

**EVALUATING DEMOCRACY AND GOVERNANCE  
ASSISTANCE**

**Gordon Crawford**

**with Iain Kearton**

**ESCOR Research Report No. 7894**

**December 2001**

# CONTENTS

<b>List of Acronyms and Abbreviations .....</b>	<b>iii</b>
<b>Summary .....</b>	<b>iv</b>
<b>1. INTRODUCTION .....</b>	<b>1</b>
<b>2. DEMOCRACY AND GOVERNANCE EVALUATIONS: DONOR EXPERIENCE.....</b>	<b>4</b>
2.1 Canadian International Development Agency (CIDA).....	5
2.2 Danish International Development Assistance (Danida) .....	12
2.3 Swedish International Development Co-operation Agency (Sida).....	18
2.4 United States Agency for International Development (USAID) .....	25
2.5 European Commission.....	38
2.6 OECD Development Assistance Committee .....	46
2.7 United Nations Development Programme .....	50
<b>3. COMPARATIVE ANALYSIS OF EVALUATION STUDIES .....</b>	<b>56</b>
3.1 Introduction to Database.....	56
3.2 Scope of Evaluation.....	58
3.4 Thematic Coverage.....	60
3.5 Methodologies and Methods .....	62
3.6 Assessing Impact.....	66
3.7 Evaluation Findings.....	70
<b>4. CHALLENGES IN EVALUATING DEMOCRACY AND GOVERNANCE ASSISTANCE</b>	<b>74</b>
4.1 DG Evaluation: Art or Science? .....	74
4.2 Impact Evaluation.....	76
<b>5. MEETING THE CHALLENGES: AN ALTERNATIVE METHODOLOGY.....</b>	<b>82</b>
5.1 Political Context Studies .....	82
5.2 Meso Level Analysis .....	84
5.3 Participatory Evaluation .....	84
<b>6. CONCLUSION: LESSONS LEARNED AND THE WAY FORWARD.....</b>	<b>99</b>
<b>Appendix One: Further donor evaluation studies .....</b>	<b>102</b>
Australia: AusAid.....	102
Ireland Aid.....	103
Japan: Ministry of Foreign Affairs and Japanese International Co-operation Agency.....	103
The Netherlands .....	105
Norway: NORAD and Ministry of Foreign Affairs.....	106
UK Department for International Development (DfID) .....	109
Club du Sahel, OECD Development Co-operation Directorate and UNDP .....	117
World Bank .....	120
<b>Appendix Two: Agencies' responses to enquiries .....</b>	<b>122</b>
<b>Bibliography.....</b>	<b>123</b>

## LIST OF TABLES AND FIGURES

<b>Box 2-1: Human Rights and Democratic Development Indicators - Illustrative List .....</b>	<b>10</b>
<b>Box 2-2: USAID’s Democracy and Governance Program Indicators - Some Examples.....</b>	<b>30</b>
<b>Table 2-1: Research methods in the Danida thematic studies .....</b>	<b>15</b>
<b>Table 3-1: Evaluation Report Coverage .....</b>	<b>58</b>
<b>Table 3-2: Donor Coverage.....</b>	<b>58</b>
<b>Table 3-3: Geographical Coverage .....</b>	<b>59</b>
<b>Table 3-4: Thematic Coverage .....</b>	<b>60</b>
<b>Table 3-5: Sub-thematic Focus.....</b>	<b>61</b>
<b>Table 3-6: Participation in evaluation .....</b>	<b>66</b>
<b>Table 3-7: Types of criteria / indicators used in impact evaluations.....</b>	<b>69</b>
<b>Table 3-8: General findings by type of impact assessment.....</b>	<b>71</b>
<b>Table 3-9: General findings by scope of evaluation .....</b>	<b>72</b>
<b>Table 5-1: Differences between conventional and participatory evaluation .....</b>	<b>88</b>
<b>Table 5-2: Levels Of End-User Participation In Evaluation .....</b>	<b>89</b>
<b>Figure 3-1: Data Collection Methods.....</b>	<b>64</b>
<b>Figure 3-2: Types of Interview .....</b>	<b>65</b>
<b>Figure 3-3: Types of Impact Evaluation .....</b>	<b>67</b>
<b>Figure 3-4: General Findings.....</b>	<b>70</b>

## List of Acronyms and Abbreviations

AusAID	Australian Agency for International Development
CIDA	Canadian International Development Agency
DAC	Development Assistance Committee (of the OECD)
DCD	Development Co-operation Directorate (of the OECD)
Danida	Danish International Development Assistance
DfID	Department for International Development (UK)
FINNIDA	Finnish International Development Agency
IDRC	International Development Research Center (Ottawa)
JICA	Japanese International Co-operation Agency
NORAD	Norwegian Agency for International Development
OECD	Organisation for Economic Co-operation and Development
Sida	Swedish International Development Authority
UNDP	United Nations Development Programme
USAID	United States Agency for International Development

## Summary

Over a decade after the introduction of development co-operation activities aimed at promoting democracy and good governance, the question of *evaluation* has become a crucial one. What has democracy and governance (DG) assistance achieved? What impact has it had on democratisation processes in recipient countries? How (in)significant is the role of external actors? Evaluating democracy and governance assistance poses considerable challenges, however, notably the establishment of linkages between DG assistance and political change. In a previous ESCOR study, Mark Robinson noted that “donors lack a systematic approach to evaluation [of democracy assistance and political aid] and there is no generally accepted methodology” (1996: ii). This research addresses the extent to which evaluation in this field has subsequently progressed. It is essentially a methodological study and has two related aims. One is to provide a ‘state-of-the-art’ review and critique of DG evaluation studies, concentrating on programme not project evaluations. The other is to develop an appropriate evaluation methodology for conducting country impact studies, that is, the impact of DG assistance on democratisation in the countries concerned.

A survey of bilateral and multilateral donors discovered that many more evaluations, and related studies, had been undertaken than anticipated, with one hundred and ten references catalogued. A concentration of effort was required, achieved in two ways. First, the experience of seven key donor agencies is explored in detail in Chapter 2, that is, those that have given most attention to issues of evaluating democracy, human rights and governance assistance, while other donor studies are covered in Appendix 1. Second, a database was constructed of the main sixty reports, studied in detail for this research. Excluded from the database were project evaluations, unless individual donor agencies had not undertaken a programme evaluation, and some case studies commissioned as part of a series, typically confining ourselves to examining one or two case studies plus the synthesis report. The database facilitated the comparative analysis of the sixty reports presented in Chapter 3, including scope of evaluation, geographical and thematic coverage, methodology and methods, degree of participation, and findings.

Studies are characterised by variability and diversity, and, despite the greater attention to evaluation issues, there is little evidence of any agreement on how to evaluate DG activities. In this respect, Robinson’s statement retains its validity. Consensus is limited to an acknowledgement of the difficulties encountered in evaluating the success or otherwise of political aid, succinctly expressed by Danida that evaluation is “faced with a set of challenges that are more pronounced than in the assessment of other types of development assistance” (Danida 2000, vol.1, p.10). Discussion of these challenges, informed by the examination and analysis of donor experience, is continued in Chapter 4, focusing on general evaluation methodology and issues of impact assessment.

### *General Methodology*

A recurrent theme in this research is the relative merits in this field of two competing methodological approaches, conventional evaluation and participatory evaluation. It is

argued that conventional evaluation, notably the logical framework approach, suffers from specific deficiencies and shortcomings, and, as an alternative, the suitability of a participatory approach is asserted. It is claimed that a participatory approach is congruent with the process of democratisation and offers particular advantages for addressing key challenges to DG evaluation.

It is noted that most donor agencies have attempted to address the dilemmas of evaluating DG assistance within the framework of conventional aid evaluation, although some have questioned its appropriateness. While a number of studies have a participatory element, this is generally limited to seeking stakeholders' perspectives as a source of data. An important distinction is between evaluations where stakeholders remain objects or become subjects of the process. Only two out of the sixty studies loosely met the three threshold criteria (Rebien 1996, pp.67-9) to be counted as a participatory evaluation.

Conventional evaluation methodology is conceived as based on traditional, scientific inquiry and relying mainly on quantitative methods. The logical framework approach (LFA) is the most clear-cut application of this conventional paradigm, with a number of agencies adopting logframe methodology, notably the European Commission, DfID, and Sida's study on the 'Evaluability of democracy and human rights projects' (Sida 2000b). Additionally, both CIDA and USAID have adopted 'results-based' approaches on an agency-wide basis, known as 'results-based management' (RBM) and 'managing for results' (MFR) respectively. Both RBM and MFR are closely related to the logical framework approach, given their mutual reliance on the *a priori* establishment of a matrix of immediate objectives and wider goals, along with objectively verifiable indicators (i.e. mainly quantitative) and means of verification. RBM and MFR are not evaluations in the sense of learning lessons, but orientated to the objective of agency accountability through an annual monitoring of short-term results.

The suitability of a conventional approach to evaluation in this field has been questioned in particular by Danida, noting "the inadequacy of conventional evaluation tools... [and that] evaluating efforts at political reform requires a different methodology (2000, vol.1, p.11). Similarly, UNDP has referred to the problems in applying existing quantitative-based methodologies and techniques and stated that "evaluation methodologies must be rethought" (UNDP 1998a, p.27). Danida is critical of the logframe approach in particular, stating that LFA analysis and political analysis "are not particularly compatible" (vol.1, p.66).

Four main shortcomings of the logical framework approach are outlined in this research: its focus at project not programme level; its applicability to 'hard' not 'soft' data; its inward orientation to project objectives; its problematic emphasis on causality and quantitative indicators. Essentially, LFA is orientated to tracking progress to pre-established objectives and cannot capture the dynamic political context in which democracy assistance activities are embedded. The logframe approach tends to assume the idea of progress, and is specifically designed to accompany this process of positive change. Two problems emerge, however. One is that democratisation is not a linear process of positive and gradual change, rather it is an irregular process following a non-linear pattern, with progressions and regressions. In the case of regressions, LFA becomes obsolete. The second problem is that the combination of inward orientation and

assumed progress means that LFA is unable to countenance negative, unintended effects of DG assistance.

In sum, such conventional evaluation approaches are rejected as unsuitable for evaluating democracy and governance programmes. Inappropriately, they offer a technical solution to a political problem. It is concluded that evaluating democracy and governance assistance is more art than science.

### *Assessing Impact*

Impact evaluation is crucial in the political aid field, pertaining to an assessment of the distinct role played by external actors in democratisation processes. From donor experience, it is also a challenging task. The most evident difficulty is that of ‘attribution’, i.e. establishing linkages between DG assistance and political changes. A number of issues emerge from the review of donor experience in conducting impact assessment. These are summarised as follows.

- *The significance of political context* as an important determinant on programme impact, and therefore the need for a background study analysing patterns of political change at the national level.
- *The multiplicity of actors and factors in complex political change* and the difficulties of differentiating the contribution of a single actor. There are difficulties in distinguishing the contribution of internal and external actors, as well as in separating out one donor from others.
- *The phenomenon of the ‘missing middle’*, requiring at times an ‘act of faith’ to leap from micro level outputs to such macro level objectives as ‘greater respect for human rights’.
- *With and without scenarios and issues of counterfactualty*. Are external actors being credited for developments that would have happened anyway, without their assistance?
- *External – internal relationships*. In partially attributing perceived (macro level) developments to the activities of external actors, have the interrelationships between internal and external actors been sufficiently addressed? External efforts may be dependent on local support. Alternatively, countervailing forces in the particular country may undermine external actions.
- *Time-scale*. How possible is it to evaluate the impact of projects and programmes that have only recently been completed, given that democratic change is a long-term process?
- *Unintended impact*. External intervention involves a dynamic, inter-active process and can have unintended side effects. Does the search for positive impact ignore the possibility of such negative impact?

One consequence of a lack of attention to these difficulties has been overoptimistic and/or exaggerated claims concerning the effects and impact of particular programmes, with limited linkages between direct outcomes of assistance and alleged impact. At times

there is a lack of methodological clarity on how impact assessments are made, and therefore it is concluded that assumptions of positive impact are made and reported, despite a lack of evidence to substantiate such claims.

Undoubtedly assessing impact poses very real challenges, yet these should not be regarded as insuperable. Attempts to demonstrate causality, especially through quantitative indicators, are rejected, however. Given the complex nature of democratisation, the establishment of *plausible linkages* between donor interventions and political change may be the best that can be hoped for. In a recent switch to a more qualitative approach, USAID (2000k) has developed a methodology for exploring the impact of DG programmes on democratisation processes at country level, with the publication of three pilot studies anticipated in 2002. While this is a most significant initiative, two shortcomings of the methodology were identified: the omission of exploring unintended negative impact, and the lack of a participatory approach. Further, it is proposed that a multi-donor study would be a more effective and efficient means of undertaking a country impact assessment. The Mali report (Club du Sahel *et al.* 1999) on overall aid effectiveness provides the only example of a collaborative venture, not specific to DG assistance, but with some useful lessons in terms of local involvement.

### *An Alternative Methodology*

The methodology proposed is specifically intended as a means to conduct impact evaluation at country level. It consists of three main elements, each considered in turn:

- Political context studies: i.e. background political analysis of the particular country context and the trends towards (or regression from) democratisation.
- A three-fold linkage between micro, meso and macro levels as an appropriate means for achieving impact evaluation.
- A participatory approach to evaluation, highlighting the perspectives of domestic actors, both governmental and non-governmental, on external efforts.

A clear and recurrent point to emerge from the examination of donor reports was the importance of DG evaluation being embedded in the political context within which democracy assistance was undertaken. This is particularly significant given that the (continually changing) context is itself a key determinant of the success or otherwise of donor interventions, as well as a key indicator of the opportunities for and constraints upon such assistance. The shortcomings of those evaluation reports that lacked the basis for such political analysis were noted. Political analysis would concentrate on democratisation trends over the specified time period, in both progressive and regressive directions, as well as identifying the major areas of change at both regime and partial regime levels. Political context studies would partly fulfil the role of a baseline study. Completion of the study would enable analysis of the extent to which the DG programme under examination has focused on those areas crucial to democratisation or not. Local academic specialists, potentially making use of democratic audit methodology (Beetham *et al.* 2001), as well as surveys of academic literature would undertake political context studies.



One conclusion of discussions of impact evaluation was to highlight the significance of meso level analysis. As proposed by Schmitter and Brouwer (1999), this can provide the 'missing link' between micro- and macro-level analysis. Evidence of micro level outcomes and impact can contribute to broader pictures of meso level impact. In turn, macro level analysis can be based more credibly on impact analyses at a range of partial regimes (ibid.), enabling evaluators to make more *plausible* connections between external support and overall political change. Additionally, meso level analysis is significant in itself for comparative studies of donor impact at the thematic or partial regime level.

In promoting a participatory approach to DG evaluation, this research is advocating 'genuine' participatory evaluation, involving the shift from participants as objects to subjects of the evaluation process. It is proposed that evaluation of external DG assistance be undertaken by a group of well-informed national actors, both government and non-governmental, and inclusive of non-recipients or and non-beneficiaries of assistance. It is claimed that a participatory approach is appropriate and beneficial for the following reasons.

First, local, pro-democratic actors, engaged in a collective process of dialogue and negotiation can most effectively address the challenging issues of evaluation identified in this research through reflection and analysis.

Second, the four characteristic features of participatory evaluation, defined by Estrella and Gaventa (1998, pp.17-27) as participation, learning, negotiation and flexibility, are all closely associated with the principles of democracy. Participation itself is clearly central to democratic processes. The democratic principles of popular control and political equality (Beetham 1999) are realised precisely through political participation, minimally in electoral processes, as well as more substantially through a variety of democratic practices. Negotiation, dialogue and compromise are central to democratic decision-making, and flexibility and learning are both held at a premium in democratic processes and practices.

Third, the process of participatory evaluation is akin to the process of democratisation itself. Democracy is constructed and crafted (and resisted and undermined) by various coalitions of domestic actors and interest groups, inclusive of elite groups. Frequently, it is local knowledge of rapidly changing regime circumstances and the alliances between national, pro-democratic actors that are central to processes of democratic transition and consolidation. Participatory evaluation provides the key input of local knowledge and analysis that is essential to an evaluation that provides a truly critical examination of external activities. The strengths and limitations of donor activities are examined in the context of the prospects for and constraints upon sustained democratisation in the particular country context. The outcome of negotiation and consensus-seeking in the evaluation process entails learning for all. For donor agencies, their past and current efforts are subjected to critical reflection and appraisal, enabling objectives to be revisited and strategies refined, informed by internal perspectives. For participant evaluators, knowledge of processes of political change is enhanced, in turn informing and strengthening local action for democratic change. Thus, the act of evaluation becomes an act of democratisation.

# Chapter 1.

## INTRODUCTION

The 1990s saw the emergence of the promotion of democracy and good governance as a major field of activity within international development co-operation, with significant aid funds now allocated to this area. Yet what has democracy and governance (DG) assistance achieved? After a decade of such activities, the question of *evaluation* has become a crucial one. Evaluation is essential to assess the achievements of DG assistance as well as to learn lessons and improve its future effectiveness.

However, evaluation of democracy and governance assistance poses considerable challenges. In a previous ESCOR study, Mark Robinson noted that “donors lack a systematic approach to evaluation [of democracy assistance and political aid] and there is no generally accepted methodology or set of indicators” (1996, p.ii). Developments since then have included a recent spate of evaluation studies, yet methodological questions remain unresolved. Two main issues are: the specification of criteria and operational indicators for judging success; the establishment of causal linkages between democracy and governance assistance and democratic changes (Carothers 1999, Schmitter and Brouwer 1999). Such difficulties are not insurmountable, though, and this research aims to contribute to their resolution. The research has two related aims. One is to provide a review and critique of evaluation studies in this field, focusing on methodologies, and highlighting both insights and shortcomings. The other is to develop an appropriate methodology in order to conduct impact evaluations in country contexts. In this respect, this report is intended as a methodological study, rather than simply summarising the various findings from a range of programme and project evaluations. Its main argument is to question the appropriateness of conventional aid evaluation methodology for evaluating democracy and governance assistance. In particular, the logical framework approach is subjected to criticism, with a participatory methodology proposed as not only offering a more suitable approach to evaluation in this context, but also one that is synergistic with the process of democratisation itself.

### *Introduction to Report*

The report is in six chapters. Following this introduction, there are four substantive chapters, plus a brief concluding chapter. Chapters 2 to 4 address the first aim of conducting a review and critique of evaluation studies in this field, while Chapter 5 develops an alternative methodology for country impact studies. Despite a focus on evaluations at the programme level, not project evaluations, the survey of bilateral and multilateral donors discovered that many more evaluations, and related studies, had been undertaken than anticipated, with one hundred and ten references catalogued. In consequence, only the experience of *key* donors is explored in Chapter 2, that is, those that have given greatest attention to issues of evaluating democracy, human rights and

governance assistance. Thus the evaluation reports and other related studies of seven bilateral and multilateral donors are examined in detail in Chapter 2, while other donor studies are covered in Appendix 1. Chapter 3 extends the coverage to embrace a total of sixty evaluation studies from the full range of donors. These are included in the research database, enabling a comparative analysis of their main characteristics and findings. Chapter 4 summarises and discusses the various difficulties and challenges in evaluating DG assistance that have emerged from the preceding two chapters, focusing on issues of general evaluation methodology and of impact evaluation. The appropriateness of conventional evaluation methodology in the DG field is questioned, including a critique of the logical framework approach. Chapter Five responds to such challenges by proposing an alternative approach to evaluation in this field, based on a participatory methodology. Finally, a brief concluding chapter is provided.

### *Introduction to Aid Evaluation*

Given that this report focuses on methodological questions and that its main argument concerns the relative merits of different aid evaluation methodologies, a brief introduction to aid evaluation in general, and to logical framework analysis in particular, is required as background information prior to examination of donor agency evaluation reports in Chapter 2.

In recent aid evaluation literature, two broad methodological approaches have been distinguished, traditional or conventional evaluation versus participatory evaluation (Cracknell 2000). Such contrasting approaches relate closely to the debate between positivist and constructivist methodologies within general evaluation literature, initiated by Guba & Lincoln's (1989) critique of the former and advocacy of the latter. Characteristics of conventional and participatory approaches to aid evaluation are as follows:

Conventional evaluation:

- rooted in positivist tradition of social science in which natural science methods are applied
- objectivity and value-free nature claimed
- cause and effect linkages identified
- management tool, concerned with aid effectiveness
- quantitative-based

Participatory evaluation:

- more critical and questioning (including motives of donors)
- issues of power addressed

- participation in evaluation, emphasising ‘negotiation’ of the whole evaluation process with participants / stakeholders
- (subjective) views and judgements of participants valued
- more process-oriented
- qualitative methods emphasised

One key issue is the extent to which these are alternative or complementary approaches. In general terms, Guba and Lincoln’s (1989, p.17) view is that “no accommodation is possible between positivist and constructivist belief systems”. In contrast, Cracknell (2000) and Rebien (1996) consider that the two approaches are compatible, with enthusiasm for a pluralist approach (Cracknell 2000, p.350). In practice, a mixture of the two occurs. As with the concept of participation, donor rhetoric has recently embraced participatory evaluation to an extent, though essentially as a means of improving data collection rather than as a full-blown alternative methodology.

Within aid evaluation, logical framework analysis (LFA) has been the dominant approach. Along with the closely related ‘results-based’ approach, LFA has been applied to the democracy and governance sector by a number of donor agencies, thus requiring its brief introduction here. LFA entails the application of a ‘conventional’ approach to evaluation, based on positivist social science, with evaluation methodology based on a theory of causality (Pawson and Tilley 1997, pp.4-8; Rebien 1996, pp.19-21). The logframe approach rests on the tracing of causal connections between project inputs (or activities), outputs and objectives, with the latter divided into immediate objectives (or project purpose) and wider objectives (or programme goal). A logical framework matrix is initially prepared at the design stage of project cycle management, inclusive of the above dimensions, plus performance indicators and their means of verification, along with a statement of the risks and assumptions involved (Cracknell 2000, pp.108-12). This then provides a means of monitoring and evaluating progress towards achievement of stated objectives, one that is essentially quantitative in nature, although the use of qualitative indicators is allowable. LFA provides a seemingly rigorous evaluation methodology, though limited to tracing the realisation of declared objectives. It is generally acknowledged, however, including amongst supporters of LFA, that it is best suited to evaluating lower level project objectives and “less well adapted to tracking performance of programmes and policies at a higher level” (Cracknell 2000, p.116).

Despite coming under increasing challenge in the 1990s from a participatory approach (Rebien 1996, Cracknell 2000), the logical framework approach continues to be popular amongst development aid agencies. But is it a suitable methodology for evaluating DG assistance? Chapter 2 includes an examination of logframe-related evaluation to the DG sphere by four agencies, CIDA, USAID, Sida and the European Commission, as well as comments on its limitations. Further, its appropriateness is questioned more comprehensively in Chapter 4.

## Chapter 2.

# DEMOCRACY AND GOVERNANCE EVALUATIONS: DONOR EXPERIENCE

### *Survey methods*

A survey was conducted in early 2001 of all OECD Development Assistance Committee (DAC) members with respect to evaluations undertaken of democracy, human rights and good governance programmes. Bilateral aid agencies were contacted, including evaluation departments, as well as foreign ministries in many instances. Multilateral organisations were similarly contacted, for instance the European Commission, the United Nations Development Programme (UNDP) and the World Bank. The response rate was somewhat disappointing, with one agency taking six months to respond, while nine agencies failed to respond at all or provided very limited information, (see Appendix Two). Other agencies, for instance USAID, CIDA and Danida, provided full accounts of their activities.

The Internet was an important resource, with the majority of reports used for this study drawn from agency websites. Nevertheless, considerable differences emerged between agencies in this respect, with some providing little or no on-line access to evaluation material (among the larger donors, DfID and Agence Française de Développement stand out in this regard), while CIDA, Danida, the European Commission, NORAD, Sida, UNDP, and USAID make considerably more effort at practising greater openness and transparency.

By such means, a bibliography of some 110 reports was built up, far surpassing initial expectations. Our interest was essentially limited to programme evaluations as distinct from project evaluations. Nevertheless, where individual donor agencies had not undertaken an evaluation study at the programme level, but did send project evaluations, then such reports were generally included in the bibliography. Out of all respondents, a relatively small number of donor agencies emerged as having given considerable attention to issues of evaluation in this field. The evaluation studies of these agencies are examined below in this chapter, comprising four bilateral agencies, (CIDA, Danida, Sida, USAID), two multilateral agencies, (the European Commission and UNDP), plus the work of the Development Assistance Committee itself. The reports of other respondents are outlined in Appendix 1.

Examining the work of donor agencies individually may create the impression, potentially a misleading one, of institutional separation. It is claimed, however, that such relative isolation reflects the actual situation, with little evidence in evaluation reports of cross-agency learning, excepting within the DAC in the period from 1993 to 1997. Each agency's studies contain virtually no references to those of other donors.

## **2.1 Canadian International Development Agency (CIDA)**

'Human rights, democracy and good governance' constitute one of six priorities for Canadian development assistance, considered important in their own right as well as integral to the goals of sustainable development. Since 1992, there has been a 'Human Rights, Democratic Development and Good Governance' division within CIDA, part of the policy branch. Significant funds have been expended in this area, inclusive of monies channeled through international NGOs. (Helmich and Borghese 1998, Annex I). One notable shift is from the initial emphasis on 'human rights and democratic development' to the more recent focus on 'governance'.

CIDA's attention to evaluation work in this area can be divided into two distinct activities. One consists of a number of specific evaluation studies of democracy and governance (DG) programmes, mainly undertaken at country level. The other entails general discussions of evaluation methodology in this field, notably the implications for the DG sector of the agency-wide adoption of the 'Results-Based Management' (RBM) system as the organising framework for aid evaluation, as well as some consideration of alternative approaches. The two sets of reports and documents are examined in turn.

### *2.1.1 Country reports*

Three reports are included here. Initially, a summary volume of five evaluation studies undertaken in the mid-1990s was published (CIDA 1996). Subsequently, in the late 1990s, two further studies were undertaken of Canadian activities, by then more commonly termed 'governance programmes', in South Africa (Sutherland 1999) and in Ethiopia (CIDA 2000b). These three reports are examined in turn.

#### *Democracy and Human Rights: What are we learning? (CIDA 1996)*

This is itself a summary of the lessons learned from five evaluation-type studies undertaken in the mid-1990s. The five studies comprised a mixture of project and programme evaluations, mostly country cases, as follows:

- the democracy programme in El Salvador (CIDA 1994a);
- aid and diplomacy in South Africa - looking at the totality of Canadian government policy (CIDA 1995c);
- a study of CIDA's bilateral human rights programmes and projects (CIDA 1994b);
- the democracy and human rights programme in West Africa (CIDA 1995a);
- the human rights programme in Sri Lanka (CIDA 1995b).

The lessons learned for future programming were five-fold:

1. The indispensability of '*solid field knowledge*', with programmes to be based on 'strong contextual analysis'. In other words, a sophisticated understanding of the political context is required, itself based on local knowledge and expertise. The importance of dialogue with *all* stakeholders is emphasised, especially consultation with local NGOs. The Sri Lankan report itself is valuable for its

- emphasis on the socio-political context as a basis for evaluating the relevance and effectiveness of the Canadian programme.
2. The significance of policy and programme '*co-ordination and coherence*' is noted, both within Canadian government activities and between donors. The value of a division of labour between donors is noted, each focusing on a limited number of issues.
  3. The establishment of '*sound programmes*' is recommended, based on a long-term commitment by the Canadian government, supporting both government institutions and NGOs, with the specific targeting of women's organisations as essential. Programmes may be either long- or short-term, however, as appropriate to the particular political situation.
  4. Specific comments on '*assessment and evaluation*' are three-fold:
    - the inadequacy of basing project evaluations merely on quantified results;
    - the value of self-evaluations by local organisations, especially where external evaluations are not possible;
    - evaluation as a collaborative process between Canadian and local partners.
  5. The need for *capacity development* within CIDA, notably staff skills; and the need to develop programme guidelines and performance indicators for monitoring and evaluation purposes.

