGO colleagues as 'donors' not 'partners'; they often resent the management and control tools thrust upon them; they feel they have to constantly react and change to the latest 'en vogue' initiatives and frameworks emanating from London HQ's (gender mainstreaming; rights based programming etc.); yet, they are reluctant to reject new initiatives or criticise tools as they fear jeopardising the scarce funds they need to survive.

This raises an uncomfortable question: are Northern development agencies in fact quite comfortable with the status quo, where Southern partners are obliged to fulfil their information, reporting and accountability needs, without any real reciprocity? To move forward, INGOs must recognise the imbalance in power relations and strive to build genuine partnerships where the needs for management, measurement and accountability are balanced and take into account the needs and capacities of their 'partners' and other stakeholders.

Assessing policy change and policies in practice

There are a number of specific implications for INGOs of trying to better understand the impact of their advocacy work. One is the need to make more systematic efforts to track influence on policy-makers, the issue of attribution notwithstanding. Indeed, given the collaborative nature of advocacy initiatives, SC UK increasingly considers that efforts to attribute advocacy success to the work of one organisation are not only methodologically problematic but miss the point. The key is to assess how working with others adds value to our work. Therefore, through our impact assessment strategies, we aim to be clear about, and consequently track, what we *contribute*, in terms of research, policy positions, influencing activities and our input within networks (Save the Children, 2003).

Perhaps the bigger challenge lies in ensuring that policy changes actually result in effective policy implementation. Evaluating policy effectiveness is a complex issue, which has not been fully addressed by NGOs, particularly in the area of micro-macro linkages. There is a growing need for NGOs to take on more thorough policy and budget analysis and monitoring. This ensuring that our operational programming is complemented with this new level of analysis and action. It may also lead us to selecting new partners to work with.

Improving Accountability

INGOs face myriad complex accountabilities to different actors, not just their donors and 'beneficiaries' (e.g. public, host government, own government, staff, trustees, etc.). More inclusive processes which involve people affected by and involved in our work can only help increase transparency and mutual understanding. Ensuring that INGOs use learning effectively from such processes, and use them as the basis of decision making is more challenging.

Given the myriad accountabilities, it becomes almost impossible to report formally to all stakeholders. Choices must therefore be made in terms of prioritising accountabilities. O'Neill (2002) talks of a move toward 'intelligent accountability' that identifies and recognises what it is most important to monitor. This approach suggests that those who are called to account should give an account of what they have done and their successes and

failures to others who have sufficient time and experience to assess the evidence and report on it.

Like other professionals, development workers' practice, is "not reducible to meeting set targets following prescribed procedures and requirements" (ibid.). Indeed, research teams in South Africa and Uganda reported that the single most important factor in successful working relationships between northern and southern agencies was personal relationships – key staff member's ability to show and engender respect, listen as well as prescribe, strive to understand the context and others viewpoints (Section 4 in BOAG, 2002). These are not qualities easily reducible to targets or prescription. Such an approach – where different actors in the aid chain collaborate to identify some priorities for monitoring and operate more on a basis of trust and respect than trying to reduce development work to a series of targets – is once again dependent on the ability of organisations develop relationships with others in the development chain that match their own rhetorical statements about partnership.

8. Conclusion

There appears to be a need for all development actors to recognise the sometimes competing/contradictory demands for measurement, management and accountability and the limitations to totally fulfilling all these requirements. This requires giving up the 'fantasy of total control' (O'Neill, 2002), and striving to find an acceptable balance between measurement, management and accountability. No-one methodology can deliver an adequate organisational performance assessment of INGO programme work; rather a range of methodologies and mechanisms, including 'impartial' external evaluations as well as participatory assessment methodologies are needed. However, development actors need to be wary of over-burdening each other, or themselves, with different regulatory frameworks. More consistency among development actors in terms of methodologies and frameworks might contribute to a decrease in this burden.

However, a key learning from SC UK's experience is that however good the mechanisms and methodologies employed, they will be largely meaningless without an appropriate organisational culture – an atmosphere of openness, reflection and constructive criticism that promotes learning within an organisation - and relationships with partners and other external stakeholders that are strong enough to enable this in practice. This requires the building of relationships of trust at all levels, and there is a role for all actors in the 'aid chain' in fostering this culture.

REFERENCES

BOAG Evaluators Group, "Breaking new ground? BOAG agencies experience of developing systems for assessing organisational performance", prepared for seminar, 10th December 2002.

BOAG Evaluators Group, "Towards Organisational Performance Assessment: experiences of strengthening learning, accountability and understanding social change", paper for DAC conference, March 2003.

Flint, M. et al. (2002) "How Effective is DFID?: Development Effectiveness Report 2001"

Fowler, A. (1997) "Striking a Balance: A Guide to Enhancing the Effectiveness of NGOs in International Development", Earthscan: London.

Hailey, J. and Sorgenfrei, M. (2003) "Measuring Success? Issues in Performance Management", paper for INTRAC Conference 31 March – 4 April, KDK, Holland.

Marsden, D. (2002) "Who is Doing the Counting and Why?", in ONTRAC, No. 22, September 2002. INTRAC: Oxford.

O'Neill, O. (2002) Lecture 3, "Called to Account", Lecture 4 "Trust and Transparency", BBC Reith lectures, March. www.bbc.co.uk/radio4

Roche, C. (2000) "Impact Assessment for Development Agencies: Learning to Value Change". Oxfam: Oxford.

Save the Children UK, (2003) "Closing the Circle: from measuring policy change to assessing policies in practice"

Starling, S. (1996) "Northern NGOs in Sub-Saharan Africa: Rhetoric and Practice", M.Phil. thesis, University of Birmingham.