(CIDA 1996, pp.1-6)

Two of these lessons for future programming are highlighted as of particular interest here. First, the importance of an understanding of the political context is stressed, and, significantly, this is perceived as based on *local* knowledge and expertise. Second, the emphasis placed on dialogue with *all* stakeholders, especially local NGOs. The application of lessons learnt to future *programming* is encouraging, yet these points are equally applicable to evaluation methodology. Yet, it would appear that socio-political analysis and stakeholder views do not feature strongly in the evaluation studies themselves. All studies employ a traditional ex-post methodology, with evaluations undertaken through document reviews, fieldwork visits and interviews with selected staff. Only two reports (CIDA 1995a & 1995c) appear to have explicitly sought to include the views of stakeholders from the recipient country or community, the other three reports focusing more narrowly on CIDA staff. Further, the authors' approach to their interview material is consistently and surprisingly low-key, used to fill in background information rather than as a key means of gathering data on programme/project performance (e.g. CIDA 1995c, p.2).

A further question raised above concerns the suitability of quantitative-based evaluation methodology in this field. This is a recurring theme in this report, addressed below in the Ethiopia case study (CIDA 2000b), in CIDA-wide discussions about methodology, and in the studies of other agencies.

*Supporting Democracy: The South Africa-Canada Program on Governance (Sutherland 1999)*

The Program on Governance initially commissioned this study, subsequently published as a book, as an evaluation for CIDA (which funds the programme) and the International

Development Research Center (IDRC) (which manages it). Its stated purpose is “to record and analyze (evaluate) the accomplishments of the South Africa-Canada Program on Governance and the lessons learned” (Sutherland 1999, p.1). It is the shortcomings of the study itself that are highlighted here, however.

This report is more of a review of activities than an evaluation. The initial chapters provide an in-depth description of the Program on Governance (Chapters 1-4), written up from office files. The fieldwork component of the study is covered in Chapter 5, consisting of 36 interviews with partners and observers from South Africa (23) and Canada (13). This is not an impact evaluation, but rather an attempt at documenting achievements, bracketed under headings such as ‘satisfaction with experiences’ and ‘direct and indirect benefits’. Interviewees were also asked to set the programme in the context of other donor activity.

The limitations of this report relate to its consistently laudatory tone, leading one to suspect that it is insufficiently analytical or critical, or, alternatively, that the Program on Governance is truly remarkable. Its almost obsequious nature is evident from the following quotations:

- “South Africans use the Program on Governance because it is effective and useful: more than half the respondents... noted that no other foreign government planned these learning opportunities so carefully” (p.54).
- “The South Africans overwhelmingly reported being satisfied and often excited and encouraged by their work with Canadians through the Program on Governance. In the 21 practitioner-partner interviews with South Africans, only 2 had criticisms of the Program on Governance’s operations” (p.56)
- “At the level of program management, the Program on Governance simply excelled. It always followed up. It solidified contracts, formalised agreements, and always answered its phones” (p.57).

To be fair, the interview record does suggest that South African participants did share this very positive view of the programme. Nevertheless, one would have to query the interview schedule and the nature of questions posed that elicit such positive responses. Despite the terms of reference, there would appear to be no lessons to be learned, apart from ‘more of the same’. The function of such an evaluation that attempts little or no critical analysis is seriously questioned.

*Democracy & Governance Programming Lessons for CIDA: Ethiopia Case Study*  
(Rawkins and Qualman 2000)

The key findings from this study undertaken by Rawkins and Qualman for CIDA were quite country specific, though with some general applicability, as follows.

The requisite of a carefully worked out democracy and governance country strategy is emphasised, based on a joint needs assessment with host country partners. In this instance, a somewhat hastily implemented programme, responding to the Ethiopian government’s request for rapid assistance, had adverse effects on effectiveness and sustainability. It is felt that “governance programming in Ethiopia would have been more



effective if the Agency had had in place a DG policy and programming framework ...before launching into project assistance” (CIDA 2000b, p.5). Similarly, agency “disbursement pressures are particularly unhelpful in this field” (p.7), with substantial outlay of staff time required before a programme can be properly designed or financial resources usefully committed.

Echoing discussions in other agencies (e.g. Danida), the appropriateness of a quantitative-based, logical framework-related monitoring and evaluation system is questioned, in this case CIDA’s ‘Results-Based Management’ (RBM) system (see below). Generally the RBM approach is most suited to projects where concrete outputs can be achieved within a specific time span. In contrast, governance is regarded as a ‘soft’ area of programming in which institutional relationships and culture are the subject of reform, with time frames hard to predict. Thus, minimally, Rawkins and Qualman argue for greater flexibility in the use of the RBM system, though a more fundamental question is whether such monitoring and evaluation systems are appropriate at all in the democracy and governance (DG) field.

Interestingly, this study reiterates a number of concerns identified in the synthesis report discussed above (CIDA 1996), suggesting that lessons had not in fact been learned. Echoing the earlier report’s emphasis on ‘field knowledge’ and ‘contextual analysis’, Rawkins and Qualman stress “information gathering and analysis” as the “single most important factor for effective and efficient DG programming (p.2). Additionally, donor co-ordination remains less than adequate. The authors find that “there has been some co-operation on governance issues” in Ethiopia, but also “some unwillingness to share sensitive information... and some resistance to working together very closely” (p.4). Finally, the earlier study’s questioning of evaluation based on quantifiable results is reaffirmed by the later report’s critique of the RBM framework. The lesson learning function of evaluation would appear to have been unsuccessful in this instance, despite the helpful synthesis of earlier lessons (i.e. CIDA 1996).

### *2.1.2 Results-Based Management and Evaluation Methodology*

CIDA’s agency-wide adoption in 1996 of ‘Results-Based Management’ (RBM) had far-reaching and problematical implications for the evaluation of democracy and governance programmes. In this section, nature of the results-based system is first outlined, followed by consideration of how performance indicators have been developed to track progress towards results in democracy and human rights. Finally, two reports question the appropriateness of RBM in the DG sphere.

The introduction of RBM in 1996 entailed the reverse of the recommendation made by the simultaneously published synthesis report on evaluations of democracy and human rights programmes (CIDA 1996), with RBM involving greater rather than less reliance on quantifiable results. A ‘result’ is defined as “a describable or measurable change ... that is derived from a cause and effect relationship” (CIDA 2000a, p.12). In other words, results are changes that are attributable to resource inputs. In the democracy and governance sphere, this is profoundly problematic. Oblivious to such issues, at least initially, the institutionalisation of RBM has led to the development of a general evaluation methodology that focuses on ‘results’, through the establishment of causal

linkages, and that is perceived as ‘impartial and objective’, with verification based mainly on quantifiable performance indicators. Key to this process are the concepts of ‘Performance Framework’ (PF) and ‘Performance Measurement Framework’ (PMF), similar to a logical framework approach. Evaluability is built-in to the project cycle through a hierarchy of objectives (outputs, outcomes and impacts) and a framework of expected results, with evaluation as an assessment of the achievement of such results. PF provides the anticipated cause and effect relationships from the level of activities (inputs) upwards to strategic goals, including assumptions and risk assessments, while PMF provides a systematic plan for measurement and verification through (mainly quantitative) performance indicators and data collection requirements. (CIDA 2000a, p.32). A participatory element is included, and indeed is said to be essential to a ‘good’ evaluation (p.26). It is arguable, however, that participation is limited to mere stakeholder consultation as distinct from a more genuine participatory approach: “Evaluations are expected to be participatory, meaning that evaluators will consult stakeholders to identify the extent of their involvement” (p.33).

*Indicators for Programming in Human Rights and Democratic Development* (Kapoor 1996)

The introduction of RBM led to a ‘preliminary study’ on the development and use of performance indicators to track progress towards results in democracy and human rights (Kapoor 1996). Kapoor’s study focuses on three main areas. He reviews the development of indicators by other donor agencies, discusses methodological challenges of measuring political development, and provides an illustrative list of possible indicators in different sub-sectors.

The review of other donors concludes that “the development of performance indicators for human rights and democratic development either has not happened or is at best at an incipient stage” (ibid., p.1), with one exception, that of USAID. Subsequently, the significance of USAID’s work in this area has been confirmed with the publication of its substantial *Handbook of Democracy and Governance Program Indicators* (USAID 1998a) (discussed below).

The challenges associated with measuring political development are perceived as three-fold. First, given that there is “no comprehensive or ‘objective’ theory / model of democracy or human rights” and that existing models tend to be ‘Western-based’, there is a danger of imposing “a foreign and inappropriate yardstick” ignoring “contextually- and culturally-specific notions” of justice, women’s rights, participation, and so on, (Kapoor 1996, pp.5-6). One possible way out of this dilemma identified by Kapoor is to tailor indicators to specific situations and to appropriate levels of political development, achievable in his view through participatory methods. A second challenge pertains to the small size of donor programmes relative to the magnitude and complexity of political change in any country. This raises the problem of *attribution*, that is the difficulty of separating out the effects of donor activities from those of numerous other factors that influence political change processes, in turn making attribution of causality to donor actions a hazardous business. The issue of ‘attribution’ is a recurrent one in this research. Another related issue concerns the constraints on the pace of change due to the particular

political context, restricting what is possible by both donors and internal agencies. Kapoor's contribution here is that the goals set by donor programmes should be modest, context-bound and measurable (p.14). A third challenge identified, but with no solution offered, concerns the generally slow and incremental nature of political change. This implies that meaningful outcomes are generally only measurable in the intermediate and long-term, whilst agency pressures are often for quick, short-term results.

Different indicators are examined, with three main types identified, quantitative, qualitative and participatory. The notion of participatory indicators is of particular interest, considered by Kapoor as highly relevant to overcoming 'Western-imposed yardsticks'. He discusses the advantages and limitations of participatory indicators, more attuned to local context, though also context-bound and less amenable to generalisation on the other. One suggested method for the formation of participatory indicators is through 'group workshops', potentially scheduled for 3-4 days with 30-40 people (p.14). This issue of participatory indicators is discussed further in Chapter 5 below. Finally, Kapoor provides a list of illustrative indicators for different sub-sectors, comprising both quantitative and qualitative indicators. Those for 'Legislative Assistance', including suggested means of verification, are given as an example in Box 2-1.

Box 2-1: Human Rights and Democratic Development Indicators - Illustrative List

**(ii) Legislative Assistance**

- strengthening of/steps towards democratically elected federal/provincial/local legislatures:
  - increased # of functioning legislatures/legislative committees
  - more timely legislative work
  - greater legislative control over government decisions/budgets/appointments, etc.
- increased laws drafted/enacted by legislatures (# of bills, hearings, etc.) regarding civil and political rights, and socio-economic and cultural rights (see below)
- increased ratification and implementation of international human rights treaties
- regular government assessments (in collaboration with NGOs) of country's human rights situation and submissions to UN human rights treaty bodies (Commission on Human Rights, Committee on Economic, Social and Cultural Rights, etc.).
- greater parliamentary/legislative consultation with citizens:
  - parliamentary commissions
  - # of citizen groups being consulted
  - survey of citizen/NGO opinions on draft laws
  - % of citizens/NGOs who believe they are being represented by/have access to MPs/MLAs
- increased opposition party power and independence in legislature:
  - extent of debate
  - review of legislation
  - inclusion on legislative committees

Source: Kapoor 1996, p.24

Notwithstanding these attempts by Kapoor to broaden the range of indicators to include qualitative ones developed through participatory processes, the appropriateness of RBM-type evaluation to the field of democracy and governance assistance remains very questionable. By definition, ‘results’ are measurable cause and effect changes. Yet such attribution is not only difficult, but also is the exact issue that requires problematisation, given the multiplicity of factors integral to processes of democratisation. Such issues have been the subject of discussion, however, in an *Evaluating Governance Programs* Workshop held in Ottawa in April 1999.

#### *Evaluating Governance Programs* (IDRC 1999)

The ‘Evaluating Governance Programs’ Workshop was hosted by the International Development Research Center (IDRC) and attended by CIDA personnel, including the Director of Evaluation. The report does not assert a single perspective, reflecting the differing views of those in attendance, and includes voices critical of the RBM-type approach. One clear statement is that, “The notion of causality in governance programming was rejected; there are too many variables at play. Evaluation of governance work is not scientific *per se*” (IDRC 1999, p.6). General discussion of RBA and Logical Framework Analysis remains more ambiguous, however. On the one hand, “Traditional evaluation approaches which demand the application of the same tool and logic model to all initiatives were rejected as irrelevant and possibly destructive” (ibid., p.8). Additionally, it was noted that some contemporary evaluation tools “(such as Logical Framework Analysis – LFA, Results Based Management – RBM, and Indicators-based studies)” were developed for ‘blueprint-type projects’ and “and have little fit with complex and iterative, governance program agendas” (ibid.). On the other hand, it was suggested that the tools were less the problem than their application, with merely their modification and customisation required to meet diverse settings.

Alternative methodological approaches were also discussed, notably a participatory approach to governance evaluation, contributed by Ilan Kapoor in the ‘Background Paper and Literature Review’. This noted the recent movement towards the adoption of a participatory approach to evaluation within international development organisations, while recognising that this remained a relatively new and uneven process, especially within donor agencies, with little application to governance assistance. The instruments of a participatory evaluation are outlined, as well as the strengths and constraints of a participatory approach. These issues are discussed further in Chapter 5.

#### *Sector Wide Approaches, Accountability and CIDA* (Schacter 2001)

A related critique of RBM, especially its insistence on cause and effect links, stems from a recent paper for CIDA on Sector Wide Approaches (SWAPs), (Schacter 2001). The author, Mark Schacter, claims that the current RBM evaluation methodology may be suited to old-style development projects, but not to new SWAPs, again raising the issue of attribution. With SWAPs it is generally *impossible* to draw direct causal links when CIDA’s input is only one of many to a sector wide strategy. A similar argument can be applied to democracy and governance assistance, where not only is an individual donor

one of many, but also the role of international actors is only one of a multiplicity of factors in complex processes of political change. Schacter argues for the current RBM system to be adjusted in two respects. One is to recognise that meaningful results mostly become apparent in the medium- to long-term, and that annual reporting requirements often cannot show such results. The other is to focus on the *logic* of interventions in relation to desired development results. The challenge here is three-fold: first, to provide evidence regarding the direction of change (positive or negative); second, to draw links between short-term developments and longer-term goals; and third, to show the role that CIDA plays (Schacter 2001, pp.7-14). Such suggestions may provide some guidance on dealing with the attributional problems inherent in evaluating democracy and governance assistance.

### *Summary*

CIDA is one of a relatively small number of donor agencies to have given considerable attention to issues of evaluation in this field. Such discussions have been undermined, however, by the agency-wide adoption of RBM as the organising framework for evaluation procedures. Although the appropriateness of RBM for the evaluation of democracy and governance programmes has been seriously questioned in discussion papers written for CIDA, the institutionalisation of RBM indicates that there is little prospect of alternative approaches being introduced at the present time.

Nevertheless, discussions have raised a number of key issues, as follows:

- The purpose of evaluation: to learn lessons and improve programming (CIDA 1996), or a managerial exercise to show results and demonstrate accountable use of public funds;
- Problems of impact evaluation and issues of causality and attribution;
- The appropriateness of a quantitative or a qualitative focus in evaluation methods;
- The introduction of participatory evaluation methods, including the development of participatory indicators (Kapoor 1996 & IDRC 1999);
- A lack of critical analysis, with evaluation studies merely serving to corroborating donor activities (Sutherland 1999);
- A failure to effectively learn and apply earlier lessons (CIDA 1996 and Rawkins and Qualman 2000).

### **2.2 Danish International Development Assistance (*Danida*)**

Human rights and democracy is described as “central in Denmark’s development co-operation”, with significant programmes in each of the 20 selected countries to which Danish aid is concentrated (Helmich and Borghese 1998, Annex I). Danida has contributed significantly to addressing evaluation issues by undertaking the single, most comprehensive evaluation by any donor agency, inclusive of a critical examination of evaluation methodology.

### *2.2.1 Evaluation of Danish Support to Promotion of Human Rights and Democratisation 1990-1998 (Danish Ministry of Foreign Affairs/Danida 2000 - 9 volumes.)*

A large-scale evaluation was undertaken in the late 1990s of Danish support to the promotion of human rights and democratisation from 1990 to 1998, resulting in a nine volume study published in early 2000. The synthesis (volume 1) was based on four thematic studies, (volumes 2 to 5: justice, constitution and legislation, elections, media, participation and empowerment), and four country case studies, (volumes 6 to 9: Ghana, Guatemala, Mozambique and Nepal). It is described as a 'lessons learned' evaluation, with the emphasis on "self-critical learning, rather than on accountability" (Danish MFA/Danida 2000, vol.1, p.1). Stated objectives were three-fold:

1. to present an overview of Danish-financed activities supporting the promotion of human rights and democracy, 1990-98;
  2. to document experience from the identification, planning, implementation and monitoring of activities in assisting in such promotion;
  3. to assess if and how Danish-financed activities have promoted democratisation and human rights;
- (vol.1, p.1).

Discussion here focuses on ways in which this study has addressed the issues pertinent to our enquiry, those of methodology and impact evaluation, rather than the study's specific findings with regard to Danish assistance.

#### *Methodology*

The study includes a useful discussion of evaluation methodology, both a critique of the application of traditional techniques and some suggestions for alternative approaches. Political aid is seen as not only new but also different from other development aid. One essential difference is that democratisation is "deeply embedded in a political process which is the outcome of a battle between contending social forces" (vol.1, p.10). Evaluation is therefore "faced with a set of challenges that are more pronounced than in the assessment of other types of development assistance" (ibid.). Given that such challenges go to the nub of our enquiry here, they are worth quoting in full:

- Varying perspectives on the subject matter: democracy means different things to different people.
- Different time horizons: agencies supporting democratisation want to see results quickly, whereas changing and, above all, consolidating regimes takes time.
- Conflictual nature of political reform: democratisation is likely to be uneven and uncertain – it is not a linear process.
- Limited value of a project perspective: building or reforming political institutions as well as empowering citizens are not single and isolated events but part of a longer and broader political process.
- Involvement of funding agencies: for various reasons, they are pulled into the reform process and thus unable to adopt an objective and detached position.

- The inadequacy of conventional evaluation tools: because of the weakness, if not absence, of objective indicators and ‘hard’ data, evaluating efforts at political reform requires a different methodology.

(vol.1, pp.10-11).

Consequently, conventional approaches to evaluation are subject to some critique, in particular the logical framework approach (LFA). Such goal-oriented types of evaluation face three problems in this field:

- insufficient baseline material that is reliable and definite;
- the many different ways of reaching the same goal, i.e. democratisation and protection and promotion of human rights;
- the difficulty of identifying the specific difference a given project activity may have made, especially any links between a single project and changes in the overall human rights and democracy situation;

(vol.1, p.11).

While it is noted that LFA has been recommended for use in the human rights and democracy field by the OECD DAC (though no reference provided, and this research has not come across such a recommendation), this study submits that LFA analysis and political analysis “are not particularly compatible” (vol.1, p.66). Even a participatory approach to applying LFA is not satisfactory as it “cannot anticipate and capture the political dynamics in which local actors will make their decisions” (ibid.). In other words, a logical framework approach can neither impart nor take into account the (constantly changing) political context in which projects and programmes are embedded. In the view of the authors, “the promotion of democratisation and human rights is too complex an activity to be subjected to a mere set of instrumentalist criteria”, and they encourage Danida “to experiment with other methods of monitoring and evaluating its work in the HR&D area” (vol.1, p.68). Unfortunately, there is limited discussion of what such alternative methods could entail, aside from reflecting the “qualitative nature of most democratisation activities” (vol.1, p.68).

Can more insight be gained from the methodology used by this evaluation study itself? The methodology adopted is described as “‘inter-subjective validation’, combined with a process-orientated approach” (vol.1, p.11). All case-studies are said to have adopted this methodology, though with flexibility encouraged. Three issues arise, however. First, the nature of this methodology remains rather imprecise and unspecific. ‘Inter-subjective validation’ appears similar to a ‘user-based’ or ‘participatory’ approach. It addresses problems of attribution through seeking the views of various stakeholders, with conclusions drawn on the basis of “their assessment of the legitimacy, relevance, management and cost-effectiveness of reform programmes” (vol.1, p.11). It is perhaps noteworthy that the ‘impact’ of programmes is not included on this list of evaluation criteria. Second, ‘inter-subjective validation’ is not itself critically examined regarding potential methodological weaknesses or problems. One possible criticism concerns the inclusion and exclusion of stakeholders. Although the terms of reference state that the

evaluations are to be conducted as a “participatory process” (vol.1, p.120), participation appears limited to interviews with ‘key stakeholders’, that is the main project beneficiaries rather than with a wider range of pertinent individuals and commentators. Positive validation is more likely to be generated from programme beneficiaries, given some disinclination to ‘bite the hand that feeds them’. Third, perhaps due to the relatively imprecise outline of the methodology of ‘inter-subjective validation’, there are not only significant differences between studies, but also the Guatemalan team states somewhat surprisingly that it was “largely free to devise its own methodology” (vol. 7, p.3).

The individual studies also comment on the general methodological difficulties faced in this area, again with negative observations about the appropriateness of traditional evaluation methodology, notably LFA. For instance, the Guatemala study notes that “all donors seem to be struggling with the same methodological problems and difficulties in terms of reliable measurement of the impact of their support to human rights and democracy” (vol. 7, p.67). The logframe approach is perceived as containing a “number of inadequacies”, and “generally not seen as useful for the more qualitative and process-oriented aspects” (vol. 7, p.67). Other studies also note the unsuitability of LFA in circumstances where donors have to respond quickly (vol. 3, p.29), and question its reliability where a “highly politicised setting” affects partner organisations’ ability to contend with other forces and political realities (vol. 9, p.62). In other words, the logframe matrix is inadequate in itself as a basis for evaluation due to the wider political context in which projects and programmes are designed and implemented. Changes and developments in this context can often be a major factor in determining success or otherwise.

## Methods

Table 2-1: Research methods in the Danida thematic studies

	Desk study only	Desk study and Fieldwork
Justice, constitution and legislation (volume 2)	Botswana*, Malawi*, South Africa (2 case studies)*, Transparency International*, International Commission of Jurists*, Penal Reform International*	Uganda, Rwanda, Vietnam, Cambodia
Elections (volume 3)	South Africa, Kenya, International Humanitarian Service	Nicaragua, Tanzania, Uganda
Media (volume 4)	Regional Media Resource Centres (Nepal); the Media Fund (Nepal); Radio CSF (Peru); El Regional (Guatemala); Bhutan Broadcasting Services; Kuensel (Bhutan); Regional TV Station (Vietnam)	University of Zambia; Press Association of Zambia; Media Institute of Southern Africa; The Nordic SADC School of Journalism; Palestinian Broadcasting Corporation
Participation and Empowerment (volume 5)	South Africa, Indonesia, Gaza/West Bank	Bolivia, Kenya

\*Denotes desk study focused on selected projects, rather than wider country programme.

Source: *Danish Ministry of Foreign Affairs/Danida 2000 – volumes 2-5*



Switching from methodology to methods, the eight volumes adopted a common approach involving a two-phase study of desk work and field work. For the country case-studies, an initial desk study based on relevant documentation is followed by the fieldwork phase, usually of two to three weeks duration, enabling more detailed investigation of selected projects in country context. The thematic studies typically entailed a desk study review of a significant number of country programmes, more than subsequently visited for field research. Only one study (vol. 3 on elections) limited its preparatory desk work to those four countries it intended to visit. For the remaining three studies, field work was undertaken in between two and six countries, with programmes in a further three to seven countries reviewed as desk studies only, as the Table 2-1 above illustrates.

Each thematic study also states that the authors sought to incorporate findings made by the evaluation teams studying Danida country programmes (i.e. volumes 6 to 9) in their conclusions.

### *Impact Evaluation*

One stated objective was to assess whether and how Danish-financed activities have promoted democratisation and human rights, amounting to an impact evaluation. Of particular interest here is *how* this objective is pursued and at what level, that is, at the micro level of project impact or the macro level of national political change. It is intriguing that the eight thematic and country studies have addressed this objective in varied ways, with most notable differences between the four country studies. While the Guatemala and Mozambique cases emphasise the problems of undertaking an impact assessment, in contrast, the Ghana and Nepal studies do attempt to provide impact evaluations. The thematic studies are also characterised by differences in approach to this key question of impact.

The Guatemala case study frankly states that it “makes very little effort to assess the impact of the specifically Danish assistance on the overall human rights situation” (vol. 7, p.55). This is due to two well rehearsed difficulties. One is the problem of singling out Danish assistance from that of other donors, and the other is the “widely acknowledged” difficulty of establishing “a direct causal linkage between a project intervention and the broader human rights situation in any country” (vol. 7, p.55). Consequently, the Guatemala case limits itself to identifying more or less successful aspects of the Danish programme. The authors of the Mozambique study acknowledge that impact assessment is useful in elucidating the “overall experiences with HR&D as an aid instrument” (vol. 8, p.49), yet have little success themselves in conducting such an analysis. They find that the “intended impact” of HR&D assistance is “often...intangible”, and that “available project documentation [is] often not... precise enough when it comes to impact assessment” (vol. 8, p.49). Thus, in contrast to the broader conceptual issues highlighted in the Guatemala case study, the Mozambique team find that impact assessment flounders principally on the more technical aspects of imprecise goal definition and inadequate documentation.

In contrast, the Ghana and Nepal teams both make more constructive attempts at impact evaluation. The Ghana study divides the human rights and democracy support into seven

'key areas' (or sectors) and attempts to assess the specifically Danish contribution to each of these. The Nepal study analyses Danish assistance by institutional beneficiary (i.e. state, local government, CSOs), and attempts to disaggregate the "immediate impact" of assistance to each institution from the "wider impact" (vol. 9, p.3). The findings of the Nepal study differs markedly from that of the Guatemala and Mozambique reports both in its content, with positive impact claimed, and in its tone. For example, regarding support to the legislature, the Nepal team find "identifiable results" such as "better public relations [legislature and citizens] partly through the establishment of an information office", while support to the electoral commission displays wider impact "in the fact that the elections have been carried out in a comparatively free and fair manner" (vol. 9, p.47). The Nepal report displays few of the concerns about causation and linkage found in other volumes. However, whether the study's findings on impact are reliable is a separate question.

The thematic studies also display differing approaches to and perspectives on the question of impact evaluation. Volumes two and three on 'legal and judicial issues' and on 'elections' respectively both concentrate on assessing micro level project impact through logframe-type linkages. The former looks at the outputs resulting from the project inputs and make a modest impact assessment, listing ways in which these outputs have "contributed to" other activities and to general processes of reform (vol. 2, pp.40-43). Annex III of this report continues in this vein, listing planned and actual outputs project by project and also providing project impact assessments. The volume on electoral support is pessimistic about the feasibility of evaluating impact at the country or sectoral level, - "stringent assessments of the overall impact on democratisation by the support rendered can probably never be achieved" (vol. 3, p.36), - yet more optimistic about the possibilities at project level. In the event, the authors are constrained from conducting project-level impact evaluations by the poor quality of project documentation, described as "virtually useless" (vol. 3, p.33). This arises partly from the acknowledgement that "Danida officials have made at best only scant use of the [LFA] method, at least in the HR&D [sector]" (vol.1, p.66). Given the absence of useable material, the authors resort to describing impact in generalised terms and from anecdotal evidence, with such assessments rendered unreliable as a result.

The media case-study comments on the "complex methodological challenge" inherent in impact evaluation, noting that relations between the media and democratisation and human rights are very complex, and that neither have indicators been developed nor impact studies completed (vol. 4, p.1). Consequently, it restricts itself to providing background analysis of the potential role that the media *may* play in the promotion of democracy and human rights (chapter 2), and of global trends in the growth of media and its relation to democratisation (chapter 3). Although directly addressed, the assessment of Danish support to the promotion of democracy and human rights is very limited, though with positive conclusions asserted.

Similarly, volume five on 'Participation and Empowerment' appears somewhat overwhelmed by the 'methodological challenge' of assessing the contribution of relevant projects and programmes to the promotion of democracy and human rights, and declines

to offer such an evaluation. Instead, the study focuses on providing a broad and general discussion of the concepts of participation and empowerment and their integration into international and Danish policy, with a general appraisal of projects in contrasting scenarios in Bolivia and Kenya.

### *Summary*

This nine-volume study represents a considerable achievement, but poses more questions than provides answers. It raises fundamental and challenging issues, notably concerning methodology and impact evaluation, as well as undertaking the most comprehensive review and analysis of a single donor's human rights and democracy programme in this field. Such issues are not resolved, however, and three critical points emerge. First, the methodological dilemmas inherent in evaluating democracy assistance are outlined and clarified, but unfortunately the contribution to their resolution is limited. The study's own methodological approach, 'inter-subjective validation', is not elaborated in sufficient detail. Similarly, and perhaps revealingly, a recommendation of the synthesis volume is to encourage experimentation with "other methods" of evaluation (vol.1, p.68), but with no discussion of what these may entail. Secondly, and relatedly, 'inter-subjective validation' involves a participatory approach in which stakeholders' assessments of reform programmes are central to evaluation conclusions. Yet the degree of participation in the eight case studies remains limited to key stakeholders, that is direct project beneficiaries and implementing staff, rather than a wider range of well-informed parties, inclusive of those without direct interests in the programme. Thus, more critical engagement with the programmes, potentially questioning their relevance and effectiveness, is avoided. Third, questions concerning impact assessment, inclusive of whether and how it is possible to achieve, and at what level, remain unanswered. The eight thematic and country studies address the issue in a variety of ways. While there could be advantages to giving the evaluation teams a 'free rein' in examining this difficult issue, there is little attempt in the synthesis volume to come to some overall consensus or conclusions.

### ***2.3 Swedish International Development Co-operation Agency (Sida)***

The promotion of democracy and human rights has become an increasingly important feature of Swedish development co-operation over the past decade, accounting for around 15 percent of Sida's overall expenditure in 1999 (Sida 2000b). Thus evaluation becomes a most significant issue, though only explored in the last few years. Such efforts can be divided into two. First, a small number of project evaluations have been undertaken, with instructive features highlighted below. Second, a wider ranging study was undertaken recently, focusing unusually on the question of the *evaluability* of democracy and human rights projects (Sida 2000b), rather than an evaluation in itself. This latter study is examined first, then the individual evaluation reports.

### 2.3.1 *The Evaluability of Democracy and Human Rights Projects: a logframe-related assessment* [Sida Studies in Evaluation 00/3]

Published in October 2000, this study was conducted for Sida by ITAD Ltd and the Overseas Development Institute, London. It acknowledges the difficulties of evaluating democracy and human rights support (Sida 2000b, p.vii), and explores the feasibility and the methodology of evaluation in this field: the same *if* and *how* questions that are central to this report. However, it addresses these questions in a very specific and, it is argued here, limited manner.

A starting point of the study appears to be that wider impact evaluation is not feasible, given that “most D/HR interventions are small-scale relative to the complexity of the problems addressed” and that it is “over-optimistic to count on significant progress towards wider objectives from donor financed D/HR activities” (p.6). Thus, “the evaluation of outcomes are generally more feasible than the evaluation of impact” (p.6). While there is some good sense in such statements, and the difficulties identified are real ones, is it valid to conclude that it is not possible to link donor activities to the overall situation of democracy and human rights in a country, and thus to abandon any attempt at wider impact evaluation? Or should ways of resolving such problems be sought? The study for Sida seeks support for its focus on micro level outcomes by asserting that “international literature on the subject also recognises the need for *specific and realistic objectives* to support evaluation initiatives” (p.vii, emphasis added), though without providing references. This emphasis on specific and realistic objectives perhaps leads to self-limitations in terms of both the study’s scope, examining individual projects only not wider programmes, and its approach, using the logical framework as an “organising evaluation structure” (p.viii). In this respect, there is a significant contrast between the Danida study, discussed above, and the one undertaken for Sida. Whereas the Danida report is characterised by the breadth of its scope, undertaking evaluations of programmes not single projects, and its *critique* of conventional (logical framework) approaches, the Swedish evaluability study is narrow, project-based and restrictively logframe-related.

The study selected 28 democracy and human rights projects in Central America and South Africa funded by Sida and examined their evaluability, that is, an assessment of the feasibility of evaluating them using the logframe approach. It was intended as the first of a two-phase evaluation project, with the second phase as a full-scale evaluation of project impact. In assessing the evaluability of each project, the methods adopted were essentially two-fold.

- A logframe matrix was identified or created where one did not exist (called an ‘evaluation pathway’).
- The evaluability of each project was then explored in terms of eleven criteria, of which four were regarded as “particularly crucial” to successful evaluation:
  - the quality of project purpose;
  - the quality of expected outputs;
  - the availability of data;

- the feasibility of attribution;

(p.x)

These tasks were achieved through a three-stage process: an initial review of documentation, followed by preliminary country visits, with the main fieldwork undertaken after an update of methodology. Stakeholder interviews were held separately with different groups in order to identify the logframes on which the projects were based, with the logframe matrix (or ‘development pathway’ where precise or predefined logframes did not exist) discussed with stakeholders in order to elicit the degree of agreement and divergences on objectives etc. Six groups of stakeholders were identified: Sida officials in Stockholm and in embassies; project managers or implementers, both NGOs and in-country government offices; project beneficiaries, both primary beneficiaries (i.e. those who benefit directly from a project) and ‘ultimate beneficiaries’ (i.e. the wider population who benefit from changes achieved by the primary beneficiaries) (p.x). Subsequently, project evaluability was assessed through an analysis of the identified logframes by means of a structured checklist, inclusive of the above criteria (pp.21-2).

The findings were that only two out of the 28 projects met all four main criteria of evaluability. Such findings are perhaps not surprising when the overall context is taken into account. A logframe-related evaluation can provide a rigorous approach to assessing the achievement of the objectives stated in the logframe matrix at the design stage. But it is a pre-requisite that a project has adopted a logframe approach from the outset. Yet this was not the case. Sida had only recently introduced a logframe approach and many projects were not based on a logframe. The attempt at retrospective construction of a matrix by the evaluators for the majority of projects was clearly no substitute for an existent logframe. The absence of pre-existing logframes meant that an assessment of evaluability by a method dependent on them produced rather predictable results.

Even if advocating a logframe approach to project management, inclusive of evaluation, it would seem apparent that a different evaluation methodology is needed for projects that pre-date the introduction of such a system. And, indeed, the ‘terms of reference’ did seem to anticipate such a scenario by requesting that the study “analyse and possibly also propose alternative impact evaluation approaches” (p.24).

One criticism of this study, however, is that little consideration has been given to such alternative approaches. The evaluators state that they re-interpreted this task, arguing that the LFA is not in itself an evaluation methodology, but a framework to structure the design of a project. In their view, this framework becomes the basis for impact evaluation, but carried out using a number of “standard evaluation methodologies”, such as document review, interviews, data collection (!), participatory techniques (ibid., p.25). These are not methodologies, however, but methods. Yet, by this sleight of hand, the evaluators shift the focus from any serious consideration of alternative evaluation approaches to the logframe-related approach, to an emphasis on remedying low evaluability through improved project design. And how is this to be achieved? By better application of the logical framework! Four ‘alternative evaluation approaches’ are outlined very briefly, for possible use in cases of low evaluability by the logframe-related

approach. These are described as ‘responsive’, ‘participatory’, ‘self-evaluation’ and ‘user-orientated’ (pp.68-70). The brevity of examination, for example of a participatory approach, is derisory, amounting to less than serious consideration.

### *Summary*

The expectation that this extensive study would address some of the difficult challenges of ‘evaluability’ in this field is sadly not realised. In contrast, the study suffers from a number of limitations and shortcomings:

- Its focus is narrow and micro, considering the evaluability of the impact of projects only, not wider programmes.
- A logframe-related evaluation provides a rigorous assessment of the projects stated objectives only. It is less successful at examining wider goals, unintended effects or a range of views on the (de-)merits of the project. The measures of success have been largely defined by the project itself.
- Its methodology is restricted to a logframe-related approach, with little consideration of alternative approaches.
- Participation in evaluation is limited. The views of different categories of stakeholders were ascertained, but within a framework delimited by the logframe and the stated objectives. A wide range of views on the appropriateness of interventions was apparently not sought.

In sum, the rigid application of a logframe approach severely restricts the significance and relevance of this study. Conclusions regarding evaluability are predictably negative and pessimistic, given the lack of existent logframe matrices, with little consideration of alternative approaches to addressing the major issues of evaluability in this field. Finally, it is noteworthy that Sida have subsequently abandoned the planned second phase of actual project impact evaluation due to the principal finding of the evaluability study that “only two projects [were] ready for logframe-related evaluation” (p.74). Thus, the disappointing outcome of this study is that the results of Sida’s efforts in support of democracy and human rights remain generally unexplored, and evaluation difficulties largely unaddressed.

### *2.3.2 Project Evaluations*

Fortunately, Sida has also undertaken a number of small-scale evaluations in this area. Three recent examples are examined below, project evaluations in Mongolia and Nicaragua respectively, and a NGO programme evaluation in El Salvador, with instructive features highlighted, both positive and negative. The El Salvador (and to a lesser extent) Nicaragua studies demonstrate the beneficial usage of participatory methods. The Mongolia study displays the limitations of a micro level project evaluation that ignores the overall political context in which support is embedded. Additionally, the challenges of impact evaluation are addressed in differing ways by the El Salvador and Nicaragua studies.

*Swedish Support to Local Self Governance in Mongolia (Sida 2000a)*

This study examined Swedish support to local self-governance in Mongolia, a Sida-funded project that formed part of a larger UNDP capacity building programme. It is an ex-post evaluation that assessed the project's impact with a view to further support. The project consisted of three elements: 1) the training of several hundred members of local government assemblies, including the training of trainers; 2) town twinning arrangements; 3) the strengthening of the local government research association (MALA). The evaluation methods entailed a desk study of project documents and fieldwork involving interviews with project officers and a large proportion (20%) of trainees. The findings were generally encouraging for all three aspects, though with some qualifications. It appears a technically good evaluation, with a particularly positive element being the clear effort made to meet people involved in the various sub-projects and to elicit their views. Nevertheless, although the consultants have fulfilled their brief in evaluating the project *per se*, the interactions between the project and the wider political situation have not been considered, neither the wider impact of the project on processes of democratisation in Mongolia nor the influence of the overall political context on the project itself.

*Access to Justice in Rural Nicaragua: An independent evaluation of the impact of rural court houses (1999b)*

This is a very well structured project evaluation, noteworthy for its methodology. Its stated objective is to assess the extent to which the existence of 122 local court houses (one in each municipality) is improving access to justice for the rural poor; and to identify the factors that inhibit access. Notwithstanding the conclusions of the large-scale *Evaluability* study (Sida 2000b), this evaluation does successfully undertake an impact evaluation primarily using logframe methodology. Assessment is made of the extent to which the project has contributed to the strategic objective of creating “an effective, impartial and independent judicial system accessible to all” in Nicaragua (Sida 1999b, p.3). The study entailed an initial review of documentation, followed by fieldwork in two phases, each of two weeks. The first phase involved interviews with judges, officials and donors, including in-depth interviews with two judges and twenty randomly selected people in order to ascertain how rural people perceived the courts. This first round of interviews enabled the evaluation team to draw up a list of issues for investigation and clarification in a second and more extensive round of interviews. These were conducted in nine municipalities with fourteen judges and 57 clients/users from a mix of criminal, civil, and extra-judicial (mediated) cases (p.18). The finding with regard to the overall justice system was that, “although some progress has been made in the last decade, the judicial system in Nicaragua is still largely inefficient and inaccessible” (ibid. p.3). Nevertheless, it was considered that the strategic objective “has been achieved to a reasonable extent”, and that the project “has contributed to the improved access to justice at the local level” (p.31).

This evaluation is instructive in two distinct ways. First, in contrast to the Mongolia study (Sida 2000a), it combines specific project evaluation with wider analysis, built on a solid contextual foundation. Part 1 provides the background of access to justice in Nicaragua, together with the history of international involvement, inclusive of Swedish support, and the history of the project. Part 2 focuses on the evaluation of the strategic objective, largely based on logframe methodology. Finally, part 3 combines both general contextual issues and specific project findings to provide an analysis of wider factors inhibiting access to justice in Nicaragua. Second, a participatory element is evident, notably through the public opinion survey undertaken, albeit very small-scale, in order to ascertain how local people perceived the rural courts, inclusive of those who were not direct users or beneficiaries. Thus, this study provides an interesting project evaluation where some complementarity between a logframe methodology and a participatory approach is demonstrated.

The authors of this report, as with the Mongolia study, have made obvious efforts to seek the views of beneficiaries. In this report, however, the use of a broader, more inclusive definition of beneficiary facilitates wider conclusions on impact to be drawn.

*Diakonia programme for democracy and human rights: the El Salvador case (1999a)*

Two aspects of this study stand out, its participatory nature and its direct attempt to address issues of impact assessment.

Ambitiously, the objective of this study was a three-fold impact evaluation of Diakonia's (a Swedish NGO) programme in El Salvador, assessed in relation to the aims of Diakonia's regional programme, to those of Swedish co-operation with El Salvador and Central America, and to democratic consolidation in El Salvador (SIDA 1999a, p.i). The study examined the achievements of the six grantee organisations supported by Diakonia in El Salvador, and the report produced is comprehensive, detailed, with many useful findings. Difficulties were experienced, however, in fulfilling the stated objectives, highlighting the challenges of conducting impact assessments in this field, accentuated in this instance by the ambitious, three-fold objectives. The evaluators found that, "It is indeed a very difficult task to assess the adequacy of the results of Diakonia's programme 'in relation to the developments in El Salvador', as it is phrased in the Terms of Reference" (p.21). They outlined three problems. First, how to define 'developments in El Salvador'? Second, how to disaggregate Diakonia's input from that of other donors? And third, how to measure the impact of a "2 year program in a field where deep changes may take a generation or more?" (p.21).

Notwithstanding such difficulties, the evaluators did attempt to provide an impact assessment. The evaluation report is in three parts. First, at the level of the individual organisations, a description is given of each grantee's aims and of what they have achieved (pp.8-13). Second, results are then assessed for their impact on:

- Diakonia's objective (itself disaggregated into 6 goals: strengthening civil society; strengthening local government; strengthening democratic culture; increasing local influence on national decision-making; strengthening local media access; strengthening the role of women and youth) (pp.14-20);



- Sida's goals in Central America (pp.20-1);
- developments in El Salvador (pp.21-4);
- "a consolidated democratic transition" (pp.24-5).

The difficulties, however, in linking the programme to national level democratisation are indicated by the brief coverage of its impact on five aspects of consolidation, no more than one or two sentences on each. Third, the report turns to more conventional evaluation issues of sustainability, cost effectiveness, programme management, and value-added contributions.

The *means* by which the assessment of impact is attempted are of interest, indicating the relative merits (in this instance at least) of qualitative and participatory methods. Initially, a quantitative methodology was intended and considerable time was spent developing an inception report with input from grantees on the types and availability of data. However, when the evaluation team arrived in El Salvador, it quickly became apparent that the data did not exist and the team concluded that "any systematic [quantitative] measure of impact is quite impossible" (p.6). Instead, they adopted qualitative methods, relying on detailed interviews with leaders of grantee organisations, on project and municipality visits, and on interviews with 'national informants' noted for their knowledge of national political developments. Significantly, the latter were especially useful in establishing the socio-political context and 'developments in El Salvador' upon which the impact of the programme could be assessed. This context is carefully used by the authors to highlight four areas of work that they considered to be "obvious weaknesses, or one could rather say missing links, in the representative local democracy in El Salvador which this program has failed to address" (p.23).

It is the participatory nature of this study that stands out, however. Not only did the evaluators give primacy to the views of local actors, including those not directly involved in the programme, but also grantees were themselves involved in developing the proposed methodology and in selecting data for inclusion.

### *Summary*

The three country studies are instructive in different ways, both positive and negative: the Mongolia report for its lack of contextualisation, neither considering the influence on the project of the wider political context nor possible impact on it; the Nicaragua study for combining logframe analysis with a participatory element and for impact assessment at the sectoral level; and the El Salvador study for its high degree of participation and attempt at macro level impact evaluation. The latter report is also a good example of the difficulties in moving direct from micro to macro level in conducting impact evaluation, displaying the phenomena of the 'missing middle', that is, lacking linkages at the meso level. This is discussed further in Chapter 4.

## ***2.4 United States Agency for International Development (USAID)***

Since the early 1990s, not only has USAID's democracy and governance (DG) programme been the largest of all bilateral donors, but the Agency has also been at the forefront of evaluation efforts. Initially, USAID's *Center for Development Information and Evaluation* (CDIE) produced a series of thematic 'lessons learned' evaluation studies. Subsequent to its establishment in May 1994, the *Center for Democracy and Governance* has assumed an increasingly extensive role in providing academic and technical support for the Agency's DG programmes world-wide, including evaluation.

Three distinct phases can be discerned in the monitoring and evaluation of DG programmes. The first phase involved the broad, thematic evaluations produced by CDIE. Second, the Agency-wide adoption of the 'managing for results' (MFR) system in the mid-1990s entailed a very different approach to evaluation, based on a corporate-style performance management scheme of annual reporting of results against targets. However, the MFR system has already undergone some reappraisal, with two changes underway in the DG sector. One is the addition of qualitative indicators to complement the quantitative ones that are used to measure progress towards objectives by US Missions in the annual MFR exercise (USAID 2000a, p.49). The other entails a more fundamental shift from this reporting exercise to more genuine evaluation activity, marking a third phase. The impact of DG programmes on country level democratisation is being investigated through country case-studies "explor[ing] the link between USAID activities and broader democratic change" (USAID 2000e, p.5). Three pilot case studies, Bolivia, Bulgaria and South Africa, are currently being undertaken (USAID 2000k).

These three distinct evaluation phases and approaches are considered in turn, followed by a brief examination of two selected project evaluations that raise interesting issues for further discussion.

### *2.4.1 'Lessons learned' - CDIE series of thematic evaluations*

Five studies were completed by CDIE up to 1998, with assistance to the following sectors evaluated: the rule of law (USAID 1994), civil society (USAID 1995a), legislatures (USAID 1997a), electoral assistance to post-conflict societies (USAID 1997e), and democratic local governance (USAID 1998b). The studies shared a broadly similar methodological approach. Impact evaluations at the thematic level were undertaken through country case studies, five or six in number. Lessons learned from the country studies were then synthesised in a summary report. CDIE developed the initial research design, including selection of country case studies. Case-study research was carried out either 'in-house' or by 'well-known scholars'. These impact evaluations at sectoral level were based on initial documentary analysis, followed by short-term fieldwork, inclusive of interviews with 'informed individuals'. The case study findings were generally presented and discussed in a workshop, prior to production of individual country reports and the overall synthesis report. The latter generally included a wide-ranging analysis of various aspects of the thematic area, and summaries of lessons learned from the case studies. The two most recent studies are examined here as examples of this approach.

The synthesis report on Democratic Local Governance (USAID 1998b) opens with a discussion of the presumed merits of democratic decentralisation: improved participation and accountability; more transparency in government; empowerment and political education of citizens; and improved state responsiveness. Reviewing existing evaluation literature on decentralization, the author finds broad agreement for the positive effects that decentralization initiatives have had on participation among citizens (such as participation in voting and other civic activities), and on the state's responsiveness to citizens. He also finds support for the beliefs that decentralization can support the transfer of some fiscal responsibilities; improve citizens' abilities in dealing with government (education); and promote leadership training among future leaders (pp.7-8). Problems identified in the same literature include: central governments' reluctance to decentralize; the danger that new layers of government will be less than democratic; problems with resource utilization, with local governments preferring to seek short-term political capital out of high profile projects rather than to plan for the long term (pp.8-10).

Evidence from the five country case studies is organised under seven headings ("Donor and host country roles"; "Resistance and political will"; "Representation, empowerment and benefits"; "Fiscal autonomy and regional equity"; "Public accountability"; "Performance and accountability"; and "National advocacy" – pp.16-43), leading to the following conclusions.

- 1) "DLG can significantly increase political participation of marginal groups and can empower geographically concentrated minorities; it appears unable to empower marginal groups that are not geographically concentrated, at least in the short run";
- 2) DLG can help alleviate poverty by strengthening the capacity of local government to deliver services that benefit the whole population, in sectors such as education, health, and water supply; it shows less promise in reducing poverty through efforts directed specifically at marginal groups";
- 3) "DLG can partly sustain itself through local revenue generation, but this will tend to exacerbate regional imbalances unless supplemented by central subsidies to poorer areas";
- 4) "When a variety of mechanisms, such as civil society, media, and political parties are used together, DLG can improve accountability of local government bodies to the citizenry; used in isolation, these instruments appear much less effective".

These conclusions then form the basis for a brief *Lessons Learned* section, which qualifies some of the assumptions underlying donor support for decentralisation. For example, the first lesson in this study ("Representation does not necessarily lead to empowerment, but is valuable in and of itself" – p.43) contrasts with the initial assertion that "the promise of democratic decentralization is to make [the rule of the people]... more direct, immediate, and productive" (p.6). The *Lessons* are therefore not particularly practical in their import, but rather suggest that donors lower their horizons slightly and take a more modest view of what decentralisation can achieve. In this respect, the lessons are analogous to some of the recommendations made to CIDA (Kapoor 1996, IDRC 1999, p.41), which advocated for a more realistic appreciation of what intervention in the DG sector can achieve.

The set of reports entitled 'From Bullets to Ballots' considers electoral assistance to post-conflict societies in five countries (Angola, Cambodia, El Salvador, Ethiopia, Mozambique, and Nicaragua). The synthesis report (USAID 1997e) considers a number of factors thought to be important in laying the ground for successful elections, including timing, design and planning (effectiveness of institutions, choice of electoral system etc), conduct of elections (lack of violence etc), as well as the outcome in terms of the effects the elections had on processes of peace building. Principal findings and judgements are:

- In many instances, elections were scheduled to occur *too soon* after the cessation of violence for agreed political changes to be implemented (e.g. demobilisation of combatants and opening of the political system). This led to their postponement in three of the six case study countries (pp.15-17).
- Difficult technical and political circumstances made conducting an election a complicated affair in each case study country. International assistance helped to create "technically acceptable" elections in five of the six countries, with donors not becoming so deeply involved in the sixth (Ethiopia). However, this process was undoubtedly complicated by "the lack of available elections officials committed to the integrity of the process rather than to a party" in each case study country (p.17). The authors note that, "In general, neutrality was better ensured by the reciprocal controls exercised by representatives of political parties than by the professionalism of the technocrats" (p.17).
- Noting the significant effects that the choice of electoral system can have on the outcome of a contest, the authors observe that in each case a proportional system was used, which also had the effect of encouraging "the participation of even small parties and the formation of new ones everywhere" (p.18). However, the vote generally split between the two main sides in the conflict, leaving smaller parties with marginal levels of support (p.19). More significantly for future assistance, though, the authors note how the main parties in two countries developed *ad hoc* power sharing arrangements as a result of close electoral results. In Nicaragua, the newly-elected President appointed a prominent Sandinista to oversee the restructuring of the military; while in Cambodia, two Prime Ministers were appointed to the coalition government, one from each of the main parties.
- Further issues highlighted concern the demobilisation of ex-combatants, effective voter registration, and countering violence and intimidation in the run-up to balloting. The authors also question the utility of the phrase 'free and fair' when applied to elections by outside observers, noting that it, along with much media reporting and many statements of governments and international organisations, does not always "reflect the complexity of the situation" (p.23).
- The ultimate test of the contribution of democratically held elections to peace building is whether a return to conflict becomes more or less likely in their aftermath. In three of the six case study countries (El Salvador, Nicaragua and Mozambique), the winning and losing parties accepted their role - after negotiation in some instances. In the three other countries (Angola, Cambodia and Ethiopia), the outcomes varied between tense peace and a return to fighting (pp.23-4).

As well as providing a synthesis of case study findings, the authors examine some of the assumptions surrounding electoral assistance, and attempt to create a sense of the limits of what elections can do. Noting that elections take place in a complex web of “existing political structures, social and cultural traditions” (p.33), the authors examine the significance in electoral processes of related factors, both specific and less tangible. The former include the transformation of political movements into parties; the creation of electoral infrastructure; independent media; and public education. While less tangible factors include “the presence or absence of democratic traditions” and “participatory social institutions” (i.e. civil society) (p.33); ethnic cleavages; and economic growth.

As is evident from the above examples, this series of studies have provided detailed analysis and useful insights into specific USAID efforts, inclusive of impact evaluations at sectoral level, as well as raising more general issues of democratic development in the respective sub-fields. Indeed, the initial studies on the rule of law and civil society were among the first serious attempts at evaluation of DG programmes by any donor. The series has discontinued since 1998, however. The reason for their cessation is unclear, but could be related to the rise of the Democracy and Governance Center and associated intra-agency ‘turf-battles’.

#### 2.4.2 ‘Managing for Results’ and Evaluation

Part of the former Clinton administration’s ‘reinventing government’ initiative and the 1994 Government Performance and Results Act, ‘managing for results’ (MFR) aims to provide accountability for expenditure of USAID funds through the precise definition of objectives or performance goals for individual projects and country programmes. Progress towards these is monitored by means of quantitative indicators, with an annual reporting of results by USAID Missions. MFR is very similar to the Canadian government’s ‘Results-based Management’ scheme. Of the two competing objectives of evaluation, ‘accountability’ and ‘lesson learning’ (Cracknell 2000, p.54), it is clear that the accountability objective is predominant in such methods. The nature of evaluation shifts from an intermittent, ex-post review of achievements, shortcomings and lessons learned, to an annual, institutionalised monitoring of short-term results. Such schemes could perhaps be more accurately described as reporting systems than genuine evaluation, which has the primary intent to improve projects and programmes on the ground. Both MFR and RBM are also closely related to the logical framework approach (LFA) to evaluation, given their mutual reliance on the *a priori* establishment of a matrix of immediate objectives and wider goals, along with objectively verifiable indicators (i.e. mainly quantitative) and means of verification.

The adoption of MFR necessitated a wholesale re-organisation of USAID projects and programmes, establishing and institutionalising a bureaucratic methodology that dominates the whole project cycle from design to monitoring and evaluation. Each USAID country programme is required to produce a regular strategic plan, organised as a hierarchical ‘results framework’ of ‘strategic objectives’ (SOs), ‘intermediate results’ (IRs), ‘sub-intermediate results’ (sub-IRs), and ‘performance indicators’. The first three elements constitute different levels of objectives, with performance indicators used to measure whether such objectives have been achieved. [It is stated, rather surprisingly,

that the terms ‘objective’ and ‘result’ are used interchangeably (USAID 1998a, p.5)]. The three levels of objectives are linked in causal hypotheses, i.e. each is perceived as an essential step leading to the next level. A ‘strategic objective’ is the most ambitious result over the 5-8 year timeframe of a country plan, with each country programme likely to entail a small number of SOs in different sectors (for example, one each in democracy, education, and gender relations). In the democracy field, SOs vary considerably in how general or specific they are defined. A USAID Mission’s strategic objective for the DG sector could simply reflect the Agency goal (‘sustainable democracy built’), or Agency objectives (for example, rule of law strengthened). Such high-level objectives can make performance measurement problematic, however, and therefore a lower-level objective, an ‘intermediate result’, can be taken as a more specific SO (for example, effective justice sector institutions). Performance indicators “answer the question of how much (or whether) progress is being made towards a certain objective” (USAID 1998a, p.7), with appropriate indicators requiring the ready availability of data sources. A ‘preference’ for quantitative indicators is noted, valued for their ‘objectivity’, though qualitative indicators are also ‘acceptable’ in order “to capture the qualitative nature of DG programs”(p.8). In fact, suggested indicators in the Handbook (see immediately below) were overwhelmingly quantitative.

To assist its Missions in the MFR exercise, a very substantial *Handbook of Democracy and Governance Program Indicators* (USAID 1998a) was produced. In over 260 pages of tables and text, an extensive range of over 500 possible indicators (‘candidate indicators’) have been developed for the levels of ‘intermediate results’ and ‘sub-intermediate results’, under each Agency DG objective. [Indicators for the four Agency objectives themselves have not been developed. It is stated that the Freedom House Index is most commonly used for measurement of overall country democratisation, although this is under review (USAID 1998a, p.14)]. Under each of the four Agency objectives, a sizeable working group completed the task of, firstly, constructing a ‘results framework’ of pertinent IRs and sub-IRs for each objective, and, secondly, developing multiple indicators by which to measure progress towards each IR and sub-IR. The result framework for each Agency objective constituted a flow chart of four to seven IRs, each of which has between one and six sub-IRs. Candidate indicators for each IR and sub-IR are rarely less than three in number and often as many as six or more, providing an indication of the overall scale of the task. Additionally, data collection methods (and their approximate cost) are suggested for each indicator, as well as a textual discussion of target setting and the interpretation of trendlines (i.e. how much progress can be expected over what period of time). In particular the *Handbook* aims to assist USAID missions with the task of identifying performance indicators through the provision of a wide and comprehensive range of ‘candidate indicators’, with the caveat that they may need to be adapted to local contexts.

Examples of the results framework and indicators are given below in Box 2-2 for each of the four agency objectives, inclusive of source data and other textual commentary for the first agency objective only.

Box 2-2: USAID's Democracy and Governance Program Indicators - Some Examples

Agency Objective 2.1: Strengthened Rule of Law and Respect for Human Rights				
Intermediate Result 2.1.1: Foundations for Protection of Human Rights and Gender Equity Conform to International Commitments				
Indicator	Definition and Unit of Measurement	Relevance of Indicator	Data Collection Methods / Approx. Costs	Target Setting / Trendline Interpretation Issues
1) Number of human rights violations filed against government security forces per 100,000.	Number of human rights abuses filed against security forces divided by total population/100,000. Filed means in a domestic court, a regional human rights tribunal, like the European Court of Human Rights or the Inter-America Court appropriate in the country.*	Direct measure of citizen confidence in the foundations & the willingness & ability of the government to address violations.	National Stats; International Organizations; State Dept. Reports. Cost: Low if there are case tracking or secondary sources. But possibly cost prohibitive if the information has to be collected.	It would depend on the specific problem as articulated by the mission. However, this indicator is most likely to be meaningful when effective means/mechanisms for filing complaints have recently been put in place. In this case, the number of violations filed should increase over the life of the strategy. It was suggested in Guatemala that this might be a good indicator only for a limited period of time. Once a large number of cases are filed, it would be impossible to track them (unless they all begin with the ombudsman).
*Comments: The indicator seeks to measure the violations that are getting beyond the ombudsman's office, thus that office would not be included in the definition of "filed." The only exception might be in a country where the ombudsman's office is not only vested with the authority to investigate but also to adjudicate human rights violations.				

Agency Objective 2.2	More Genuine and Competitive Political Processes
Intermediate Result 2.2.1	Impartial Electoral Framework
Indicator 1)	Degree to which electoral law/rules conform with international standards

Agency Objective 2.3	Increased Development of Politically Active Civil Society
Intermediate Result 2.3.1	A Legal Framework to Protect and Promote Civil Society Ensured
Indicator 2)	a) % of target CSOs that think the registration process is: 1. timely (or has become more timely, etc.); 2. Transparent; & 3. Low-cost. OR b) general CSO sense that process is timely or untimely, transparent or not, low or high cost.

Agency Objective 2.4	More Transparent and Accountable Government Institutions
Intermediate Result 2.4.1	Increasing Government Responsiveness to Citizens at the Local Level
Intermediate Result 2.4.1.3	Mechanisms of Participation Increased
Indicator 2) a)	total # or average # of people attending town meetings organized by local govt.; or b) % of local govts. holding more than X town meetings in the last year with more than Y people attending

Source: USAID 1998a

The sizeable teams dedicated to the production of this Handbook represented a very significant investment of Agency personnel and finance. It would appear to have been an investment of questionable worth, however. No sooner had the Handbook been published in 1998, than Agency guidance turned away from predominantly quantitative tools and towards a more qualitative approach. This turn around was quite remarkable. Initially, the ‘Indicators Handbook’ was lauded as “a first step toward... the complex effort of measuring democracy” (USAID 1999c, p.4). A year later, however, two annual reports discussed the shortcomings of quantitative indicators and signalled a shift to qualitative measures (USAID 2000a and 2000b). The *1999 Agency Performance Report* stated that:

“Measuring democracy is not a science. We acknowledge that attempting to gauge democratization quantitatively *fails* to provide the information we need to measure success... Over the past year, our experience has shown that it has *not been instructive* to compile and analyze quantitative democracy program performance measurement data from the missions on a global or regional basis.” (USAID 2000a, p.48, emphasis added).

Subsequently, activities have been two-fold to develop a qualitative approach. First, as part of the MFR exercise, new qualitative indicators are being developed for use in measuring progress towards strategic objectives and intermediate results. Such qualitative indicators are intended to complement the existing quantitative ones, with publication anticipated for 2002. Second, more significantly for evaluation efforts, the methodology for undertaking qualitative case-studies in programme impact has been developed (USAID 2000k), with pilot country case-studies underway in 2001. This important initiative is discussed more fully in 2.4.3 (below). Prior to this, however, a critique of the MFR system as an evaluation approach is examined, focusing on the inadequacy of its quantitative nature.



*Carothers and 'the inadequacy of numbers'*

Thomas Carothers (1999), a noted commentator on US democracy assistance, has strongly critiqued the MFR system. His criticisms are essentially two-fold. One is an attack on its quantitative basis, 'the inadequacy of numbers', while the other is a critique of the distorting effects of the system itself. The critique is forceful and merits a detailed examination.

While acknowledging that the MFR system has obliged practitioners 'to flesh out' such vague and loose concepts as 'strengthening civil society' or 'making government more responsive' (p.291), Carothers is nonetheless vigorous in his critique:

“The effort to assess the impact of democracy programs by using highly reductionist indicators is a *deeply flawed* undertaking that is consuming vast resources, producing little useful insight or knowledge, and introducing *serious distortions* into the designing and implementing of such aid” (p.291, emphasis added).

He perceives that the “core problem is that democratization in any country cannot be broken down neatly and precisely into a set of *quantitative bits*” (p.291, emphasis added). In other words, there are difficulties both with the division of whole, complex processes into a series of fragments and with trying to quantify them. Using such “informational bits as criteria of success without grounding them in sophisticated, deep-reaching analyses of the political context produces superficial and dangerously misleading pictures” (p.291). Carothers cites the example of Cambodia in 1997 as evidence, where USAID reported that “progress against indicators exceeded expectations” with regard to its democracy programme, seemingly (and ‘almost surreally’) oblivious to the recent coup that had derailed the democratic transition (pp.291-2). As regards quantification, the trouble is that:

“In most cases, numbers tell very little, and what they do tell is unclear. Reducing large elements of democracy ... down to two or three extremely narrow quantitative indicators does irremediable violence to those concepts” (p.293).

It is the “false dream of science..., the belief that all those messy particularities of people and politics can be reduced to charts and statistics” (p.293). Quantitative indicators used to measure the success of legislative strengthening projects are given as an example. Two such indicators are the *number* of public hearings held or the *number* of bills initiated by the legislature as distinct from the executive branch. In Carothers' view, such increases may be for a variety of reasons, some with little to do with democratisation, yet it will be reported as a favourable result. (pp.292-3).

Additionally, not only is the quantitatively-based results system criticised for its ineffectiveness in assessing democracy aid, but also for its distorting effect on the design of new programmes. Projects are designed “that will produce quantifiable results” and “the universe of programme design shrinks to match the indicators” (p.294). In effect, “The evaluation tail begins to wag the program dog” (p.294).

Carothers' criticisms would appear to have had some influence, given the recent shift towards more qualitative methods. While welcome, the introduction of qualitative indicators into the MFR framework would seem of less import, given that it is essentially a managerial tool to ensure accountability for efficient expenditure of public funds. The MFR system is essentially oriented to provide a positive 'spin' on short-term Agency achievements, with little to say about the effects of Agency activities on complex, long-term processes of democratisation. In contrast, the qualitative case-studies on programme impact would seem a more significant initiative.

#### *1.4.3 Country case-studies in programme impact*

This recent initiative by USAID is the most direct effort to date to address the difficult methodological question of *how* to determine the impact of DG programmes on democratisation processes. A research design and protocol has been prepared for pilot case-studies (USAID 2000k). Although the three country studies, Bolivia, Bulgaria and South Africa, remained incomplete at the time of writing, the research methodology is of much interest in itself, and is outlined below. Publication of the pilot case-studies is scheduled for 2002.

As previously discussed, a major methodological challenge for evaluators concerns whether political changes in the recipient country are the result of particular democracy assistance efforts, described as the problem of attribution in this research and dubbed by Carothers as the 'causal conundrum' (1999, p.283). USAID intends to address this challenge through a three-stage research process: a) a political context study; b) an account of USAID programmes; and c) identifying and tracing impact. These steps are examined in turn.

First, a political context study is deemed essential to provide information on democratisation trends in the country concerned over the specified time period (1995-2000 for the pilot case-studies). This is completed as a desk study, focusing on "particular areas of significant change (positive or negative), the principle sources of change, and the principle constraints to it", inclusive of key actors (USAID 2000k, p.1). To facilitate this assessment of political change, Linz and Stepan's framework of five categories or 'diagnostic variables' is adopted: state coherence (a precondition for democracy); political society; civil society; the rule of law; good governance (or a 'usable state', able to deliver goods and services) (Linz and Stepan 1996). Such a political context study provides both (retrospective) baseline data and an assessment of subsequent changes at a broad level of analysis.

Second, USAID's DG programme in the country concerned is examined, inclusive of objectives, strategy and activities over the period of study, with initial identification of impact from available documentation and key interviews (USAID 2000k, p.8).

Third, and most significantly, research attempts to trace programme impact and connect it to higher levels of political change, thereby examining whether and in what ways USAID programmes contributed to democratisation at a country level. It is recognised that "iron-clad causal connections should not be expected" (USAID 2000k, p.1), but 'plausible

connections' are sought between the intervention and the political change. The method is known as 'process tracing' and has a distinctive logical framework feel to it: a logical sequence is traced from USAID activities to outputs (i.e. concrete results such as numbers of election officials trained) to outcomes (e.g. voter registration systems) to higher level impacts (clean elections). One difference from typical logical framework analysis, however, is that investigations are not limited to pre-established objectives within a project matrix. Research focuses on apparent correlations between programme impact and areas of political change, tracing connections until a possible relationship "becomes too tenuous to be persuasive" (p.5). Evidence of impact is sought from a variety of sources, with validation from two or more required, i.e. a process of triangulation. Research also examines instances where immediate programme impact is identified, yet where overall political change in related areas was absent or negative. This raises the possible scenario that DG programmes helped prevent or slow down democratic regression or backsliding (p.4). Finally, evidence of impact is aggregated and considered against the overall political context, with judgements made concerning the contribution and significance of the USAID programme. Questions to be answered include whether the agency was active in those areas most crucial to democratisation and/or those most critical to possible backsliding.

Steps one and two are undertaken by a small team of USAID personnel and external consultants as a desk study, inclusive of drawing up preliminary hypotheses about areas of programme impact and political change. The third step is based on fieldwork, conducted over a two to three week period, involving further document collection and interviews with "mission and embassy staff, partners, political actors, and host country experts" (p.9). It is further stated that "particular attention must be paid to interviewing actors who hold divergent views about both political trends and USAID assistance" (p.11). A single local expert would appear to be contracted for this fieldwork phase only.

USAID is commended for shifting to more qualitative assessments and for directly tackling the methodological challenges associated with impact evaluation at country level. Given successful outcomes of the three pilot cases, it is understood that further country case-studies will be undertaken in subsequent years.

Two shortcomings in the methodology are perceived, however, one relating to unintended consequences and the other to the lack of local participation. These are examined in turn.

It is noted that the search for positive impact includes the examination of circumstances where no progressive political change is observed, entailing the difficult counterfactual of analysing whether external assistance could have prevented (or slowed down) democratic backsliding. Yet an investigation into unintended negative impact is omitted from the research agenda. Schmitter and Brouwer have addressed this issue head-on. They state: "it is not sufficient just to measure the extent to which a specific goal has been reached. The 'complete evaluator' has to deal with the entire array of changes emanating from a specific programme or project" (Schmitter and Brouwer 1999, section IV, 2). External intervention involves a dynamic, inter-active process and can have unintended side effects, with an intensification of ethnic divisions and conflict probably being the greatest fear. Less dramatically, a distorting effect on the nature of civil society is possible as

some NGOs “adapt to the discourse of the donors” in order to “enjoy funding, visibility and a certain influence” (section IV, 2). One potential negative consequence is the intensification of upward linkages and accountability to donors, while downward linkages and accountability to members detrimentally affected.

The issue of participation is a multi-faceted one, with only some aspects examined briefly here. (The alternative methodology proposed in Chapter 5 emphasises a participatory approach, and thus issues of participation are examined in more detail there.) Local participation in the USAID studies appears very limited, restricted to one local consultant and interviews with local political actors, used mainly as a source of information. It seems that little or no domestic contribution is made to the political context study, undertaken in Washington prior to fieldwork. Yet, in making assessments of “the general direction of change, particular areas of significant change (positive or negative), the principal sources of change, and the principal constraints to it” (USAID 2000k, p.5), who is better informed than local political scientists and other local experts? They have invaluable knowledge of rapidly changing regime circumstances and of the alliances between national, pro-democratic actors that are so central to processes of democratic transition and consolidation. Participation at this level appears restricted to one local consultant at the later fieldwork stage only.

There is little or no opportunity for active local participation in the research from the design and planning stages onwards. Instead, the evaluation adopts the traditional approach of research by a small team of (almost all) donor country personnel, flying in and out on a brief fieldwork mission. Such evaluation methods are at odds with the consensus that democratisation is essentially an internal process, with the contribution of external actors being relatively peripheral. Evaluation provides an opportunity for critical reflection on democracy promotion activities, for eliciting the views of knowledgeable, domestic pro-democracy actors on external efforts as well as on democratisation trends, with the aim of re-orientating strategy in the light of such discussion and analysis. Yet, if evaluation studies are themselves donor-focused and conducted by external contractors / consultants, this can become an opportunity lost as the moment for critical thinking descends into mere corroboration and affirmation of donor efforts.

While the stated intent by USAID to include divergent views is exemplary, the degree of inclusion or exclusion of various stakeholders remains an open question. To what extent will indirect stakeholders or non-recipients of DG assistance be included, for instance? Indirect stakeholders are those groups and individuals who are *not* recipients of the assistance, i.e. with no *direct stake* as such, but who have a critical engagement with democratic processes in their country, and thus have an interest in and are affected by the political interventions of external agencies. Categories of ‘non-recipients’ include those denied DG assistance, or excluded from applying by domestic legislation and other measures associated with a (semi-)authoritarian state, or self-excluded by their opposition to democracy assistance as external manipulation (Brouwer 2000, pp.10-11).

Issues of impact evaluation and of a participatory approach are further addressed in Chapters 4 and 5 respectively.

#### 1.4.4 Project evaluations

USAID has also conducted a number of specific evaluations of selected projects and/or programmes. Two are examined here, one from 1995 and the other from 2000, chosen partly because of the interesting commentary on participatory evaluation methods and on impact evaluation that they give rise to.

##### *Evaluation of the Impact of the Centro DEMOS Programme, El Salvador (1995)*

Entitled *Evaluate the Past to Build the Future*, an evaluation was carried out in 1995 on the Centro DEMOS project in El Salvador, a training programme aiming to promote tolerance among political elites in the post-Peace Accord period. The objective was to evaluate impact “on the level of tolerance of direct beneficiaries, associated institutions, and the reconciliation process in El Salvador” (USAID 1995b, p.3). Of particular interest is its participatory nature. A Washington-based contractor, Management Systems International (MSI) was responsible for the evaluation. It is stated that the methodology was drawn up over 3 weeks “with the participation of all interested parties”, with objectives, indicators, means of verification, methods and executors all agreed upon (p.4). The ‘interested parties’ were:

- 2 representatives of MSI (including the project leader);
- 3 representatives of USAID/ODI;
- 2 representatives of ICAS (the Institute for Central American Studies, a US NGO);
- 1 representative of FUNDEMOS (the governing body of Centro DEMOS);
- 3 representatives of Centro DEMOS;
- 2 representatives of the EX-CEDES Association (an organisation of former participants in the Centro DEMOS programme).

(p.4)

Methods included a review of documents (including the political context), a questionnaire to 120 former students, and interviews with 34 political leaders, Centro DEMOS staff and 55 former students. The in-country evaluation appears to have taken over four months (September 4 – December 10), although the number of hours and weeks actually spent on the evaluation is not stated. One in-country partner makes the criticism in the report that working practices could have been more effective if “everyone on the evaluation team - including the consultants - lived permanently in the country” (p.44), indicating that this was not a full-time exercise for all involved.

Findings regarding the evaluation process are of interest. Despite the comparatively high level of participation, this remains qualified. It was felt that the degree of participation was adversely affected by the contract requirements which “did not permit the flexibility needed to develop a complete, truly participatory process” (p.45). In practical terms, this meant that USAID required a complete draft report before evaluators left the country. This deadline (a standard contractual condition) hampered the evaluation team’s efforts

to arrive at findings by consensus by emphasising the final product over the process or quality of reporting. The authors noted that the deadline prevented the working through of differences of opinion in the group. The report further recommended the adoption of “internal processes of continuous monitoring which involve the active participation of all parties” (p.46), in addition to the occasional, external-led evaluation. Such monitoring was required in the original programme specification, intended to ensure ongoing relevance and appropriateness, but not implemented.

### *Impact Assessment of OTI programme in Indonesia (2000)*

An ‘impact assessment’ of a USAID programme in Indonesia during FY 1999 was carried out by PricewaterhouseCoopers. The programme was implemented by the Office of Transition Initiatives (OTI), based in USAID’s Bureau for Humanitarian Response. OTI provides short-term (two years) project assistance only, but claims to be able to provide crucial support by virtue of its ability to implement projects very rapidly. Projects are intended to serve as potential bases for future work and may be referred on to other funding streams after the initial two years with OTI.

OTI’s programme in Indonesia comprised five thematic areas of work, (civil society groups, elections, media, governance, and civil/military relations), aiming to achieve six broad objectives. These ranged from the more modest aim of promoting dialogue among opposing groups in conflict areas to the more ambitious aims of changing the role of the military and providing new Regional Parliaments with the capacity to manage their regions’ resources (USAID 2000c, pp.2-3). Only the first three thematic work areas were evaluated in the report, identified as “having the greatest overall impact on peace and democracy in Indonesia” in the year in question (FY1999)” (p.4). In fact they contained by far and away the largest incidence of projects in the overall programme (259 out of a total of 285).

Evaluation objectives were ambitious, attempting the difficult task of a macro-level impact evaluation. This aimed to assess “the impact of OTI transition programs in Indonesia during FY1999” (p.3), and to determine “whether and how OTI contributions have assisted the transition to peace and democracy” (p.4). Methods were described as “quantitative, qualitative, and anecdotal” (p.1), although the nature of ‘anecdotal’ methods is not made clear. In addition to a documentation review, the report is based largely on fieldwork comprising interviews with grantees and focus group discussions with project beneficiaries. The incorporation of a participatory dimension would seem constructive, though the author raised concerns regarding ‘who participates’ and ‘whose voices are heard’. She noted that: “OTI staff... identified the participants of the focus group discussions and arranged all logistics for the meetings. We do not know what kind of bias may have been present in the selection” (p.4).

Findings were of “positive and meaningful impacts” in each of the three thematic areas (p.5). For instance, reviewing a training project for journalists on the electoral process, significant outcomes were identified as the acquisition of a better understanding of electoral rules and of electoral fraud. The report also notes a number of problems with evaluating impact, notably that of the civil society programme, “not because individual

grants were not successful but because there were few commonalities among the civil society activities included in this assessment” (p.14). This implies that the overall programme was somewhat *ad hoc*, lacking a coherent strategy, and that its aims of increasing CSO involvement in voter education and of promoting peaceful political participation, were either too broad or not met.

The report also seeks to analysis the impact of the overall country programme (p.4), but is unable to do so because of the lack of available data in many instances (p.5). Nevertheless, the author does make positive impact assessments, based on heavily cited interview evidence, for example, “OTI was like a midwife helping our voter education efforts be born” (p.18), with illustrative statistics on project output, for example, “Radio broadcast on women’s participation in political transition: 13 million listeners as beneficiaries” (p.18).

Thus, it is notable that impact evaluation is more possible in terms of tracing linkages from project outcomes to the general sectoral level. In contrast, attempts to evaluate impact on the programme goal of overall democratisation became more tenuous and contrived.

The report also raises an interesting question about the timing of evaluations. The author finds that the time elapsed between the implementation of some projects and her study visit was 16 to 18 months in some cases. As a result, “Grantees and OTI/Indonesia staff members found it difficult to remember details about activities... some grantee organizations no longer existed, the primary grantee contact had moved to another organization, or the OTI/Indonesia staff member no longer worked in Indonesia” (p.22). Therefore the author recommends that evaluations are conducted within two months of fiscal year end (p.22). However, such a short time scale severely restricts the type of impact evaluation that can be conducted, accentuated in the case of democracy assistance where impacts are more long-term in nature.

### *Summary*

The above two studies raise pertinent evaluation issues. Questions concerning the nature and degree of participation emerge from both. The El Salvador report shows a relatively high degree of participation, but with two qualifications. The ‘interested parties’ are limited to project beneficiaries and do not encompass a wider circle, while the participatory process was restricted by time and resource constraints. The Indonesia report indicates how participatory techniques, such as focus groups, remain subject to manipulation by donor agencies, with or without malevolent intent. Additionally, this latter study raises questions concerning ex-post impact evaluations, notably their timing and the better grounding of impact analysis at the sectoral level. The issue of tracing linkages in impact evaluation and the importance of sectoral or meso level is discussed in subsequent sections.

## ***2.5 European Commission***

The European Commission’s own development aid programme is characterised by institutional complexity and fragmentation, composed of a number of ‘regional’ programmes each with a distinct legal, financial and administrative framework (see

Crawford 2000b, pp.93-8). At policy level, the promotion of human rights and democracy has been prioritised since the landmark Council of Ministers Resolution of November 1991, confirmed most recently by the European Commission Communication of May 2001 on 'The European Union's Role in Promoting Human Rights and Democratisation in Third Countries'.

### *2.5.1 Evaluation studies*

To date, four evaluation studies have been conducted on European Commission assistance to democracy and human rights, each evaluation restricted to a particular budget line(s) or regional programme. An initial report on positive measures to support human rights and democracy in developing countries was undertaken by the German Development Institute and published in 1995 (Heinz et al. 1995), followed by three studies concentrating on different regions. An evaluation of the PHARE and TACIS Democracy Programmes (PTDP) (1992-97) in Central and Eastern Europe was undertaken by ISA Consult and the European Institute, University of Sussex (European Commission 1997). An evaluation of the MEDA Democracy Programme (MDP) from 1996 to 1998 was undertaken for the European Commission by Nadim Karkutli and Dirk Bützler (European Commission 1999). Most recently, an evaluation of 'positive actions in the field of human rights and democracy' (1995-99) in the African, Caribbean and Pacific (ACP) nations was undertaken by Franklin Advisory Services (European Commission 2000a).

The latter three regional evaluations are examined here in an integrated manner, undertaking a comparative analysis of evaluation objectives, methodological approaches, and findings.

### *Objectives*

The three evaluation studies are characterised by difference. There is no consistency or uniformity in the evaluation approach, objectives or methodologies, in line with the general fragmentation and incoherence of Commission activities in this field (Crawford 2000).

The PTDP study is the most ambitious of the three. Its objective is an ex-post impact evaluation, investigating the contribution of the PTDP to democratisation processes in nine selected countries (out of 26), i.e. an impact evaluation at the macro, national level.

Objectives of the MDP study are two-fold: an analysis of the general intervention strategy at country level, and an evaluation of 34 (out of 124) selected projects in six (out of 12) countries. It is described as "an objective ex-post examination" of the strategy used and of the impact of projects (European Commission 1999, p.14). Thus, impact evaluation is at the micro level of selected projects only.

The ACP evaluation is different again. Its stated objectives are to evaluate the adequacy and coherence of the EC's policy 'instruments', and the relevance, effectiveness, efficiency, impact and sustainability of its 'positive actions' (European Commission 2000, p.18). Like the MDP study, impact is investigated at the micro level only,



examining projects in seven selected country cases (out of the 77 ACP countries), namely Senegal, Rwanda, Nigeria, Malawi, Haiti, Ethiopia and the Great Lakes region (Congo-Tanzania). Additionally, this study entailed a thematic focus on conflict prevention and electoral assistance.

One interesting element of commonality, however, is that all three evaluation studies themselves added a further objective: an assessment of programme management by the Commission.

### *Methodologies*

There are major differences in methodological approach between the PTDP study and the MDP and ACP studies. The latter two essentially rely on the standard logical framework approach to aid evaluation, while the PTDP study tailors its methodology to the specific task of evaluating democracy assistance.

#### *PTDP*

The PTDP study adopts a two-stage approach to its objective of macro-level impact evaluation. The first stage involves a democracy assessment of the strengths and weaknesses of the nine countries, using a methodology developed previously by Kaldor and Vejvoda (cited in European Commission 1997, p.13). The second stage entails an examination of projects and programmes to assess their contribution to democratisation in each country, that is, investigating to what extent they have contributed to the strengths and addressed the weaknesses. Eight evaluation criteria were invoked, fairly traditional in character: relevance, consistency, adequacy of procedures, cost-effectiveness, impact (both intended and unintended), sustainability, replicability, visibility (European Commission 1997, p.12). A combination of quantitative and qualitative methods was used, with statistical analysis complemented by qualitative background studies (stage one above) and fieldwork interviews and roundtable discussions.

In more general terms, the PTDP study makes useful contributions to discussions of methodological issues in this field. First, democratisation is acknowledged as a complex process influenced by a multitude of interrelated factors, including democratic institutions and processes, economic development and political culture. Second, it is recognised that the impact of a single programme (PTDP, for instance), far less that of a single project, can never be measured exactly, given that it is very difficult to single out its influence from the range of factors affecting the democratisation process. Four main problems concerning 'attribution' are identified.

- 'With and without' situations or issues of counterfactualty: rarely is it possible to gauge what would have happened without the external support, given complex and turbulent political contexts.
- Intangible lines of causality, with indirect performance indicators (for example, increased number of legal drafts) which could have happened anyway.

- Outputs of a specific donor or programme (for example, PTDP) can be indistinguishable from those of other donors.
- Possible outcomes of donor efforts may be contradicted or cancelled out if not supported by other local policies and processes.

(European Commission 1997, p.15).

Unfortunately, despite recognition of such methodological problems, only a limited attempt was made to address and potentially resolve such issues, perhaps due to the exacting time constraints under which the evaluation was undertaken. The study's findings of a significant impact on strengthening the NGO sector (p.5) (discussed below) fall back on a relatively unsophisticated and self-evident linkage between projects and their contribution to national-level democratisation, given the overwhelming focus of the PTDP on the non-government sector.

### *MDP*

The methodology of the MDP study was divided into two aspects, corresponding to the two distinct objectives. First, background studies of the six country cases were completed and priority issues of democratisation ascertained for each country, pertinent to the evaluation of the intervention strategy. Second, as required by the terms of reference, project evaluation was based on the logical framework approach (LFA), with each project's logframe used to assess the relevance, efficiency, effectiveness, impact and sustainability at project level. Fieldwork methods included questionnaires, interviews and roundtable discussions. Interestingly, reservations were expressed as to the appropriateness of the logical framework approach. It was emphasised that such political programmes have "many 'soft' components with regard to project results, objectives and impact... very difficult to measure in terms of the logical framework" (European Commission 1999, p.15). Notwithstanding the terms of reference, the study intriguingly cites support from the Commission itself, quoting the statement that "human rights projects are quite different from infrastructure construction programmes, and the evaluation and selection criteria should therefore be based on a different approach" (European Commission 1995, cited in *ibid.*). It would appear, however, that little subsequent progress has been made in developing such an alternative approach, with continuing reliance on conventional evaluation methodology.

### *ACP*

In contrast, the ACP study enthusiastically adopted the logical framework "as the cornerstone of its methodology", describing it as "the most widely accepted planning tool" and a "useful quality assessment tool" (European Commission 2000a, p.18). Although the authors clearly state that "the evaluation formulated its analysis... in logical framework categories" (p.65), the precise way in which this methodology was employed is neither well explained nor appears unproblematic, with logframes having to be (re-)constructed at both project and programme levels. At project level, it is stated that the logframe "often has to be reconstructed *a posteriori* from the project documentation" (p.18), presumably due to absence or inadequacy. A programme logframe was also constructed in order to evaluate the positive measures as a single project, thus enabling

general conclusions to be made without relying on an aggregation of “variable individual project results” (p.65). Somewhat bewilderingly, two programme logframes were then drawn up. One is described as the ‘*de facto*’ or ‘restored’ framework (pp.66-71), based on existing Regulations and other Commission documents. While the other is labelled as the ‘optimal’ or ‘recommended’ framework (pp.73-78), “as defined by the evaluation findings” (p.65). What is clear is the authors’ commitment to a logframe-based evaluation study. What remains unclear is not only how this methodology was operationalised, but also its appropriateness in this context, which remains unquestioned. Additionally, the role played by the country case studies in the report and what research methods were used are not clearly elucidated.

One aspect common to all three studies is the relative lack of participatory methods. Participation by stakeholders and beneficiaries of assistance is not completely absent, but limited and instrumental, used more as a means of gathering information, than valued as an essential and dynamic feature of evaluation in itself. The PTDP and MDP studies both use roundtables in the country studies as a method by which to ascertain opinions, while the ACP report uses stakeholder interviews for general validation purposes.

### *Findings*

The findings of the PTDP and the MDP studies are generally positive, while the ACP evaluation is more mixed in its overall findings. The most critical comments in each report are reserved for management and procedural issues.

#### *1. PTDP*

The overall finding is that “the PTDP has been of considerable value for the development of democracy and civil society in Central and Eastern Europe” (European Commission 1997, p.7). The most significant impact is judged as its contribution to the growth of a lively NGO sector in all nine countries examined, regarded in turn as crucial to the democratisation process (p.76). It is noted that it is difficult to distinguish the impact of the PTDP from other Western democracy assistance, and that democracy assistance itself is a “comparatively minor instrument in assisting the process of democratisation” (p.76). Nevertheless, it is held that such assistance has been important financially, psychologically and politically to non-governmental groups. It is believed that foreign funds have helped create “a moral community” essential to the construction of a democratic political culture (p.77).

Findings on management of the programme were more mixed, noting delays in financial disbursements and the potential adverse effect on impact. Programme management up to the selection of projects is deemed relatively satisfactory, contracted out to the European Human Rights Foundation, but subsequent procedures for issuing contracts and making payments, in the hands of the Commission, are severely criticised for serious delays, undermining projects and damaging the reputation of the programme (p.80).

One interesting recommendation pertains to the formalising of a participatory dimension to programme planning. It is noted that NGOs in recipient countries have helped shape priorities informally, sometimes through their Western partners, alerting the EU to new

issues. It is recommended that discussion of priorities for democratisation in each country should be formalised through regular round tables, like those organised for the study, involving recipients, EU delegations and outside experts (p.78). If implemented, two outstanding issues would concern the relative weight of internal and external personnel and the question of who participates from the recipient side.

## 2. MDP

The finding concerning the intervention strategy was basically that there wasn't one: "there has been no official formulation of a particular strategy" (European Commission 1999, p.32). In other words, project selection had been reactive and application-led, thus relatively *ad hoc*. Yet, rather than being critical of the lack of country-based strategies, the evaluators attempt to turn this into a positive. Such a 'reactive strategy' (as generously described) is "not to be seen as being negative" as "it guarantees a bottom-up approach" and "usually the human rights and democracy organisations on the ground have the best judgement on priorities for their country" (p.32). Whilst acknowledging that this latter statement appears to correspond to the approach favoured in this paper, which itself emphasises the essential contribution of local actors to developing a democratisation strategy, there are potential dangers in such a positive perception of a reactive, application-driven approach. Responding to applications remains a relatively arbitrary process and does not 'guarantee' the inclusion of the *range* of local actors, both government and non-governmental, who may perform key functions in the democratisation process.

The strengthening of civil society is again highlighted as the main positive contribution of the programme, unsurprising given that 96 percent of projects were to NGOs. Such concentration of assistance is not only noted satisfactorily, but also the "continuation and expansion of the support to local civil society" is recommended, perceived as playing "a crucial role" in the "long and hard road to democratisation in the Arab World" (p.2). Whilst the significance of civil society to democratisation is not questioned here, an almost exclusive focus on channelling support to NGOs contains an inherent danger. It disregards the importance of democratising government institutions, of supporting pro-reform elements *within* government, and of strengthening government-related organisations that perform 'watchdog' and accountability functions, for instance, parliament, a human rights ombudsman, the auditor general.

Specific findings with regard to the 34 selected projects were generally positive for each of the five criteria, though with some qualifications. Impact at micro level was perceived as mainly good, though variable between different types of projects. One exception to the positive picture, however, was the negative findings with regard to the adequacy of procedures and management. Substantial delays were common due to understaffing and overextensive procedures, viewed as not only having adverse effects on the efficiency, effectiveness and impact of projects but also on the credibility of the EU as a donor.

## 3. ACP

The findings of this study are more mixed, though again it is issues of Commission management that receive the harshest criticism, viewed as having had a particularly adverse effect on achievements. Such difficulties are not presented separately but

integrated into the overall conclusions on the five stated criteria, as follows. Findings are closely related to the logical framework approach that both permeates and limits this study.

The *relevance* of measures is generally very good, though with qualifications that a longer-term strategy is needed to replace short-term preoccupations and that some areas are relatively neglected, for example, governance, conflict, economic, social and cultural rights (European Commission 2000a, p.6).

As regards project *impact*, the country studies generally note positive contributions, but it is recognised that impact is difficult to ascertain due to the attributional problems associated with single projects as triggering a particular change, as well as the lack of specific and realistic objectives by which to assess impact (p.7).

*Coherence*, understood as the quality of the linkages between results achieved and objectives, is perceived as particularly problematic, due to:

- Weaknesses in the logical framework, focusing again on poorly defined specific objectives (or project purpose), often insufficiently related to the particular country context or a particular aspect of the EU's programme, resulting in a "problematic gap between the stated immediate objectives of the projects, and the general objectives" (p.8).
- Weaknesses in programme management, with the Commission's ability to set abstract goals not matched by its ability to put these goals in practice. The authors find that poor management affects all stages of the programme cycle, from project selection and planning, to monitoring and evaluation. It is noteworthy that, unlike the PTDP, the Commission dealt with all aspects of programme management with ACP countries, with no contracting out of parts of the programme cycle. Importantly, the study also found evidence of good project management on the ground, underlining that problems lay principally in Brussels. This leads to recommendations for greater involvement of EU Delegations in the identification of projects and in programming generally. Whether this would lead to greater participation by *domestic* actors in the development of country strategies remains an open question, apparently not one addressed in the evaluation report.

*Efficiency and effectiveness* are found to be difficult to assess due to the lack of specific programme/project objectives, as referred to above. Again, the harshest critique is reserved for poor project management at the Commission level. While project management is relatively good at country level, the process of internal reform in Commission structures has led to a shift of emphasis in programming away from Delegations to specialised units in Brussels (p.9), contrary to the report's recommendations. The report concludes that this trend has led in fact to a *decrease* in the effectiveness and efficiency of individual projects since late 1998 (p.9).

In sum, compared with the other two regional studies, the authors of the ACP study identify greater problems with EC interventions in the field of democracy and human rights. Predominantly, however, these pertain to design (logframe approach) and management issues (Commission structures and division of labour), rather than addressing more major questions such as the development of coherent democratisation

strategies at country level and the respective roles of internal and external actors in this task.

### *Shortcomings and limitations*

While these evaluation studies provide important and useful insights, they also display a number of shortcomings in the areas of methodology, impact evaluation, political context, and lack of critical analysis.

In the realm of methodology, neither the MDP report nor the ACP study engages with the inherent difficulties in this field, relying on the logical framework approach, though with some reluctance on the part of the former. In contrast, the PTDP study engages most usefully with methodological problems, especially those associated with impact evaluation. Nevertheless, only a limited attempt is made to potentially resolve such issues, perhaps due to the exacting time constraints under which the evaluation was undertaken. Hence, impact evaluation was not successfully undertaken. The study's findings of a significant impact on strengthening the NGO sector (European Commission 1997, p.5) rely on a relatively unsophisticated linkage between projects and their contribution to national-level democratisation, given the overwhelming focus of the PTDP on the non-government sector. Indeed, Schmitter and Brouwer (1999, section IV, 3) specifically cite the PTDP evaluation as an example of the shortcomings of macro level impact evaluations which "tend to assume rather than to prove the impact of individual projects on the entire process of democratisation".

The importance of background studies for country evaluations, analysing the political context with regard to changes towards (or away from) democratisation, is recognised to some extent by the PTDP and MDP evaluations, with the former placing most emphasis on this as an integral part of its methodology. However, greater use could be made of the background studies in both instances. In the PTDP study, its function is limited to evaluating one criterion out of the eight listed (i.e. the *relevance* of projects and programmes), whereas the background studies are only used for assessment of the intervention strategy, not for impact evaluation, in the MDP study.

Although criticisms are made of the management of democracy and human rights assistance, ones that are common to all Commission development programmes, the findings of all three studies are generally positive in relation to the substantive issues of democracy promotion, though less so in the ACP study. This indicates a lack of critical analysis, tending to endorse Commission activities, and limiting their value as a rigorous assessment of European Community performance in this field. This was most evident in the MEDA evaluators attempt to turn their negative findings (the absence of country strategies and ad hoc, reactive programming) into a positive as 'a bottom-up approach'.

These issues are taken up and discussed further in Chapter 4.

## 2.6 OECD Development Assistance Committee

Although not a funding agency, the Development Assistance Committee (DAC) of the Organisation for Economic Co-operation and Development (OECD) plays an important role in aid evaluation issues, notably through its Working Group on Aid Evaluation (previously named 'Expert Group on Aid Evaluation'). In the mid- to late-1990s, the DAC made a significant contribution to policy and practice in the area of 'Participatory Development and Good Governance' (PDGG) through two co-current institutional processes. The 'Expert Group on Aid Evaluation' formed a 'PDGG Steering Committee' in March 1993 that undertook a three-year study, publishing a *Synthesis Report* of its findings (OECD 1997a). Simultaneously, an 'Ad Hoc Working Group on PDGG' was established and organised a series of thematic workshops, resulting in a *Final Report* (OECD 1997b). These are examined in turn.

### 2.6.1 Evaluation of Programmes Promoting Participatory Development and Good Governance: *Synthesis Report* (OECD 1997a)

The work of the Expert Group on Aid Evaluation focused on five themes within the PDGG area, with responsibility for each taken as follows: legal systems (US); public sector management (UK); decentralisation (Norway); human rights (The Netherlands); and participation (Sweden). Public sector management was itself sub-divided into three aspects, privatisation, institutional strengthening and evaluation capacity building. A paper was written on each theme, or three in the case of public sector management. The *Synthesis Report* (OECD 1997a) is composed of the executive summary of each thematic paper, preceded by an introductory synthesis chapter, itself providing a further summary of each theme as well as identifying cross-cutting issues. The individual thematic papers often re-emerge as separate publications, for example the *Evaluation of Decentralisation and Development*, published by the Norwegian Foreign Ministry (1997a). Despite its title, the *Synthesis Report* is less an evaluation and more a review of donor activities and experiences in the selected areas, with lessons learned drawn from each. A strength of the Report is that it gathers together the collective knowledge and experience gained by a range of donor agencies. A weakness is that it constitutes a review of existing documents, inclusive of some evaluation studies, but with no fresh fieldwork or country studies undertaken.

There are no common methods between thematic papers, far less methodology. However, some of the chapters do specifically discuss the issue of evaluation methodology in this field. For instance, the human rights paper identifies three possible approaches to evaluation: project effectiveness; consultative (i.e. participatory) evaluations; impact evaluation of programmes, against the background of different political systems. It is noted that impact evaluation requires an appropriate methodology that remains to be developed, inclusive of baseline studies and concrete performance or progress indicators (OECD 1997a, p.84). Interestingly, the participation paper distinguishes between 'evaluation of participation' and 'participation in evaluation', that is a participatory

evaluation. The conclusion drawn is that, “Despite participatory rhetoric, there is very little evidence of genuine participation in evaluation” (p.23).

The overall objective of this study was to learn lessons from donor experiences and these were outlined in the introductory synthesis chapter. The three initial lessons are outlined below, concerning issues that pertain to the ‘process’ of external support for democratisation.

- *Reform efforts require political commitment and local constituencies supportive of change*, i.e. political will as a requisite for successful PDGG reforms. Where this does not exist then donor assistance should be limited to helping create that political will.
  - *Donor advocacy of ‘home-grown’ initiatives is likely to be more successful than donor-driven reform efforts*, i.e. reform needs to be an indigenous effort, not externally imposed. Host country commitment and ownership is again emphasised as key to success, with participatory strategies needed to build indigenous support for reforms, especially where negative impacts are likely (e.g. civil service reforms, privatisation).
  - *Participation and participatory approaches should be important ingredients in all donor PDGG assistance efforts.*
- (p.24).

Such lessons appear positive in their emphasis on local (domestic) commitment and control, at both governmental and civil society levels, and on participation. Yet the instrumental intent behind a participatory strategy and the apparent ‘hands-off’ approach is rendered visible by the aim of fostering ‘host country commitment’ to donor demanded reforms that may encounter opposition. Rather than domestically generated reform and participation being valued in themselves, they form a donor *strategy*, in the sense of a purposeful design, by which external-driven efforts are perceived as best achieved.

One lesson pertains to the evaluation process itself, with emphasis on the use of “quantitative measures and complementary qualitative measures, and... participatory evaluation techniques”. The advocacy of participatory evaluation is worth quoting in full:

“Truly participatory evaluation techniques have many advantages that donors should consider, such as promoting learning and encouraging client/stakeholder ownership of and use of the evaluation results. While donors’ rhetoric favors using more participatory evaluation, actual practice is still limited if not rare. Donors often use rapid appraisal [participatory] techniques that involve ‘listening’ to customers and stakeholders, but they typically stop short of bringing them fully into the evaluation process. Donor concerns for independence, objectivity, accountability, and ‘control’ over evaluations may put some practical limits on use of participatory evaluation techniques in some cases.” (p.26).

There are two clear, if somewhat paradoxical, messages here. One is to stress the advantages of a *genuinely* participatory approach to evaluation. The other is to recognise the gap between donor rhetoric on participation and their reluctance to loose control over the evaluation process.



### 2.6.2 *Final Report of the Ad Hoc Working Group on Participatory Development and Good Governance, Parts I & II (OECD 1997b)*

The *Ad Hoc Working Group on PDGG* existed co-currently from 1993-96, holding a series of thematic workshops, at which a small number of invited NGOs and academics also participated. The workshops addressed somewhat broader issues, asked more academic questions, and entailed more roundtable discussion, though the selected themes generally coincided with those of the *Expert Group on Aid Evaluation*, with the exception of the omission of public sector management and the addition of civil society. Lead responsibility was taken as follows, with discussion papers produced prior to each workshop: the role of donors in the democratisation process (DAC); civil society and democratisation (Sweden); human rights in development co-operation (Canada); legal systems (US); and decentralised co-operation (US).

The *Final Report* (OECD 1997b) is divided into two parts. Part I is a policy statement endorsed by the DAC High Level Meeting, whilst Part II provides summaries of the discussions at the five thematic workshops.

The central feature of Part I is eight key conclusions, synthesised from the overall discussions and activities of the Group. These are at a very generalised level and, rather than based on lessons from experience, merely provide a concise affirmation of current donor beliefs and rhetoric in this field. The first asserts the ‘new orthodoxy’ (Leftwich 1996, p.4) that ‘*Democratisation and Good Governance are Central to the Achievement of Development Goals for the 21<sup>st</sup> Century*’ (OECD 1997b, p.3), i.e. that democratic governance is essential to progress in the developmental goals of poverty reduction, gender equality, basic education and health standards and so forth. Other key conclusions avow the common donor declarations on ‘partnership’ [*Base Partnerships on a Common Understanding of Development – Governance Linkages*] and ‘local ownership’ [*Take Local Ownership Seriously*], and affirm the significance of non-governmental actors as advocates of reform in the political and economic systems [*Strengthen the Strategic Role of Civil Society*] (ibid.). The final key conclusion, perhaps of more practical consequence, declares the importance of both a strategic approach and a long-term perspective to governance assistance [*Operate in a Long-term Strategic Framework*], with particular reference to fragile democracies in post-conflict situations (p.4). Issues of evaluation and evaluation methodology are not directly addressed, but there would appear to be two implications from the final point. One re-affirms the difficulty of undertaking a meaningful impact evaluation either during or immediately after a programme, when affects may be more long-term (Sida 1999a). The other indicates the influence of the overall political context of a country on programme success or failure, especially how conducive that context is to ongoing democratisation and the relative strength or weakness of pro-democratic actors.

Additionally, Part I of the report outlines ‘Action-oriented Outcomes’, including the formation of an Informal DAC Network on PD/GG as one of many ‘products’. Other recommendations include:

- “Pilot cases for learning how to develop locally-owned in-country processes for dialogue and co-ordination on PD/GG issues”;
- the permanent incorporation of PD/GG into the evaluation agenda of the DAC Expert Group on Aid Evaluation; and
- a new thematic emphasis in PD/GG work on women in development, corruption and the media;

(OECD 1997b, p.5).

It is difficult to ascertain the extent to which these recommendations have been implemented in practice. However, rather than the stimulation of such efforts by the two reports, it appears that governance activities went into a lull after 1997. Certainly, the Informal Network on PD/GG was relatively inactive from its inception in 1997. In May 2001, along with the Institutional and Capacity Development Network (I/CD Network), it was subsumed into a new DAC Governance Network (GovNet). At the time of writing, information on GovNet’s proposed activities was not available.

Part II of the report provides interesting synopses of the workshop discussions, but with little specific consideration of evaluation issues apart from a few brief mentions of agencies’ evaluation findings.

Both reports represent valuable contributions to discussions on the role of development aid in promoting democracy and good governance, with lessons learned and key conclusions particularly useful. Questions and problems of evaluation methodology were not addressed in any detail, however. Further, given how rapidly this field has continued to evolve, aspects of the reports already appear somewhat dated.

### *2.6.3 Searching for impact and methods: NGO evaluation synthesis study. A report prepared for the OECD/DAC Expert Group on Evaluation (1997c)*

Although covering quite different ground from that of democracy and governance, a recent DAC synthesis study of NGO evaluations provides useful findings and observations on matters of evaluation *process*, especially relating to (pseudo-) participation and impact evaluation. The study has some parallels with this research, involving a desk review of evaluations of NGO-implemented projects, although the focus on *project* evaluation is distinct. Its objectives were two-fold: to assess the impact of NGO development interventions by means of a study of evaluation reports; and to review the methods used in those reports. Interest here focuses on the issue of methods, with two points highlighted.

First, despite the rhetoric about partnership and participation, evaluation exercises continue to be perceived as essentially “a top-down, externally-driven exercise”, viewed “more as an audit to ensure that funds are well spent than a process dominated by the desire to learn in order to enhance future impact” (OECD 1997c, chapter 6, no page numbers given). This appears to be a common feeling amongst those who are subjects of the evaluation, at whatever level: “This is not only a complaint which northern NGOs

make about donor-commissioned evaluations; it would appear to be complaint which southern NGOs make about northern NGO evaluations and which community-based organisations make about southern intermediary NGOs” (chapter 6). Clearly, such feelings and perceptions require addressing by all evaluations with a focus on learning lessons rather than financial accountability, and by those with a participatory dimension in particular.

Second, how judgements concerning impact are made was questioned in a number of the studies examined. While most studies “describe in detail the specific factors against which they assess impact, they provide very little detail about precisely how they arrive at the judgements made” (chapter 7.1). It is stated that “extensive comments and judgements [are] made on impact... [but]there is a paucity of information detailing *how* these judgements were made” (chapter 7.1). The lack of such methodological clarity leads the report to deduce that impact assessment is frequently a matter of ‘personal judgement’. The general conclusion drawn is that donor-commissioned evaluation studies do not provide sufficient guidance on important ‘*how*’ questions as, in this instance, “how to assess impact, how to undertake cost-effectiveness analysis with minimal data, how to assess NGOs’ ability to innovate, how to assess NGOs’ flexibility, or how to undertake a gender or environmental analysis of NGO development interventions” (chapter 7.1). This point is of particular pertinence with regard to impact assessments at country level of donor democracy and governance programmes, where methodological difficulties are especially complex.

## ***2.7 United Nations Development Programme***

With the publication of a policy document in January 1997 (UNDP 1997a), the promotion of governance became a major theme of UNDP’s activities, described as the fastest growing area of its work and may indeed now be the largest area. Good governance features in UNDP’s mission statement, and it is stated that good governance underpins *all* four aspects of UNDP’s overall objective of sustainable human development, i.e. eliminating poverty, creating jobs and sustaining livelihoods, protecting the environment and promoting the advancement of women (UNDP 1997a, chapter 1, p.2).

This section examines three evaluation studies undertaken by UNDP, a general review, and a regional study on Latin America and the Caribbean, and a thematic study on decentralisation and local governance.

### ***2.7.1 UNDP and Governance: Experiences and Lessons Learned (1998a)***

Little more than a year after its policy paper in early 1997 (UNDP 1997a), UNDP published a review of its activities in the governance field and the lessons learned (UNDP 1998a). Although not an evaluation as such, there are clear overlaps with an evaluative study given the aim of learning lessons from experiences. The stated purpose is “to trace the evolution of UNDP’s approach to governance and to summarise the key trends and programme activities” (UNDP 1998a, p.2). The methodology is essentially a review of

relevant documentation from UNDP projects, workshops and conferences, including project evaluations, of which two or three are listed in the bibliography. The resulting paper is self-admittedly “descriptive”, regarded as defensible as “much of the UNDP experience in the field of governance is relatively new, and the availability of evaluative surveys is limited” (p.3).

Two introductory chapters provide an overview of, first, the rise of governance as a key element in a new development model, and, second, the evolution of governance policy within UNDP. Subsequently, the report devotes a chapter to each of UNDP’s five governance focus areas, providing a summary of activities to date, including country examples, and of emerging lessons. Surprisingly, only the chapter on ‘governing institutions’ tackles the question of evaluation. It notes that UNDP’s experience in this particular governance field is relatively new and that “project evaluation... is problematic but essential” (p.27). Problems outlined are the familiar ones of causality and the inappropriateness of quantitative approaches: “existing methodologies and techniques often pos[ing] problems in demonstrating causality among data, outputs and inputs, and governing institution programmes often do not lend themselves to quantitative analysis” (ibid.). The evaluation of success is consequently “often dependent on *perceptions* of institutional or societal changes” (emphasis added), with such interpretations presumably deemed both subjective and insufficiently linked to project activities. It is noted that “qualitative indicators have not been developed for many of these types of projects” and that “evaluation methodologies must be rethought” (ibid.). Action in this respect appears to be promised in the report’s overall conclusion with the statement that “Evaluation methods will be developed further and tested” (p.43), outlined as one of a small number of planned initiatives in the governance field.

### *2.7.2 Evaluation of the Governance Programme for Latin America and the Caribbean (1998b)*

At approximately the same time as the general review of its governance activities (UNDP 1998a), an evaluation of UNDP’s governance programme in Latin America and the Caribbean was completed (UNDP 1998b). Objectives were stated as follows:

- to determine to what extent the co-operation provided by UNDP had contributed to the achievement of concrete results;
- to establish the relevance of UNDP’s activities;
- to summarise lessons learned;

(p.10).

The coverage of this study was very broad, both geographically and thematically, with the latter comprising: the promotion of democratic governance; consensus-building; governance and human rights; democracy and citizen participation; citizen auditing of the quality of democracy; reform of the justice system; public safety (i.e. the protection of persons and property); reform of the electoral system; modernization of public

administration; decentralization. A number of projects were evaluated under each thematic heading, a total of 31 in all.

At the time of this evaluation, UNDP did “not have a defined set of results indicators for projects in the area of governance” (p.4), and it is stated that it was not possible to measure results in this evaluation. In common with many of the other reports covered in this review, the authors instead draw a number of qualitative conclusions (p.4). Problems arise, however, with the lack of methodological clarity and with the lack of transparency concerning the conclusions drawn, as explored below.

It is unclear whether the report intends to draw conclusions on impact or not. The Foreword states that the aim of the paper is to “obtain an overall view of what UNDP has achieved in the region” (p.3). In some places, the authors recognise that it is “extremely difficult to evaluate in isolation the impact and the successes of [a] project” (p.34), yet attempt to do so. Elsewhere they note simply that “the scope of the evaluation does not allow us to measure the impact of the project” (p.20). Whether or not the authors are looking at achievements or at impacts, they consistently highlight project outputs rather than outcomes. The combination of over-ambitious coverage and the lack of specific performance indicators leads to a study that is more a *review* of activities than an evaluation, despite document title.

Nevertheless, at times the study does draw impact-related conclusions, but these are somewhat undermined by methodological problems. For example, the authors find positively that activities relating to consensus building in societies traumatised by violence, and those relating to the defence of human rights, had “the greatest political and social support” (p.6), but it is not made clear what measures of support have been made, and how this conclusion is reached. While it may be reasonable to assume that citizens will support processes that lead to the cessation of violence, as in the authors’ example of Panama (pp.26-31), the authors make no effort to gauge popular support for elite consensus-building negotiations, even anecdotally.

In another sector, a project promoting women’s participation in public life is said to have had a “clear” impact on participants, “reflected in ... an increase in the proportion of female councillors and mayors elected” (p.45). This would appear to be a rather grandiose claim, insufficiently grounded in evidence and with little consideration of the overall political context. On the one hand, the scale of the project inputs are limited (namely 15 seminars held across Brazil to provide information for women candidates); while, on the other, the range of factors that contribute to a successful election campaign are extensive (not least the standing of the party in the locale; the candidate’s support network; campaign finances; media presence, etc.). In this example, the linkage between the inputs and the success claimed for the project (an increase from 3,085 women councillors to 6,536 – p.45) is tenuous when one considers that the government had reserved a quota of 20% of seats for women, and therefore, that the number of women returned could reasonably be expected to go up! Although the project may well have helped to promote competition for these new vacancies, it almost certainly did not single-handedly cause three and a half thousand new women councillors to be elected. In this example, the evaluators can be said, at best, not to give sufficient credit to external

factors (although in fairness they do acknowledge their existence); or, at worst, to claim successes for projects when the results were overwhelmingly achieved by other means.

The report also draws conclusions with regard to project weaknesses, particularly concerning a lack of coherence and sustainability. Regarding the former, it was stated that, “In some national programmes, the projects on governance appear just to be a collection of initiatives that are not always structured around a coherent strategy” (p.6). While the meagreness of government contributions meant that project achievements were often unsustainable after project completion (pp.6-7).

### *2.7.3 The UNDP role in decentralisation and local governance: A joint UNDP-Government of Germany evaluation (2000a)*

This report focuses on an important sub-theme of the governance agenda, decentralisation and democratic local government, and undertakes a geographically wide-ranging evaluation, with case-studies of activities in Guatemala, Mali, Philippines, Thailand and Uganda. It is a troubling report, however, containing inconsistencies and contradictions, as well as an apparent predilection for positive findings, despite the lack of supporting evidence at times.

The study is conducted by UNDP / BMZ, but it evaluates only UNDP support in this sub-field. Its stated objective is to identify “the main issues of decentralization and local governance at macro level”, and assess “the results of UNDP-supported programmes in these areas” (UNDP 2000a, p.2) against the background of “decentralisation policy in each country” (p.3). The consultants review the UNDP’s Strategic Results Framework, its programme planning, management and reporting tool, in the context of its decentralisation activities. Yet, the evaluation team then dismiss the Strategic Results Framework as a means for measuring impact “since little has been done in terms of definition, at the outset [of the programmes] of clear indicators and baselines for measurement” (p.15), and team instead chose to use a selection of the UNDP’s Higher Level Goals as a basis for the evaluation, namely:

- a) national programme conceptualisation;
- b) achievement of sustainable human development (SHD) objectives;
- c) developing capacities of programme countries;
- d) supporting people’s participation;
- e) fostering social cohesion;
- f) promoting gender and equity issues;
- g) building national ownership of programmes.

(p.16).

The choice of these goals for this evaluation is confusing since they skew the evaluation findings away from the Terms of Reference, as well as giving themselves a much harder,

if not impossible task. We have commented repeatedly on the difficulties of assessing the impact of DG programmes on democratisation, far less than the range of development goals outlined above. The first two points covered by the ToR are: “Identification and analysis of the main issues in decentralisation and local governance”; and “Assessment of results of UNDP-supported programmes” (p.60). Logically, it can be argued that the report, having identified the main issues in decentralisation, should have moved on to look at how the UNDP’s programmes address these issues. Instead, the authors choose to evaluate the programme in terms that are only indirectly relevant to decentralisation (agency level goals rather than programme or project specific objectives).

After the design of the evaluation framework, countries were selected for evaluation on the basis of their government’s active policy commitment to, and on-going UNDP involvement in, decentralisation. Country visits of between one and two weeks duration were undertaken with the five-member team working in groups of two and three, together interviewing 300 people in connection with the evaluation. The consultants were assisted by other UNDP staff (for the desk report), and national experts employed in each country to “provide basic information on national policy as well as on the country programme of UNDP”(p.3).

Following production of individual country reports (summarised in an annex), a synthesis report was presented to a “special stakeholder group” at UNDP headquarters in New York, and a subsequently revised version to a high-level workshop held in Berlin (ibid.).

Problems with this evaluation are most evident at the level of findings, partly stemming from the adoption of the higher level goals as performance criteria. One qualification repeatedly emphasised was that the programmes were only recently initiated and that decentralisation is a long-term process (p.15), indicating the difficulties of drawing conclusions at such an early stage of programmes. Yet, rather confusingly and contradictorily, it is stated that “UNDP was found to have shown early signs of contributing to seven higher-level goals... [but that] The evaluation team did not have time to verify or validate these findings” (p.16). The findings regarding achievement of each of the seven goals are given in a few brief paragraphs (pp.16-7), although the team do not draw clear links between claims of successes (or shortcomings) and the promotion of decentralisation. For example, under the goal of social cohesion, it was stated positively that, “In all 5 countries UNDP had been active in fostering the promotion of human rights”, (p.17). Two questions arise, but unanswered: why does promoting ‘social cohesion’ get interpreted as promoting ‘human rights’?; and how is the promotion of human rights (or social cohesion) related to decentralisation and local governance?

The team’s findings are logically troublesome. They write positively on the contribution of the interventions towards the achievement of the “broader development goals” of sustainable human development, but can find no “empirical proof that decentralization promoted economic growth, sustainable human development or poverty alleviation” because interventions are too recent to have had meaningful effects (p.vi).

Interestingly, some of the failings discovered in the programme are replicated in the evaluation itself, notably in relation to participation. For example, the authors state, “Few of the projects and programmes were found to address the needs of civil society and especially the private sector adequately” (p.16). In conducting their research, the team

claim in one instance to have “met and listened to approximately 300 people from all walks of life” (pp.2-3), but later revisit this, writing: “The people the team interviewed were mainly public servants and not representatives from the private sector (hardly any) and civil society (very few)” (p.19 – parentheses omitted). Another observation concerned the importance of adapting programmes to local country contexts and some failure to do so, questioning the reliance on and quality of input from ‘international experts’, given some “lack of sensitivity and understanding of the national situation” (p.24). A greater role for national experts in the evaluation process would also seem an appropriate deduction.

While this is a clearly structured report, there are troubling inconsistencies in its approach and methodology which lead one to seriously question whether it adequately addresses the key issues relevant to decentralisation. It appears to be another instance of seeking to present positive findings on impact at macro level, yet with little evidence on which to base such claims, nor indication of how such conclusions were reached.

### *Summary*

These three UNDP reports are all unsatisfactory in different ways. The general report and the Latin America and the Caribbean study are more reviews than evaluations, despite titles. Both the regional and thematic studies suffer from a poor methodology, with insufficient thought given to how to evaluate impact, especially at macro level. Findings and conclusions are not convincing in either of these studies. Optimistic claims regarding impact are made in both cases, without supporting evidence in the ‘decentralisation’ study and without establishing clear linkages from micro level projects to macro level assertions in the Latin America and Caribbean study. Levels of participation are also low in all cases, coinciding with a UNDP publication on the merits of participatory evaluation (UNDP 1997b), examined in Chapter 5.

Issues raised in all the individual reports examined in this chapter are returned to and discussed in Chapter 4. Prior to this, Chapter 3 examines a wider range of some sixty evaluation studies, highlighting their similarities and differences in a comparative manner.



## Chapter 3.

# COMPARATIVE ANALYSIS OF EVALUATION STUDIES

This chapter is based on a larger data set of some sixty evaluation reports in the democracy and governance field, predominately programme rather than project evaluations. These were selected in turn from the one hundred and ten studies revealed by our research into donor-sponsored democracy and governance evaluation, with selection criteria outlined below. The chapter undertakes an analysis of the sixty reports, covering their scope, thematic emphasis, methodologies and methods, impact evaluation, and findings.

The initial section introduces the database and methods employed in undertaking the comparative analysis, followed by sections presenting the analysis and findings on each aspect.

### *3.1 Introduction to Database*

#### *Description of database*

Analysis and comparison of reports was facilitated by the creation of an Access database. For each report, basic bibliographical information was recorded, along with information under the following categories. Information was collected either under closely defined headings or as general textual commentary, as appropriate.

- scope of the report (whether project; programme; discussion of evaluation issues, strategy etc.)
- geographical coverage (name of country/ies or region/s)
- donor coverage (coded bilateral and multilateral; single or multi-donor)
- principal themes (coded democracy, governance, human rights, and combinations)
- sub-themes (free text)
- evaluation objectives
- methodology
- methods
  - desk study (yes/no)
  - interviews 1 (listing the types of interview conducted)

- interviews 2 (listing other means of obtaining stakeholder input – e.g. focus groups; surveys)
- socio-political contextual study (yes/no), (establishing whether an analysis of the overall context was conducted or obtained)
- impact evaluation (yes/no)
- type of impact evaluation (at micro, meso, macro levels)
- whether explicit criteria used or developed for the purpose of the evaluation (yes/no – plus another category for those based on logframe analysis)
- findings
- comments on participation
- general comments

The sub-categories under ‘Methods’ provided the basis for the analysis contained in section 2.5 below. To facilitate analysis, we extracted the data to a spreadsheet, creating a matrix showing the combinations of methods used in each of the 60 reports.

### *Selecting the reports*

As stated, of the 110 reports unearthed by this research, 60 were studied in detail for this report and appear in the project database. What were the criteria for inclusion or exclusion from the database? In general, project evaluations were excluded, given that our interest was limited to programme evaluations, though some were included where individual donor agencies had not undertaken an evaluation study at the programme level. A considerable proportion of ‘excluded’ reports are ones that were commissioned as part of a series, for example the USAID/CDIE evaluations of Elections; Democratic Local Government; or Legislative Assistance (see USAID 1997a; 1997e; and 2000d for details). Such series generally consist of five or six country case studies and a synthesis report, and for reasons of time, we typically confined ourselves to reading one or two case study reports plus the synthesis report. This allowed us to gain a sense of findings and working methods. A similar category of exclusions encompasses studies that were also included in later synthesis reports (such as CIDA 1996), although they were not necessarily commissioned as part of a series (for instance CIDA 1994a and CIDA 1995c). Some reports were dismissed because, after preliminary reading, we decided they were not strictly relevant for the purposes of this study. In this category come reports such as SIDA 2000e (on gender and empowerment); the Technical Publication Series by USAID’s Center for Democracy and Governance which, despite ‘indicating best practice and lessons learned’, tend to focus on future strategies rather than evaluating past experiences, for example, USAID 1999f (on media and democracy strategy), USAID 1999d (on Political Party Development Assistance). A related category of reports are those that are not strictly evaluations, but do contain highly relevant discussions about evaluation methodology, e.g. IDRC 1999. Such reports are not included in this database, but have already been examined in chapter 2.

Resulting from this selection process, a database of some sixty evaluation reports was built up, analysed below.

### 3.2 Scope of Evaluation

Table 3-1: Evaluation Report Coverage

Donor country programme	30
Multiple case studies (country and/or thematic) from donor's global programme	16
Donor project	13
Multi-donor programme	1
<b>Total</b>	<b>60</b>

Source: research database

Table 3-1 indicates the coverage by subject matter of the evaluation reports examined. It shows that half of the evaluation studies (30) focused on a single country programme by that individual donor, for example, the South Africa-Canada Program on Governance (Sutherland 1999). There were also a significant number of studies by individual donors (16) that incorporated multiple country case studies from their global programmes, generally from three to six, often with a thematic focus, for instance USAID's report on 'Democratic Local Governance' (USAID 1998b). Thirteen project evaluations were also included, usually where the donor agencies concerned had not undertaken programme evaluations, plus some instances where project evaluations were instructive in highlighting particular methodological issues, such as Sida's reports on local self-governance in Mongolia (Sida 2000e) and on rural court houses in Nicaragua (Sida 1999b). Only one report covers the activities of multiple donors in a particular country, that is a study of overall aid effectiveness in Mali, initiated and published jointly by the Club du Sahel, UNDP and the Development Co-operation Directorate of the OECD. This is not specific to democracy and governance assistance, but included as it is instructive on issues pertaining to undertaking a collaborative evaluation on a participatory basis (see Appendix One).

Table 3-2: Donor Coverage

Single bilateral donor	42
Single multilateral donor	7
Multi donor	11
	60

Source: research database

This Table indicates the coverage by donor of the 60 evaluation reports included in the database. The large majority examined the country or thematic programme of either a single bilateral donor, for instance Danish support to human rights and democracy in Guatemala, or that of a single multilateral organisation, for example UNDP's 'Governance Programme for Latin America and the Caribbean'. In eleven instances, reports cover the efforts of more than one donor, termed 'multi-donor' here. It must be clarified that, with the exception of the Mali study (Club du Sahel *et al.* 1999), these are not collaborative ventures between donors. Most commonly, they entail studies by a

single agency, for example USAID’s CDIE reports on ‘democratic local governance’ and ‘electoral assistance’, which take activities by other donors into their purview to a limited extent, though effectively concentrate on the activities of that agency. Three studies are more genuinely multi-donor in scope, although this is more a quirk of how a particular project or programme was funded and/or implemented, for instance a UNDP evaluation of assistance to Tanzania’s National Assembly, funded by various bilateral donors. Only the Mali report (Club du Sahel *et al.* 1999) is a genuinely collaborative study of the development assistance efforts of the range of donors active there. This lack of donor co-ordination in conducting evaluations is somewhat disappointing, if not unexpected. The need for co-ordination in the democracy promotion field, both between donors and with host countries, is much proclaimed, but not accomplished. This is disappointing as collaborative evaluations of lessons learned from past activities on a country basis could potentially lead to more collective efforts in the future.

### 3.3 Geographical Coverage

Table 3-3: Geographical Coverage

Country-focus [listed by region]	<p>Angola (1); Burundi (1); Ethiopia (3); Ghana (4); Kenya (3); Lesotho (1); Mali (3); Mozambique (9); Namibia (1); Rwanda (1); South Africa (9); Tanzania (2); Uganda (4); Zambia (1); Zimbabwe (1)  Sub-total: African nations (44)</p> <p>Bangladesh (1); Bhutan (1); Cambodia (3); Indonesia (3); Mongolia (1); Nepal (7); Pakistan (1); Papua New Guinea (1); Philippines (4); Sri Lanka (1); Thailand (2); Vietnam (4)  Sub-total: Asian nations (29)</p> <p>El Salvador (6); Guatemala (6); Honduras (1); Jamaica (1); Nicaragua (3, excluding 1 sub-state); St. Kitts-Nevis (1)  Sub-total: Central American and Caribbean nations (18)</p> <p>Bolivia (3); Chile (1); Peru (1)  Sub-total: South American nations (5)</p> <p>Palestinian Territories (3)  Sub-total: Middle East (3)</p> <p>Ukraine (1)  Sub-total: Eastern European nations (1)</p> <p style="text-align: right;"><b>Total</b> 91</p>
Sub-state focus	<p>Karnataka (India) (1)</p> <p>Autonomous Atlantic Regions (Nicaragua) (1)</p> <p style="text-align: right;"><b>Total</b> 2</p>

Regional focus	Africa, Caribbean and Pacific (2)	
	Asia and Latin America (1)	
	Latin America and Caribbean (1)	
	Central America (1)	
	Mediterranean (2)	
	Sub-Saharan Africa (1)	
	West Africa (1)	
	Central and Eastern Europe (1)	
Global (8)	<b>Total</b>	18

Source: research database

### 3.4 Thematic Coverage

Thematic coverage is examined in two main ways, one general and one more specific. First, evaluation reports were categorised in terms of their focus on the main concepts of democracy, human rights or good governance, or combinations, as identified by the donors themselves. Second, reports were categorised in terms of any sub-thematic emphasis.

Table 3-4: Thematic Coverage

Theme	No. of reports	Sub-themes (Incidence)
Democracy	19	Civil society (5); Decentralisation (7); Elections (5); Institutional strengthening (2); Legislative strengthening (1); Media (1); Tolerance (1);
Human Rights	8	Civil society (1); Community policing (1); Rule of law (4); Without sub-theme (2)
Good Governance	5	Institutional strengthening (2); Legislative strengthening (1); Without sub-theme (2)
Democracy/ Human Rights	14	Accountability (1); Aid management (1); Civil Society (2); Decentralisation (2); Elections (4); Empowerment (1); Human Rights (Civil and Political Rights) (1); Legislative strengthening (1); Media (3); Participation (2); Rule of law (2); Without sub-theme (6)
Democracy/ Governance	7	Gender (1); Poverty reduction (1); Institutional strengthening (1); Legislative Strengthening (4); Media (1); Prisons (1); Political parties (1)
Democracy/ Governance/ Human Rights	2	Without sub-theme (2)
General	5	Aid management (1); Decentralisation (1); Evaluation methodologies (1); Human rights (civil and political) (1); Participation (1); Institutional strengthening (1); Review of evaluation reports (1); Rule of law (1); Without sub-theme (1)
<b>Total</b>	<b>60</b>	

Source: research database

Table 3-4 indicates that the notion of democracy is the most central, confirming it as the overarching concept to which the other two are interrelated (Crawford 2000a, p.24). The so-called ‘like-minded’ donors, plus the European Commission, continue to identify their programmes as supporting ‘human rights and democracy’, though essentially their focus is on civil and political rights. The term ‘governance’ is favoured by the multilateral organisations, such as UNDP and the World Bank, and by a few bilaterals, for example, UK DfID and AusAid. CIDA appears to have moved from a focus on ‘democratic development and human rights’ to a more recent emphasis on ‘governance’, while USAID sticks with the phrase ‘democracy and governance’ to describe its programmes. Such differences are not necessarily merely ones of labelling, but can also lead to differential emphases in programme activities. Nevertheless, it is evident that democracy and democratisation remain of central concern in this overall field of political aid.

Table 3-5: Sub-thematic Focus

Accountability	1
Aid management	2
Civil society	8
Community policing	1
Decentralisation	10
Elections	9
Empowerment	1
Evaluation methodologies	1
Gender	1
Human Rights Civil and Political Rights	2
Institutional strengthening	6
Legislative strengthening	7
Media	5
Participation	3
Political parties	1
Poverty reduction	1
Prisons	1
Review of evaluation reports	1
Rule of law	7
Tolerance	1
Without sub-theme	13
<b>Total</b>	<b>82</b>

Source: research database

As 13 reports had no sub-thematic emphasis, 47 reports generated 82 sub-themes. Five sub-themes stand out: decentralisation, elections, civil society, rule of law and legislative strengthening. These indicate those areas of greatest interest to donors in terms of evaluating the effectiveness of their interventions, and are also suggestive of those sub-fields of democracy promotion that attract most donor attention and finance. A focus on the evaluation of electoral activities is fairly unsurprising, given the initial emphasis by many donors on issues of transition in the first half of the 1990s and beyond. Other government-related themes concentrate more on those institutions, parliament and the legal system, that perform accountability functions in relation to the executive, in line with liberal notions concerning the curtailment of the arbitrary exercise of power by governments through democratic procedures. Similarly, the relative emphasis on

evaluations of civil society assistance reflects donor interests in strengthening the non-government sector in its watchdog and advocacy roles in the democratic process (Van Rooy 1998). Perhaps most surprising is the focus on decentralisation. While decentralisation is clearly an important theme of democratisation, with the potential for enhanced participation in democratic decision-making at local levels, donor advocacy of decentralisation can be more problematic. First, decisions on decentralisation pertain to constitutional issues concerning the respective roles and responsibilities between different levels of government, to be determined endogenously, without intrusion on national sovereignty. Second, decentralisation in itself does not necessarily strengthen democracy, but depends on local power structures and the democratic character of local government (Crawford 1995, p.67-8). Media assistance is another specific area that has been subject to a number of evaluation efforts.

### ***3.5 Methodologies and Methods***

#### *Methodologies*

Methodology is distinguished here from methods. The former is a theoretically informed approach to evaluation, generally located within a broad theoretical paradigm, while the latter are techniques, not necessarily specific to a particular methodology, though some methods correlate more closely with one methodology.

Categorising evaluation methodology was not straightforward. For a significant number of reports (13), the nature of the methodology was neither explicitly discussed nor implicitly inferred. Others did not adopt a strict methodology as such, most commonly where the study was more a general review of policy implementation in the political aid field (12 reports in total), for instance the Norwegian evaluation report on human rights and democracy (Ministry of Foreign Affairs 1997b) (see Appendix One). However, referring back to the introductory discussion of the main two evaluation approaches, conventional and participatory, it can be stated confidently that the large majority of studies followed a conventional methodology, though with variations. The most common pattern entailed an initial desk review of project / programme documents, followed by a short field trip, with the evaluation report based on information gathered. Stakeholders or beneficiaries generally played a passive role, as sources of information only. Quantitative methods were often preferred, viewed as more reliable, for example, in USAID's (1998a) Handbook of Indicators. The use of quantitative indicators was favoured in particular in attempting to trace 'cause and effect' linkages, for instance, in UNDP's evaluation of the governance programme for Latin America and the Caribbean (1998b). External consultants (that is, from donor countries) were invariably hired to conduct evaluations, preferred for their assumed expertise and independence. The methodology on which such methods are based is traditional, scientific inquiry, which perceives that a mixture of external expertise and established research design will provide objective and value-free findings. Those studies that are specifically based on a logical framework approach or 'results-based management' (USAID, CIDA) fit in most closely with this conventional paradigm, though the latter 'results-based' enquiries are conducted through internal reporting mechanisms. Despite the questioning of its suitability in this field by some donors (Danida 2000, vol.1, p.66), the ongoing influence of the logical framework

approach is evident from the Sida (2000b) 'evaluability' study, an extensive report dedicated to this approach, and from ten studies being explicitly based on logframe methodology. Donor agencies rather than evaluators determine evaluation methodology on the whole, for instance, through application of a standard agency-wide approach to evaluation (for example, CIDA 2000a), and through determining 'terms of reference', with limited examples of evaluators being able to develop their own criteria and indicators.

Although a number of studies have a participatory element, discussed below, very few could be described as participatory evaluations. Many evaluations include stakeholder perspectives, but an important distinction is between evaluations where stakeholders remain objects or become subjects of the evaluation process (Rebien 1996, p.67). A participatory dimension may be integrated in order to represent beneficiaries' opinions in the evaluation report, as observed with the use of participatory data collection methods in studies examined here. Yet, such methods are essentially a means of improving data collection, while the stakeholders remain as *objects* of the evaluation. A more genuinely participatory methodology entails stakeholders as *subjects* of the evaluation process, determining objectives and identifying information needs.

Rebien has developed three threshold criteria to be met for an evaluation to be counted as participatory: (a) stakeholders must be involved as active subjects rather than as passive objects or data sources; (b) stakeholders should be involved in at least the design of terms of reference, data analysis and use; and (c) at least beneficiaries, field staff, intervention management and donor representatives should be involved (1996, pp.67-9). In applying these three criteria to the sixty evaluation studies examined, only two loosely fulfil such criteria and approximate the designation of being truly participatory. Coincidentally, these are both in El Salvador: a relatively small-scale USAID evaluation of the Centro DEMOS programme (USAID 1995b), and the Sida-funded Diakonia programme for democracy and human rights in El Salvador case (Sida 1999a). The Centro DEMOS report involved recipients and beneficiaries in the design stage, including the development of the 'terms of reference', and in conducting the evaluation. Even so, its participatory nature was qualified by some participants who drew attention to the downside of the financial costs of external consultants and the constraints this placed on time (i.e. the budget was quickly depleted) and on the workings of the participatory process. The evaluation of Diakonia's programme involved grantees in developing the proposed quantitative methodology and in selecting data for inclusion.

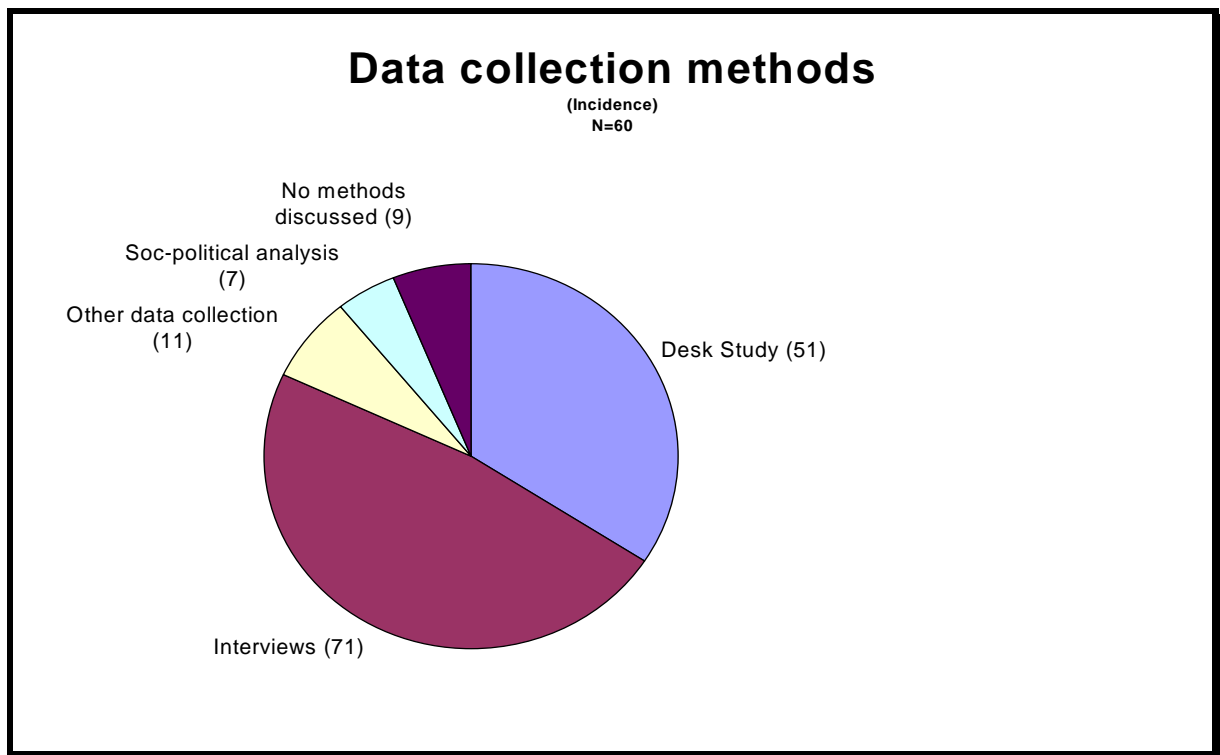
Nevertheless, while no official agency has adopted a thoroughgoing participatory approach ('negotiation mode'), some convergence between the two distinct approaches is evident, with increasing pluralism. On the one side, greater flexibility with the application of LFA is suggested due to its difficulties in dealing with 'soft' data (for instance, European Commission 1999b), with recognition that causal attribution is difficult if not impossible (Schacter 2001). On the other side, integration of participatory evaluation methods is increasingly emphasised, significantly by the DAC Expert Group (1997a), with consultation of 'stakeholders' becoming almost obligatory.



## Methods

This section analyses the main methods used in the reports examined. The most common format of these evaluation studies entailed an initial study of documentation, followed by a brief fieldwork visit (commonly 2-3 weeks duration) to examine the political aid programme(s) in question. Thus, three distinct methods were identified: a desk study of documentation; fieldwork interviews as a means of data gathering from a range of stakeholders; and socio-political analysis, pertaining to whether a study of the overall country context was conducted or obtained. Other less common data collection methods, such as focus groups, were classified under miscellaneous.

Figure 3-1: Data Collection Methods

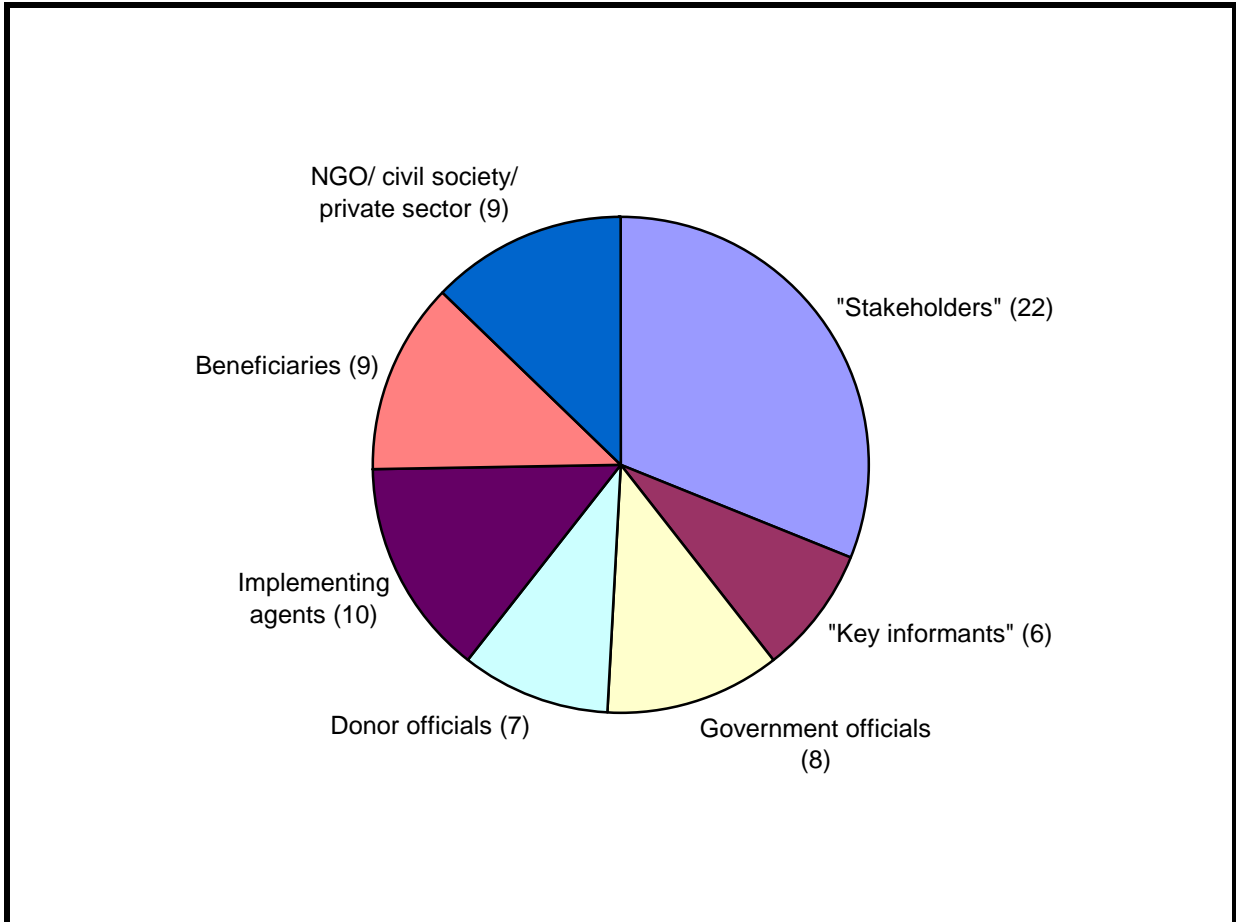


Source: research database

Figure 3-1 shows that desk studies of documentation and interviews were the most common methods employed. Of the 60 reports examined, nine did not discuss or outline the methods used. Of the remaining 51, all either explicitly state that they have undertaken a desk study, or can reasonably be assumed to have done so. In eight cases, a desk study was the only method employed. A desk study plus interviews were the most common combination of methods. Interviews were conducted in 39 studies. Interviews were categorised by type of interviewee, giving 71 incidences of interviews in the 39 studies. (See Figure 3-2 below). Disappointingly, a background study analysing the social-political situation in a particular country was undertaken in only seven instances. Of this small group, three studies were explicit about the methods drawn on, including a democratic audit (European Commission 1997); newspaper surveys (USAID 1996b); and reviews of academic literature (USAID 1998b). The remaining four reports did not specify the specific methods used to develop an analysis of the political context in which

the agency was working (SIDA 1999c, Dutch Ministry of Foreign Affairs /IOB 1999, USAID 1995b and CIDA 1995b). Other miscellaneous methods of data collection (eleven incidences) entailed: questionnaires conducted among implementing agents (2 incidences) and beneficiaries (1 case); an informal survey of organisations participating in the programme (1 case); public opinion surveys (3 incidences); and focus groups (4 incidences).

Figure 3-2: Types of Interview



Source: research database

Interviews were clearly the main means of information gathering. Figure 3-2 sub-divides the 71 incidences of interviews by type of interviewee. Categorisation was based on information available. At times, information was limited to donor descriptions of 'stakeholders' or 'key informants'. Such categories are clearly less precise and likely themselves to contain instances of interviews with beneficiaries or implementing agents etc. No duplication has occurred, however. It is apparent that a range of individuals and groups involved in project or programme implementation were commonly interviewed as part of the evaluation process. Whether the interviews were used merely for purposes of 'factual' data collection or whether the views and perspectives of interviewees on the

programme were sought is a separate question, of course. What is evident from this analysis, nonetheless, is that there is a limited focus on those actors that are less directly involved in the programming process. Most frequently, interviews are conducted with those directly involved in project or programme implementation, including representatives of donor agencies and recipient governments, as well as implementing agents and beneficiaries. Only in nine out of 71 incidences were interviews categorised as with civil society and other groups, often less directly involved, though affected by the project/programme nonetheless. (It is possible that more incidences are concealed in the general ‘stakeholders’ category, but based on the reports where more detailed information was available, it is felt that consultation of *indirect* stakeholders only occurs in a small percentage of cases).

Table 3-6: Participation in evaluation

Workshop	1
Roundtable discussion	3
First draft of report discussed by stakeholder group	2
First draft discussed by donor group	3
Evaluation ToR co-developed by recipients/ beneficiaries	3
Recipients/ beneficiaries engaged in conducting evaluation	2

Source: research database

As referred to above, a total of eight reports involved a degree of participation in the evaluation, although in only two cases were Rebien’s (1996) threshold criteria met for the title of a participatory evaluation. The variety of participatory methods found in the eight reports, and the number of incidences, is presented in Table 3-6. This excludes the more routine inclusion of interviews with ‘stakeholders’, beneficiaries and implementing agents as a simple data gathering exercise. Participatory methods range from initial Workshop and Roundtable discussions to decide the scope of the study, to joint development of the terms of reference, engagement in conducting the evaluation, and input into the draft report.

Another report with a strong participative element is the Mali study on overall aid effectiveness (Club du Sahel et al. 1999) (see Appendix One). Although not specific to DG assistance, this collaborative donor evaluation is instructive for its relatively high degree of involvement of national actors, including the use of Malian consultants as evaluators and of a Malian NGO to provide ‘civil society perceptions and proposals’. The latter was achieved by various means of ‘direct participation’: a survey questionnaire, a public conference and radio discussions in local languages.

### 3.6 Assessing Impact

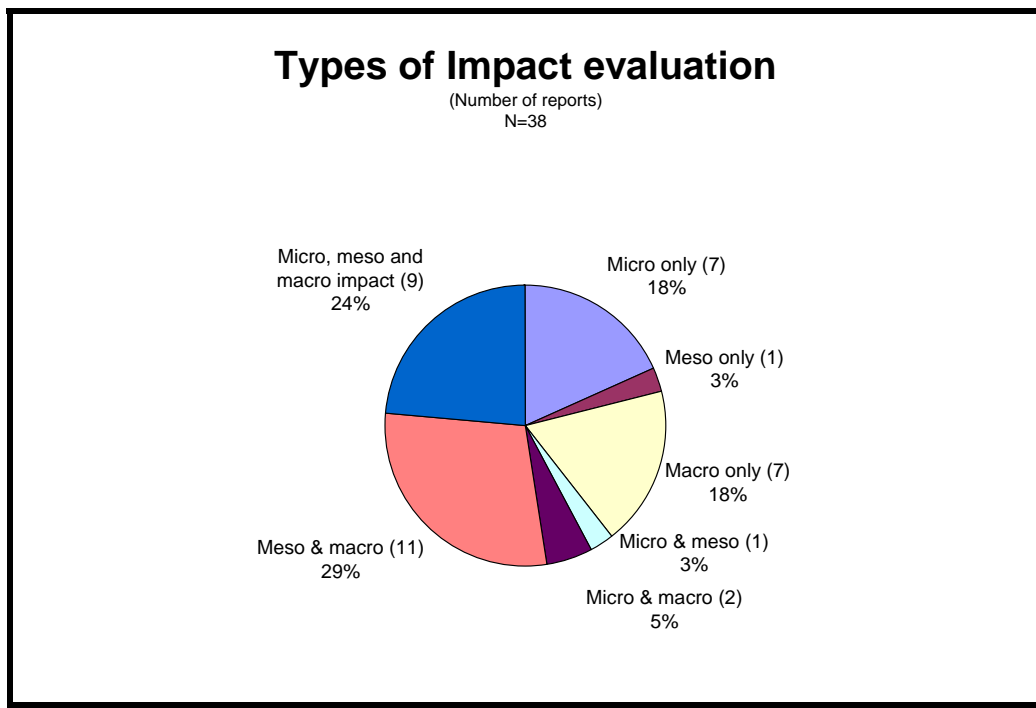
The impact of a project or programme (i.e. the changes that have occurred as a result of the intervention) can be assessed at three distinct levels: micro, meso and macro.

- **Micro** is the impact of the project / programme at the local level, i.e. the impact on single organisations and institutions, such as the strengthening of a specific NGO or the legislature, often corresponding to the fulfilment of project / programme objectives, at times as stated in a project’s logical framework matrix.

- **Meso** is the impact at sectoral or thematic level, e.g. the impact on civil society (rather than on a specific NGO); on the legal system (rather than on the judiciary or police); on democratic decentralisation; on women’s role in politics, and so forth. The meso level can also be described as the *partial regime* level.
- **Macro** is the impact at the national level in terms of the overall political regime and its movement along the democratisation continuum, either progress towards (or regression from) a democratic polity.

Of the sixty evaluation studies examined, 38 had undertaken some form of impact assessment. These 38 reports were categorised in terms of the level at which impact was evaluated, that is at micro, meso or macro levels, and combinations of such levels. This categorisation is illustrated in Figure 3-3. The reports are fairly widely distributed amongst the various types of impact evaluation, although relatively few studies examined impact at the meso level or combined micro and meso analysis. A number of studies have evaluated impact at the micro or macro level only, seven in each case.

Figure 3-3: Types of Impact Evaluation



Source: research database

Examples of different types are given as follows. Sida’s (2000a) evaluation of its assistance to local government in Mongolia is an example of micro level evaluation, evaluating the programme’s impact in terms of its own objectives rather than its possible impact on central-local government relations, or on the local/national political context. The programme is delivered by a variety of means, including the training of trainers in techniques for working in local government, and twinning arrangements between Mongolian and Swedish cities. Instances of the micro level focus are provided by the indicators of success. Commenting on the training element in the programme, the authors

note: “Another important factor for the success of the project is that the subjects of the training [i.e. its content] have proved to be valid for the daily work of the participants (Sida 2000a, p.1). Another indicator of success is the “clear evidence of increased knowledge and capacity” found among those trained (p.1). In neither case is there any discussion of the possible wider effects.

The European Commission’s (2000a) evaluation of its country strategy in Papua New Guinea provides a good example of a study looking mostly at the macro (regime) level impacts of a programme. The report’s clear focus on the ‘big picture’ is due to its concern with the effectiveness of the Commission’s *country* strategy, with clear terms of reference providing guidance. The authors’ findings are negative regarding impact in the governance realm. They state that the Commission’s concerns about governance has led to the creation of parallel institutions, rather than to efforts to strengthen existing ones. This has meant that a potentially important area for governance work has been ignored, and unless corrected, may have long term implications for government capacity and local ownership of initiatives (p.48).

A significant number of studies undertook a linked examination of impact at meso and macro levels (11 instances), and also at all three levels (9 instances). For the former category, the report by Biddle, Clegg & Whetton (1999) of DfID/ODA support to the police in developing countries, provides a good example of a focus on meso and macro level impact [see Appendix One]. The report is based largely on existing project evaluations and is guided by the question, “To what extent have projects been able to achieve effective policing and how could effectiveness be increased?” (p.43), encompassing the intermediate or meso level, and also dwells on macro level impacts, measured in terms of “wider criminal justice and good government goals” (p.i). Support to community policing is regarded as “special UK contribution to developing more accountable and accessible policing”. At the intermediate level, training has increased officers skills and led to better equipped police forces, but this has failed to contribute to macro level goals, because of an insufficient grasp of the importance of political will and of the commitment of police leaders. The authors find that “assisted police forces have been reluctant to mainstream community policing”; and that “Senior management have been concerned that community policing, whilst useful in image terms, could undermine their effectiveness in combating crime” (p.iv).

Regarding impact evaluation at all three levels, Sida’s (1999b) report on the impact of rural court houses in Nicaragua provides a good example. The terms of reference for this report stress that it is concerned with assessing the “extent to which this project has had an *impact* on the access to justice in Nicaragua, or – speaking in LFA terms – to what extent the project has contributed to the development objective” (p.3, emphasis in original). At this level, findings are negative: “the judicial system in Nicaragua is still largely inefficient and inaccessible. There are serious problems and shortcomings...” (p.3). At the intermediate level, however, the project “has contributed to the improved access to justice at local level”, by improving the status, visibility and (indirectly) the independence of judges and the court system (pp.3-4).

Table 3-7: Types of criteria / indicators used in impact evaluations

	Using logical framework criteria	Using other explicit criteria	Not using explicit criteria (pragmatic evaluation)	Total
Studies attempting impact assessment	12	9	17	38

Source: research database

A key question pertains to the criteria used in evaluating impact. From Table 3-7, it is disappointing that in almost half of the cases, no explicit criteria were stated, referred to here generously as a ‘pragmatic evaluation’. In other words, an assessment of impact was made, but without specifying the basis for, or the evidence to support, that judgement. For example, CIDA (1995b) finds positive outcomes for the Canadian programme supporting democracy and human rights in Sri Lanka, noting especially that the programme as a whole has increased NGOs’ abilities to address human rights questions (p.2), while not providing specific examples of such practice. Rather, the author takes a ‘pragmatic’ line, evaluating the programme in terms of a general feeling about its effects in terms of human rights. Consequently, he is positive about the outcomes of particular projects even where they do not accord with the objectives that the funding mechanism is supposed to fulfil. Other examples where impact evaluation appears to be based more on a hunch than specific assessment criteria include NORAD’s (2001) evaluation of support to human rights NGOs in Pakistan (see Appendix One), USAID’s (2000c) impact assessment of support in Indonesia, and its evaluation of post-conflict elections in Cambodia (USAID 1997c).

This concurs with the conclusions drawn in the DAC NGO study on ‘searching for impact’(DAC 1997c). It is recalled that this commented on the lack of information on how impact assessments are made, deducing that commonly such assessments are a matter of ‘personal judgement’ (see 2.6.3).

Where criteria were clearly stated (21 cases), 12 were based on the logical framework approach, that is using those criteria defined in the original project logframe. These reports vary considerably in the types of impact evaluation they make. Four evaluate impact at micro, meso & macro level (e.g. SIDA 1999b, discussed above); three at meso and macro level (e.g. evaluation of the European Commission Special Programme on South Africa, SPM Consultants 1996), and five at micro (e.g. Dutch Ministry of Foreign Affairs / IOB (1999). In the remaining nine instances, evaluators developed their own criteria or indicators against which to gauge programme outputs and/or outcomes. Of these, the European Commission’s evaluation of its country strategy in Papua New Guinea stands out as adopting the most ambitious criteria. This macro level evaluation states: “The EC’s performance in a country is to be principally assessed in terms of its impact on the EC’s priority objectives, in particular governance, poverty reduction, gender equality and environmental protection. These objectives are embodied in the Maastricht Treaty and in the revised Lomé Convention” (European Commission 2000a,

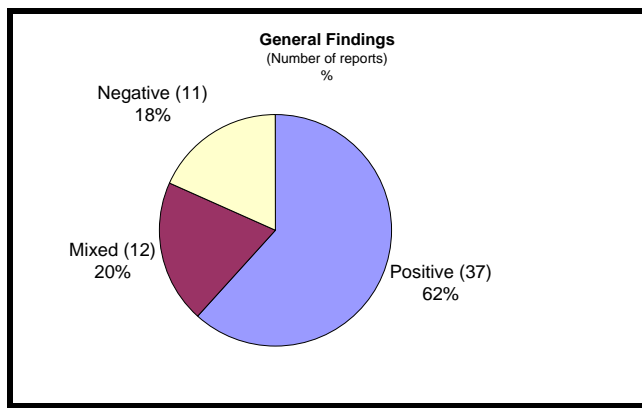
p.17). Determining appropriate criteria ex-post is more difficult where the focus is on more specific lower-level outcomes. It is interesting to note that none of the nine evaluations in this category attempt an impact assessment purely at micro level and that all reports have a macro element.

### 3.7 Evaluation Findings

This section examines the general findings of the reports studied and offers explanations with regard to the relative number of positive and negative findings.

Evaluation reports were categorised in terms of their overall findings, as shown in Figure 3-4. This indicated that evaluation findings were positive in the majority of instances, in 37 out of the 60 reports. Only in eleven studies were overall negative findings reported. In a further twelve studies, findings were mixed, neither overwhelmingly positive nor negative.

Figure 3-4: General Findings



Source: research database

Although almost two-thirds of evaluations reported positive findings, it is *not* concluded here that this accurately reflects the relative success and failure of donor political aid programmes. Rather, it is suggested that the overall nature of the findings is influenced by the function and character of the evaluation studies themselves. Four such factors are put forward.

First, one function of evaluation is as a managerial exercise to endorse donor agency activity. A number of mechanisms at different stages of the evaluation process, largely controlled by donor agencies, serve to influence positive outcomes. At the design and planning stages, donors determine of the terms of reference and methodology and delimit the framework of the investigation. At the implementation stage, consultation can be limited to a narrow range of stakeholders, commonly the beneficiaries of assistance.

Second, relatedly, key actors in the evaluation process often exist in a behoven relationship to donor agencies. Evaluators may be independent in theory, but dependent

on ongoing donor contracts in practice. Recipients of aid may be reliant on further assistance for their sustainability of their organisation or project. Neither is unlikely to ‘bite the hand that feeds them’.

Third, a lack of rigour has been apparent in some evaluation studies. This has been most noticeable in relationship to assessment of impact, where almost half of studies did not clearly specify assessment criteria, and where evaluation appears to be based on ‘personal judgement’, that is, unsubstantiated assumptions.

Fourth, a lack of critical analysis has been noted in some evaluation reports, for example, those on the ‘South Africa-Canada Program on Governance’ (Sutherland 1999), the European Commission’s (1999b) MEDA Democracy Programme, and DfID’s (1999) ‘Evaluation of the Indonesia National Police Management Training Project’ (see Appendix One). Consequently, affirmation of donor strategy and programmes is the likely outcome. A systematic evaluation requires not only a rigorous methodology, with assessment criteria clearly specified, but also the ability and predilection to engage critically with the material. This is especially important when the nature of external donor intervention concerns the difficult and contentious terrain of political institutions and processes.

It is postulated here that such factors can induce positive findings and partly account for the low incidence of negative findings.

It may also be instructive to examine those studies that had negative findings. Do they exhibit common features that may partly account for such findings? Two factors can be identified.

Table 3–8: General findings by type of impact assessment

	Generally positive	Mixed	Generally negative	Total
Micro only	6 (4)	1 (1)	0	7 (5)
Meso only	1 (0)	0	0	1 (0)
Macro only	3 (1)	0	4 (3)	7 (4)
Micro & meso	1 (1)	0	0	1 (1)
Micro & macro	1 (1)	1 (0)	0	2 (1)
Meso & macro	8 (4)	2 (1)	1 (0)	11 (5)
Micro, meso and macro	6 (3)	3 (2)	0	9 (5)
<b>Total</b>	26 (13)	7 (4)	5 (3)	38

Source: research database

Notes: No impact assessment undertaken = 22 reports

The number given in parentheses is the number of reports that explicitly state the indicators used to orientate findings.



The first factor concerns the type of impact evaluation attempted. Correlating findings with impact evaluation in Table 3-8 produces five reports with negative findings. It is notable that four of the five assessed macro impact only, while the fifth assessed meso and macro impact. In contrast, all but one of the micro impact assessments entail overall positive findings. This reinforces the difficulties of undertaking impact assessment at macro level without tracing clear linkages from micro and meso levels. A related phenomenon, however, has been a tendency to grandiose, but unsubstantiated claims of positive impact at the macro level (e.g. UNDP 1998b).

Second, the extent of participation in the ‘negative’ reports may also be a factor influencing their findings, that is, the greater the degree of participation, the higher the likelihood of more critical commentary and negative findings. Among these eleven reports were three that are characterised by their participatory elements, namely beneficiary involvement in developing the terms of reference (Club du Sahel et al. 1999); stakeholder involvement in discussing the final draft (European Commission 2000a); and similarly the presentation of a synthesis report to a ‘special stakeholder group’ (UNDP 2000a).

These two features tend to be replicated in the further twelve studies where findings were more mixed, neither overwhelmingly positive nor negative. First, correlating with impact evaluation (see Table 3-8) again shows a tendency to focus on the macro level: all but one of the seven reports attempting an impact assessment examines macro level impacts in combination with a lower level. Second, these studies as a whole used a variety of additional information gathering techniques, including participatory methods, for instance: public opinion surveys and focus groups (both USAID 1998b); workshops (USAID 1997e); questionnaires (USAID 1995b) and socio-political analysis (USAID 1998b and USAID 1995b). Additionally, the Danida study (2000, vol.9) on participation and empowerment comes into this category.

Table 3-9: General findings by scope of evaluation

	Generally positive	Mixed	Generally negative	Total
Country Programme	18	5	8	31
Country case studies from donor's global programme	10	6	0	16
Project	9	1	2	12
Multi-donor programme	0	0	1	1
<b>Total</b>	<b>37</b>	<b>12</b>	<b>11</b>	<b>60</b>

Source: research database

The breakdown of findings by the scope of evaluation indicates that negative findings occur most frequently in evaluation studies of country programmes. Eight such instances are concentrated in this category, a quarter of country programme evaluation reports. In

contrast, no examples were discovered of negative findings in the 16 instances of multiple country case-studies, commonly used in evaluations of thematic areas. This would imply that a country focus enables, though does not guarantee, a more detailed examination of impact and other effects within the particular country context, itself more likely to induce critical and potentially negative conclusions. In contrast, thematic studies, inclusive of a number of country cases, may not examine the individual country context in such detail.

## Chapter 4.

# CHALLENGES IN EVALUATING DEMOCRACY AND GOVERNANCE ASSISTANCE

The discussion of key donor evaluation studies (Chapter 2) and the comparative analysis of the wider range of reports (Chapter 3) have revealed a variety of issues and problems faced in evaluating democracy and governance assistance. These are summarised here under two main headings considering general evaluation methodology and impact assessment. The first questions the suitability of conventional evaluation methodology in this field, based on a critique of LFA. The second section examines the range of difficulties and problems associated with conducting an impact evaluation.

### *4.1 DG Evaluation: Art or Science?*

The range of reports examined (both above and in Appendix One) are characterised by diversity, covering programme and project evaluation, various thematic foci and a range of country evaluations. Yet one common feature is for almost all donor agencies to acknowledge the difficulties of evaluating the success or otherwise of political aid. Danida has expressed this most succinctly, stating that evaluation is “faced with a set of challenges that are more pronounced than in the assessment of other types of development assistance” (2000, vol.1, p.10). However, none have satisfactorily resolved these problems. Most agencies have attempted to address such dilemmas within the framework of conventional aid evaluation, although a few have questioned and criticised such approaches. From the preceding chapter (section 3.5), it is recalled that the large majority of studies adopted a conventional methodological approach, based on traditional, scientific inquiry and relying mainly on quantitative methods. The logical framework approach, inclusive of ‘results-based’ approaches (USAID, CIDA), is the most clear-cut application of this conventional paradigm. However, the appropriateness of a conventional approach has also been questioned, most strongly by Danida, noting “the inadequacy of conventional evaluation tools... [and that] evaluating efforts at political reform requires a different methodology (2000, vol.1, p.11). Similarly, UNDP has noted the problems in applying existing quantitative-based methodologies and techniques and stated that “evaluation methodologies must be rethought” (UNDP 1998a, p.27). Intriguingly, the European Commission stated the view that “human rights projects are quite different from infrastructure construction programmes, and the evaluation and selection criteria should therefore be based on a different approach” (European Commission 1995), yet in subsequent terms of reference for the MDP and ACP evaluation studies have relied on the logical framework approach that is commonly used for such infrastructure projects. Finally the Japanese International Co-operation Agency (JICA) expressed the need for the “development of new criteria and methods of

evaluation”, remarking that objectives are qualitative and difficult to quantify (JICA 1995, Chapter 3, 2.1.4).

This section questions the suitability of a conventional approach to evaluating democracy and governance assistance, based on a critique of the logical framework approach.

#### *4.1.1 Logical framework analysis: technical solution to a political problem?*

Logical framework analysis is the most commonly applied conventional approach within aid evaluation. Its relevance and appropriateness for the evaluation of political aid is strongly questioned here.

As discussed above, the logical framework approach (as used by the European Commission, Sida and DfID), as well as variations such as ‘results based management’ (CIDA) or ‘managing for results’ (USAID), provides a hierarchy of linked objectives that act as evaluation criteria, with progress towards them assessed by means of (mainly quantitative) performance indicators. It is an approach that has generated considerable controversy, especially its notions of causality and its reliance on quantitative data. Partly stemming from the general critiques of positivist approaches to evaluation (Guba and Lincoln 1989), LFA within aid evaluation has come under increasing pressure from alternative approaches that emphasise participatory and qualitative dimensions (Rebien 1996, Cracknell 2000, p.178), notably for evaluating social development projects (Marsden and Oakley 1990). A review of such debates is not possible here, and the critique offered in this research focuses specifically on the shortcomings and limitations of LFA as a means of evaluating democracy and governance assistance. Four main problems with LFA are briefly examined: its project not programme focus; its applicability to ‘hard’ not ‘soft’ data; its inward not outward orientation; its problematic emphasis on causality and quantitative indicators. It is noteworthy that criticisms include those stemming from donor agencies, notably Danida, who stated that logical framework analysis and political analysis “are not particularly compatible” (Danida 2000, vol.1, p.66), indicating the contested nature of evaluation in this field.

The logical framework approach is narrowly geared towards *project* evaluation, particularly where clear outputs can be achieved within a specific time span, i.e. ‘blueprint-type projects’ such as infrastructural projects, where ‘*hard*’ quantitative data is more readily available. It is less appropriate for evaluation of wider *programme* goals. This is especially true in an area like democracy and governance, a ‘*soft*’ area of programming in which institutional relationships and culture are the subject of reform, where time frames are hard to predict, and change is difficult to measure (CIDA 2000b).

LFA is inward-orientated, inverting evaluation towards pre-determined project objectives. In contrast, evaluation of political interventions requires an outward orientation, able to capture the dynamic political context in which such interventions are embedded. This is particularly important given that the overall context is itself a significant factor in influencing the success or otherwise of external donor interventions, for instance, the relative strength or weakness of domestic pro-democratic actors. Thus, the nature of democratisation, and of programmes intended to assist such processes, are

not appropriate to logframe-type analysis: LFA “cannot anticipate and capture the political dynamics in which local actors will make their decisions” (Danida 2000, vol.1, p.66). Indeed, the logframe approach tends to assume the idea of progress, being purposely designed to accompany this process of positive change. Two problems emerge, however. One is that democratisation is not a linear process of positive and gradual change, rather it is an irregular process following a non-linear pattern, with progressions and regressions. In the case of regressions, LFA becomes obsolete. The second problem is that the combination of inward orientation and assumed progress means that LFA is unable to countenance negative, unintended effects of DG assistance.

Other criticisms broaden this discussion and reject the ‘false notion of science’ with regard to causality and the emphasis on quantification, recalling Carthorse’ (1999) forthright critique of USAID’s ‘managing for results’ system. Given the complex nature of democratisation and the variety of factors involved, both structural and agency-related, the establishment of *plausible linkages* between donor interventions and political change may be the best that can be hoped for. Carothers was especially critical of attempts to demonstrate causality through quantitative indicators, described as a “deeply flawed undertaking” (1999, p.291), reducing complex political processes to a few numbers, for example the increased *number* of public hearings or the *number* of parliamentary bills initiated by the legislature as indicators of successful legislative strengthening activity. Such increases may be for a variety of reasons, “possibly having little to do with the aid project or with increased openness and accountability” (ibid., p.292), yet portrayed as a successful ‘result’. A similar example was found in this research. On the basis of the election of increased numbers of women councillors and mayors, success was claimed for a UNDP project aimed at promoting women’s participation in public life in Brazil, yet with little consideration of the range of factors, including the introduction of a quota system reserving seats for women, that could account for the increase (UNDP 1998b, p.45).

The overall purpose of the related ‘results-based systems’ is also recalled. They are not evaluations in the sense of learning lessons and contributing more effectively to democratic reform in the countries concerned, but orientated to demonstrate donor agency achievements in contemporary performance-fixated management culture.

In sum, logical framework analysis is rejected here, as are ‘results-based’ approaches, as inappropriate for evaluating democracy and governance programmes. They are narrow in focus, more pertinent to the limited functions of project cycle management. The (pseudo-)scientific approach, largely based on quantitative methods, is unable to cope with the dynamic of the political context in which DG activities are embedded. LFA offers an inappropriately technocratic solution to a political problem. It is concluded that evaluating democracy and governance assistance is more art than science.

## ***4.2 Impact Evaluation***

Impact evaluation is a key dimension of aid evaluation in general, and, arguably the most crucial in the political aid field, pertaining to the distinct role played by external actors in democratisation processes. From the donor experience documented above, it is also a challenging task. The most evident difficulty is that of ‘attribution’, i.e. attributing

causality and impact is often problematic. How, for instance, is it possible to differentiate the particular contribution of a specific donor from the range of factors and actors that affect the overall process of democratisation in a country? Other problematic issues regarding impact evaluation have become evident in the preceding chapters, for example, with one particular concern being the gap between micro level outcomes and macro level impact, referred to as the phenomena of the ‘missing middle’. This section looks specifically at these two main challenges.

One consequence of a lack of attention to the difficulties entailed has been overoptimistic and/or exaggerated claims concerning the effects and impact of particular programmes, with limited linkages between direct outcomes of assistance and alleged impact. This was observed in a number of reports, for example: in the claimed accomplishments of the South Africa – Canada Program on Governance (Sutherland 1999); in the lack of micro-macro linkages in the impact assessment of the PHARE/TACIS programmes of the European Union (European Commission 1997); and in the claims of positive impact in Danida’s study of support to elections (2000, vol.3). Similarly, it was observed that USAID’s impact assessment of its OTI programme in Indonesia (USAID 2000c) was more successful at tracing the linkages from project outcomes to sectoral impact, whereas attempts to evaluate impact at the macro level of overall democratisation became more tenuous. Such findings were corroborated by those of a DAC study into NGO impact assessment. This noted the lack of methodological clarity on how impact assessments were made, deducing that impact evaluation was frequently a matter of ‘personal judgement’ (OECD 1997c, chapter 7.1). In other words, assumptions of positive impact are made and reported, despite a lack of evidence to substantiate such claims.

Another factor that has been given insufficient attention is the effect of overall political context on impact at macro level, highlighted by Kapoor. He notes that governance programmes may be adversely affected by factors over which they have little control, that is, a restrictive political context, for example, lack of political will by political elite to carry out democratic reform (Kapoor 1996, p.41). Therefore no impact may be visible or measurable, despite *bone fide* attempts. A lack of attention to overall context may also affect judgements about impact in the opposite direction, with exaggerated claims about external efforts where its reliance on local support is not recognised.

### *Attribution*

The problems of attribution in conducting an impact evaluation have been raised by a number of donors, though perhaps expressed most succinctly by the evaluators of the European Commission’s PHARE/TACIS Democracy Programme. They outlined the following difficulties:

- ‘with and without’ situations or issues of counterfactualty: i.e. how to gauge what would have happened without the external support;
- intangible lines of causality;
- outputs of a specific donor indistinguishable from those of other donors;
- possible outcomes of donor efforts maybe cancelled out if not supported by other local policies and processes.

(European Commission 1997b, p.15).

Given such difficulties, the rather pessimistic conclusion of Danida's nine-volume study was that "identifying a wider impact of specific inputs is in most cases nearly impossible... [with] no firmly substantiated evidence of wider impacts, only those that can be assumed" (2000, vol.1, p.38). Although the common terms of reference for all eight country and thematic studies included an assessment of whether and how activities had promoted democracy and human rights, individual studies differed in how they addressed this aspect. The Guatemala study bluntly addressed the difficulties entailed, stating unapologetically: "very little effort [was made] to assess the impact of the specifically Danish assistance on the overall human rights situation. It is for the most part difficult to single out the Danish assistance from that of like-minded donors and, as is now widely accepted, a direct causal link between a project intervention and the broader human rights situation in any country is very difficult to determine" (vol.7, p.55). In contrast, the Nepal study concentrates initially on more direct effects at the thematic level (e.g. decentralisation, elections and media), and then makes connections to questions of wider impact. Based on this linkage to effects at the thematic level, the comments about wider impact can be made with greater confidence. For instance, concerning Danida's activities in electoral matters, "it seems likely that monitoring of elections, along with the strengthening of the Electoral Commission, has contributed somewhat to achieving the degree of fairness observed" (vol. 3. p.55). A qualification remains, however: "It is difficult to assess to what extent the assistance had decisive impact on the elections, as the Commission, of course, possessed certain capabilities and capacities before" (vol. 9, p.47). Nevertheless, perhaps this study provides a pointer as to how impact at the macro level can be ascertained, based on analysis from the thematic level (referred to here as 'meso' level), although this is not further developed by Danida.

As well as suggesting *how* wider impact can be ascertained, unwittingly the Nepal study also provides a hint as to *by whom*. It is stated that: "It is difficult for Danes and other non-Nepalese to fully understand the Nepalese environment... it is difficult to understand the dynamics of the environments in which projects are implemented and to determine to what extent certain changes in the environment can be attributed to the implemented project activities or not" (vol.9, pp.92-3). The implication, not developed by Danida in this study, could be to suggest the merits of a more participatory approach where well-informed Nepalese actors are centrally involved in analyses of impact. The only study examined in this research to implement such an approach to impact assessment was the evaluation of Norway's support to human rights NGOs, which gives primacy to the implementing NGOs rather than the aid agency / evaluators in defining results and attributing causation (NORAD 2001) (see Appendix One). Such a qualitative approach was limited to implementing agencies, however, not a wider range of actors.

The evaluators of Diakonia's activities (a Swedish NGO) in support of human rights and democracy in El Salvador also found difficulty in complying with the requirements to undertake an impact assessment of the programme "in relation to the developments in El Salvador" (Sida 1999a). Difficulties were stated as three-fold: "First of all, there is of course no general agreement on what 'developments in El Salvador' are. Second, it is impossible to identify exactly the significance of Diakonia's contribution given that many other donors also support the activities of Diakonia's POs [People's Organisations]. Third, we are speaking about a two-year program in a field where deep changes may take

a generation or more” (ibid., p.21). Notwithstanding, an attempt was made to evaluate the programme’s impact on “a consolidated democratic transition”, though the measure of difficulty experienced is indicated by each of five aspects of consolidation being dealt with extremely briefly in the space of 1 or 2 sentences (ibid., pp.24-5).

This study also raised one important, but less commonly expressed problem, that is the need to assess and analyse the pattern of political change at the national level. It was felt that this was an essential, yet lacking, pre-requisite in order to undertake an impact evaluation. Again, this confirms the significance of political context studies. The relative impact of external interventions is partly determined by the overall context in which they occur, for example, the degree of government commitment to democratic reforms and/or the strength of pro-democratic movements. Impact may appear positive because the overall setting is favourable to democratic reform, or, alternatively, negative because the obstacles are considerable. Political context studies need to include analysis of the opportunities and constraints at the partial regime or thematic level, given the potential for variations between different areas and for donor support to be inappropriately targeted.

#### *Tracing micro-macro linkages*

Donor discussion of the problems of impact evaluation has focused strongly on impact at the national or macro level. Yet the impact of external interventions can be explored at different levels. The analysis undertaken in Chapter 3 examined impact assessment at the three levels of micro, meso and macro. Findings were that a meso level assessment acted as a link to macro level impact evaluation in a significant number of instances, just over fifty percent of reports where impact assessment was undertaken. This was encouraging, given previous analysis that “most evaluations either focus on measurable project outputs or seek evidences of impact in terms of contribution of donors to the macro-level political change” (Robinson 1996, p.iii). Similarly, Schmitter and Brouwer (1999, section IV, 3) have commented on the shortcomings of macro level impact evaluations which “tend to assume rather than to prove the impact of individual projects on the entire process of democratisation, even in the absence of any clear causal relationship between the two”, citing the evaluation of the EU’s PHARE and TACIS Democracy Programme as an example of this gap or ‘missing middle’ between micro and macro level analysis. Schmitter and Brouwer specifically propose the addition of meso level analysis as an important linkage, enabling a more reliable assessment of the impact of external interventions generally. It would appear that there is some evidence of this occurring, albeit in a relatively imprecise and undeveloped manner. Meso level analysis is discussed further in Chapter 5.

#### *Summary*

It is clear that impact evaluation at country level is a challenging undertaking. Seven main issues are summarised as follows.



1. **The significance of political context** as an important determinant on programme impact, and therefore the need for a background study analysing patterns of political change at the national level.
2. **The multiplicity of actors and factors in complex political change** and the difficulties of differentiating the contribution of a single actor. There are difficulties in distinguishing the contribution of internal and external actors, as well as in separating out one donor from others.
3. **The phenomenon of the ‘missing middle’**, requiring at times an ‘act of faith’ to leap from micro level outputs to such programme objectives as ‘greater respect for human rights’.
4. **With and without scenarios and issues of counterfactualty.** Are external actors being credited for developments that would have happened anyway, without their assistance? Baseline data certainly helps in providing clear evidence of the developments that have occurred, but any attribution of these to external assistance remains hard to prove. At most, external actors may be able to show some correlation between the nature of their assistance and such developments.
5. **External – internal relationships.** In partially attributing perceived (macro level) developments to the activities of external actors, have the interrelationships between internal and external actors been sufficiently addressed? External efforts may be dependent on local support. Alternatively, countervailing forces in the particular country may undermine external actions.
6. **Time-scale.** How possible is it to evaluate the impact of projects and programmes that have only recently been completed, given that democratic change is a long-term process?
7. **Unintended impact.** External intervention involves a dynamic, inter-active process and can have unintended side effects. Does the search for positive impact ignore the possibility of such negative impact?

These are very real challenges. On the one hand their significance should not be underestimated, while on the other they should not be regarded as insuperable. From previous discussion it is clear that USAID has made most progress in addressing such challenges and developing a methodology for exploring the impact of DG programmes on democratisation processes at country level. Notwithstanding, two shortcomings of this methodology were identified above, the omission of exploring unintended negative impact and the lack of a participatory approach.

All three elements of the alternative evaluation methodology proposed in the next chapter are aimed at facilitating impact evaluation. Without overly anticipating what is to come, the arguments are as follows. First, a political context study provides evidence regarding the direction of change, both positive and negative, and enables the assessment of impact to be made in that light. Second, a meso level analysis provides the ‘missing link’ between micro and macro level analysis, enabling improved links between short-term outcomes and longer-term changes. Third, similar to the above discussion on Danida’s Nepal report, a participatory approach enables well-informed local actors to determine

the extent to which changes in the political environment are attributable to external efforts or not.

It is further contended that the above difficulties of specifying the particular contribution of external agencies become less crucial if internal forces more effectively guide democracy assistance programmes. It is also towards that end that a participatory approach to evaluation is advocated in the next chapter.

## Chapter 5.

### MEETING THE CHALLENGES: AN ALTERNATIVE METHODOLOGY

In developing an alternative approach to evaluation in this field, one that is essentially *participatory* and *qualitative* in nature, aims are two-fold: one relatively pragmatic and the other of a more political nature. The first aim is to contribute to resolving the methodological difficulties encountered in this field and therefore provide a more accurate assessment of democracy assistance activities. In this respect, evaluation can fulfil its objective of learning lessons from past efforts more effectively, and thereby help to improve future strategy, notably at country level. The second aim is to shift the notion of evaluating democracy assistance from one that is perceived as a technical exercise to one that involves a participatory political process, thereby becoming an integral part of democratisation itself. In this way, evaluation seeks to assert domestic influence and control over external assistance, as befits the nature of democratisation.

The methodology is proposed specifically as a means to conduct impact evaluation at country level, though the participatory approach has the potential for application to other types of evaluation. To a large extent, the methodology addresses the same question as the recent USAID country case-studies. That is, to what extent have DG programmes had an impact on political change in the countries concerned? Yet the participatory approach also keeps the door open to further questions about the role of external actors at different societal levels, and to issues of democratisation more generally. There are three elements to the proposed methodology:

- Political context studies: i.e. background political analysis of the particular country context and the trends towards (or regression from) democratisation.
- A three-fold linkage between micro, meso and macro levels as an appropriate means for achieving impact evaluation (Schmitter and Brouwer 1999).
- A participatory approach to evaluation, highlighting the perspectives of domestic actors, both governmental and non-governmental, on external efforts.

This latter proposition is the most central to the proposed methodology, providing the underlying theme. The three elements are addressed in turn, the first two relatively briefly, with more detailed attention to the main concept of a participatory approach.

#### *5.1 Political Context Studies*

A clear and recurrent point to emerge from the examination of donor reports was the importance of DG evaluation being embedded in the political context within which democracy assistance was undertaken. This is particularly significant given that the

(continually changing) context is itself a key determinant of the success or otherwise of donor interventions, as well as a key indicator of the opportunities for and constraints upon such assistance. The shortcomings of those studies that lacked the basis for such political analysis were noted.

Prior to an assessment of the contribution of external democracy promotion efforts in a particular country, it is necessary to gauge the political reforms that have occurred in that country over a given time period, covering at least the period of the assistance programme. Similarly, USAID's research design for its country impact studies has also emphasised the significance of examining the political context. Such political analysis would concentrate on democratisation trends over the time period, in both progressive and regressive directions, as well as the major areas of change at both regime and partial regime levels. Partial regimes are the various sectors, as well as linkages between them, that make up the political regime as a whole. Instances of partial regimes include: electoral systems, the division of power between the executive, legislature and judiciary, civil-military relations, the NGO environment and so forth (Schmitter and Brouwer 1999, section IV 4.). Political context studies would partly fulfil the role of a baseline study, generally not feasible in this field. Completion of the study would enable analysis of the relationship between the DG programme under examination and political change, that is, the extent to which the programme has focused on those area crucial to democratisation or not.

Political context studies would also identify the opportunities for and challenges to democratisation. A particularly important aspect is what Carothers has termed the 'missing link of power', that is "the power relations that underlie and in many ways determine a country's political life" (1997, p.122). It is clearly crucial to understand and analyse the various economic, social and political forces, including their relative strength, that influence and shape the prospects for democratic reform.

In contrast to USAID, the approach advocated here is that such studies would be undertaken by local academic specialists, potentially making use of democratic audit methodology (Beetham *et al.* 2001), as well as surveys of academic literature and other sources. A particular focus would be on the analysis of various social forces, both elite and mass, and the interrelationship between their deep-seated interests and democratic reforms. The only study unearthed in this research where local specialists were involved to this degree was in the Mali study on aid effectiveness (Club du Sahel *et al* 1999), not in fact a DG evaluation, but instructive in its use of Malian consultants as the evaluators.

Political context studies are invaluable not only as essential background for the assessment of the contribution of external actors to democratic reform processes, but also for future strategic planning, indicating priority areas and emerging issues. They enable evaluation of external efforts to be based on an understanding of the wider political forces and power relations at work within a society, and for future strategies to be developed in the light of such knowledge.

## ***5.2 Meso Level Analysis***

The conclusions of previous discussions of impact evaluation were to highlight the significance of meso level analysis. As proposed by Schmitter and Brouwer (1999), this can provide the ‘missing link’ between micro- and macro-level analysis. The lack of meso level analysis can lead to grandiose and unfounded claims of macro level impact, based inadequately on micro level analysis only.

The benefits of a meso perspective are its recognition that “even if donors provide assistance to a single NGO, this affects a large part or even the entire NGO community and the relationship between the community and the state/government” (Schmitter and Brouwer 1999, section IV, 4). Meso level analysis goes beyond the impact of a single project and examines changes that have occurred to a specific and limited set of organisations and institutions at the partial regime level, while avoiding the more speculative leaps of faith directly from micro to macro level. Thus, given the wide gap between analysis at micro and macro levels, the meso level can act as an important linkage (ibid.), enabling a more reliable assessment of the impact of external interventions generally. Evidence of micro-level outcomes and impact can contribute to broader pictures of meso level impact. In turn, macro level analysis can be based more credibly on impact analyses at a range of partial regimes (ibid.). Additionally, meso level analysis is significant in itself for comparative studies of donor impact at the thematic or partial regime level.

Thus, although a genuinely participatory approach defies prescription, it is proposed that tracing micro-meso-macro linkages would enable evaluators to make more plausible connections between external support and overall political change.

## ***5.3 Participatory Evaluation***

The significance of a participatory approach within aid evaluation is increasingly recognised. Indeed, Cracknell (2000, p.178) recently stated that, “The trend towards the use of participatory methods in monitoring and evaluation is undoubtedly the most significant change currently taking place in the field of evaluation”. This is reflected in this research in two ways. First, a limited number of evaluation studies involved an element of participation (see Table 3-6), with a larger incidence of interviewing ‘stakeholders’ for data collection purposes. Indeed, CIDA’s general evaluation guide states that a participatory element is essential to a ‘good’ evaluation (2000a, p.26). Second, general discussions of DG evaluation methodology have included favourable references to a participatory approach in a number of instances, notably in the OECD DAC (1997a) synthesis report on ‘participatory development and good governance’, in Danida’s (2000) methodology of ‘inter-subjective validation’, and in Canadian discussions of ‘evaluating governance assistance’ (IDRC 1999). Nevertheless, a healthy dose of scepticism is in order when considering the degree of participation envisaged. The CIDA evaluation guide (2000a), based on a ‘results-based’ approach, limits participation to mere stakeholder consultation, and similarly participation in Danida’s (2000) studies is restricted to ‘key stakeholders’, that is, project beneficiaries and implementing agents. The important distinction is recalled between the use of

participatory data collection methods, where ‘stakeholders’ remain as passive objects of the study, and a more genuinely participatory evaluation, where participants are engaged as subjects of the evaluation process. It is further recalled that only two out of the sixty evaluations examined here could be described as truly participatory, as defined by Rebien’s (1996) three threshold criteria.

The gap between donor rhetoric and practice is explicitly recognised in the DAC synthesis report, stating bluntly that, “Despite participatory rhetoric, there is very little evidence of genuine participation in evaluation” (OECD 1997a, p.23). Donor agencies may be prepared to ‘listen to stakeholders’ but “typically stop short of bringing them fully into the evaluation process”, partly due to an unwillingness to cede ‘control’ over evaluations (ibid., p.26).

This ambiguity arises partly from “participatory evaluation ha[ving] inherited some of the basic weaknesses of the participation concept itself” (Rebien 1996, p.65), notably the variable definitions and (mis)use of the term. Like participation, there are differing interpretations of what participatory evaluation means. Essentially, PE can be viewed either as an absolute or relative concept, with very different practices in consequence. As a relative concept, PE is not especially controversial or radical. It is quite compatible with conventional evaluation, with many current evaluation studies, as shown here, containing a participatory dimension that is basically tacked-on to a traditional approach. As Rebien states, however, the danger of such a watered down version is that “participatory evaluation runs the risk of being reduced to a mere catchword with no real substance” (1996, p.84). In contrast, an absolute concept of PE gives rise to issues of compatibility with conventional evaluation, as Cracknell acknowledges: “The participatory approach, in its full rigour, is clearly incompatible with the traditional approach” (2000, p.333). In this more ‘genuinely’ participatory approach, participant evaluators are subjects of the process, determining the nature of the investigation, not mere sources of information for a donor-led enquiry. In this truly participatory form, participation becomes integral to the whole evaluation, occurring throughout, from design and planning stages onwards. Such an approach is less common, of course, as evidenced here.

In promoting a participatory approach to DG evaluation, this research is advocating ‘genuine’ participatory evaluation, and views varieties of ‘limited’ and ‘controlled’ participation as less than sufficient, not representing the essential shift from participants as objects to subjects of the evaluation process. Whether donor agencies are willing to introduce a truly participatory approach to evaluation depends on the extent to which they are willing to make the requisite cultural change and share power and control more widely. Despite the concept of ‘participation’, as well as related notions of ‘partnership’, ‘ownership’ and ‘empowerment’, becoming central to current donor agency rhetoric, there is considerable scepticism concerning their willingness to actually share decision-making power, as these concepts imply.

This section will examine a participatory approach to DG evaluation in five parts. The first two will provide an introduction to participatory evaluation in general, focusing initially on its key characteristics and defining principles, and then highlighting the differences between participatory and conventional evaluation. Turning to applying a participatory methodology specifically to DG assistance, three sections explore, first,

issues of congruence and compatibility between a participatory evaluation and democracy promotion, second, the practicalities of undertaking a participatory evaluation, and third, some challenges faced in conducting a participatory DG evaluation.

### *5.3.1 Participatory Evaluation: key characteristics and defining principles.*

Participatory evaluation (PE) is not a new concept and its own history can be traced back at least twenty years to discussion by contributors in the volume edited by Fernandes and Tandon (1981) of participatory research and evaluation in India. More recently, UNDP have provided a useful definition and summary of PE. It “involves the stakeholders and beneficiaries of a programme or project in the collective examination and assessment of that programme or project. It is people-centred: project stakeholders and beneficiaries are the key actors of the evaluation process and not the mere objects of the evaluation” (1997, Part Two, p.1). Its key characteristics include:

- drawing on local resources and capacities;
- recognising the innate wisdom and knowledge of end-users;
- demonstrating that end-users are creative and knowledgeable about their environment;
- ensuring that stakeholders are part of the decision-making process;
- using facilitators who act as catalysts and who assist stakeholders in asking key questions.

(UNDP 1997, Part Two, p.3).

‘Stakeholders’ is clearly a key term and the ‘stakeholder approach’ has become increasingly common in the development parlance of both aid agencies and NGOs. Stakeholders are those “individuals, groups, organisations and institutions who directly and indirectly influence but who are also affected by the actions or development interventions of others” (Estrella 2000, p.1). Stakeholders can “include, among others, beneficiaries; project or programme staff; researchers; local and central government politicians and technical staff; funding agencies” (ibid.). The involvement of various stakeholders in the evaluation process is potentially a significant departure from the traditional evaluation approach that is donor-driven and funded, and oriented to their needs. Yet, key questions remain with regard to which stakeholders are included and what level of participation is entailed. For example, are indirect stakeholders and non-recipients (as discussed above in 2.4.3) included? Tokenism and co-option are familiar outcomes for radical ideas.

#### *Principles of Participatory Evaluation*

What is the purpose of PE? In addressing this question, it is useful to examine the underlying principles of PE. These have been outlined as four-fold by Estrella and Gaventa (1998), the principles of ‘participation’, learning, negotiation and flexibility. These are examined in turn.

*The principle of participation.* Self-evidently, the defining feature that distinguishes PE from conventional evaluation is the emphasis on participation. Yet, as discussed, a key

issue remains the variation in levels and degree of participation. Estrella and Gaventa suggest that there are two main ways by which participation in evaluation can be characterised: “by *whom* it is initiated and conducted, and *whose perspectives* are emphasised?” (1998, p.17). The first concerns whether the evaluation is externally-led, internally-led or jointly-led, while the second examines which stakeholders are emphasised – all major stakeholders, beneficiaries or marginalised groups (ibid.). An externally-led PE generally involves an external evaluation team being contracted to conduct a participatory assessment of a project or programme, seeking the perceptions of various stakeholders and beneficiaries and presenting a more ‘balanced’ and ‘objective’ account and analysis of opinions. An internally-led PE is generally initiated and undertaken by those individuals and organisations directly involved in the project or programme. Such evaluations are viewed as contributing to local capacity building. The jointly-led evaluation is somewhat of a hybrid, but generally aiming to involve a more diverse set of stakeholders and to assess efforts from a wider range of perspectives. (Ibid., pp.19-20). With regard to whose perspectives are emphasised, again it is evident that the definition of major or relevant stakeholders remains open to interpretation and could be limited to those that have a major stake (i.e. a particular interest in the project or programme). ‘Beneficiaries’ are more specifically restricted to those that have a direct involvement in the project or programme, whereas marginalised groups, defined as “the least powerful, visible, and assertive actors” remain most likely to be excluded. (Ibid, p.22).

*The principle of learning.* PE is characterised as a process of individual and collective learning for the variety of parties involved (ibid.). It is a learning process that entails reflection, analysis and action, with resultant improvements in the programme examined. A further outcome is local capacity building, with stakeholders more able to analyse their wider environment and take appropriate action.

*The principle of negotiation.* With multiple stakeholders engaged in PE, it becomes “a social process for negotiating between people’s different needs, expectations and world-views” (ibid., p.24). At one level, this relates strongly to Guba and Lincoln’s (1989) ‘fourth generation evaluation’ methodology that is fundamentally a process of negotiation between different perceptions of social realities. The value of negotiation can be recognised, however, without adopting their constructivist and subjectivist approach in which there is no social reality other than a negotiated consensus between actors. As Estrella and Gaventa point out, the negotiation process is also a “highly political exercise, which necessarily addresses issues of equity, power and social transformation” (1998, p.25). Thus, negotiation interacts with the other two principles above, with the capacity to empower participants / recipients. In Freirean manner, their active participation can result in better understanding of their environment (‘conscientisation’), leading to further action. At a more practical level, negotiation is integral to the whole evaluation process, commencing with the initial selection of criteria and indicators (ibid., p.25).

*The principle of flexibility.* There is no blueprint for carrying out PE. Rather “flexibility and experimentation are regarded as integral aspects” (ibid., p.26). The process of PE should be contextualised, i.e. adapted to project- and programme-specific contexts, with a consequent wide range of practices.



Thus, it becomes clear that participatory evaluation is valued not merely as a means towards improved data collection and better representation of stakeholder views in evaluation studies, but as an end in itself. It is argued below that participatory evaluation, in the specific DG context, is congruent with the process of democratisation itself.

### 5.3.2 Differences between Participatory Evaluation and Conventional Evaluation

Table 5-1: Differences between conventional and participatory evaluation

	Conventional evaluation	Participatory evaluation
Who	External experts	Community people, project staff, facilitator
What	Predetermined indicators of success, principally cost and production outputs	People identify their own indicators of success (which may include production outputs)
How	Focus on 'scientific objectivity'; distancing of evaluators from other participants; uniform complex procedures; delayed, limited access to results	Self-evaluation, simple methods adapted to local culture; open, immediate sharing of results through local involvement in evaluation processes
When	Usually upon completion; sometimes also mid-term	Merging of monitoring and evaluation; hence frequent small-scale evaluations
Why	Accountability, usually summative, to determine if funding continues	To empower local people to initiate, control and take corrective action

Source: Narayan-Parker 1993, p.12

Two tables are reproduced here, useful in highlighting the distinctions between conventional and participatory evaluation. Table 5-1 from Narayan-Parker presents the differences between CE and PE as an ideal-type construction. Somewhat in contrast, Table 5-2 from UNDP (1997) presents a continuum of levels of participation from low to high. The following differences between the two forms are explored, with reference to the two tables.

*Aims and purpose of evaluation.* CE is generally donor-focused and donor-driven, oriented to fulfilling accountability and management requirements regarding funds expended, and to demonstrating the overall efficiency and effectiveness of programmes. This is not incompatible with learning lessons and improving programmes, of course, but is more a matter of emphasis. However, it is the donor agency that determines the terms of reference and what use is made of the results. In a truly participatory evaluation, the stakeholders and beneficiaries decide the terms of reference, conduct the research, analyse the findings, and make recommendations (UNDP 1997, Part Two, p.3). The intent is to directly feedback results into corrective action. Wider aims can include local capacity building and even empowerment of local people, congruent with the aims of democracy and governance assistance itself.

Table 5-2: Levels of End-User Participation in Evaluation

Dimensions of evaluation/Levels of participation	Low	Medium	High
Evaluation initiator	Commissioned or obligatory evaluation typically part of programme development. Meets institutional needs. Evaluation done to, on or about people.	External evaluator invites end-users to assist in one or more evaluation task(s).	Evaluation in which end-users collaborate with external facilitator or among themselves to assess, review and critically reflect on strategies formulated for them.
Purpose	Justify or continue funding. Ensure accountability. Levels of funding or sustained support.	Gain insights into development activity from end-users' perspective. Shift focus from institutional concerns to end-user needs and interests.	Promote self-sufficiency and sustainability by linking end-users to evaluation planning cycle. Develop relevant, effective programme decision-making based on end-user views, opinions, recommendations. Increase ownership in & responsibility for success-failure of development interventions.
Questions-maker(s) ?	Agency heads, administrators, outside clientele, persons distances from evaluation site.	End-users with external evaluator at various stages of evaluation generally determined by the evaluator.	End-users, external facilitator, persons most affected by development intervention.
5 Method(s)	Established research designs, statistical analyses, reliance on various quantitative methods. Product (findings) oriented (mathematical in nature). Dominated by math whiz kids.	Qualitative methods favored but also includes quantitative methods. Values a process focussed on open-ended inquiries. Uses methods that give voice to voiceless.	Relies on highly interactive qualitative methods but does not disregard quantitative tools. "The process is the product". Inventiveness and creativity encouraged to adapt the methods to the context being evaluated.
Evaluator's versus Facilitator's Role	Evaluator takes lead in designing evaluation. Formulates questions/survey forms with no input from those evaluated. Steers overcome by setting design. Assumes objective, neutral, distant stance.	Evaluator works collaboratively at various stages with end-users. Is partner in evaluation and imparts evaluation skills. Shares lead with end-users.	Evaluator becomes more of a facilitator. Facilitator acts as catalyst, confidante, collaborator. Takes lead from end-users. Has few if any pre-determined questions.
Impact/Outcome	Reports, publications circulated in house. Findings rarely circulated among end-users. Findings loop into planning stage with little input from end-users.	Shared data-gathering but limited participation in data analysis. End-user views loop into planning stage. Increased understanding of end-user experiences.	End-user more capable of meaningful decision-making based on effective involvement in evaluation. Findings become property of end-users or community. Participation in analysis is critical.

Source: Reproduced from UNDP (1997, Part Two, pp.4-5), *Who Are the Question-makers? A Participatory Evaluation Handbook*,

*Who evaluates?* In CE, external expert consultants (invariably from Western countries) are contracted to undertake the evaluation in line with the terms of reference. They direct and control the evaluation, with limited input from local stakeholders or beneficiaries in terms of the questions asked, the types of information gathered and reflected on, or the use of findings (Estrella and Gaventa 1998, p.14, citing Rubin 1995, p.20). In general terms, this sums up the nature of the evaluation studies examined here. In PE, the overall process is determined (through negotiation) by the group of key stakeholders, which, at project level, could include local community representatives and project staff. Often an (external) facilitator is engaged, acting as a catalyst and collaborator, and whose role could include “guiding the process at critical junctures and consolidating the final report, if necessary, based on the findings of the stakeholders” (UNDP 1997, Part Two, p.3). The role of a facilitator will be addressed at other points in this chapter.

*What is evaluated?* In CE, this is partly determined by the stated objectives of the project or programme and by the terms of reference, which generally state the evaluation criteria. A project logframe will also suggest indicators and means of verification. In PE, there is more scope for criteria and indicators to be identified by the relevant stakeholders, again through a facilitated process of negotiation.

*How is the evaluation conducted?* Regarding evaluation methods, CE generally involves a desk review of project / programme documents, followed by a short field trip, with the evaluation report based on information gathered, as has been common practice in the evaluation studies examined here. Stakeholders or beneficiaries play a passive role, as sources of information. Quantitative methods are often favoured, viewed as a more reliable measure of success or otherwise. The methodology on which such methods are based is traditional, scientific inquiry, which perceives that a mixture of independent external expertise and established research design will provide objective and value-free findings. In contrast, PE recognises “the wide range of knowledge, values and concerns of stakeholders and acknowledge that these should be the litmus test to assess and then guide a project’s performance” (UNDP 1997, Part Two, p.3). Methods are interactive and qualitative, though the use of quantitative materials is possible. There is an emphasis on flexibility and innovation in developing methods for gathering and analysing information as appropriate to the particular context. The broader methodology is drawn somewhat from Guba and Lincoln’s ‘fourth generation evaluation’, with its emphasis on negotiation, though the overall theoretical perspective of phenomenology does not have to be accepted.

*What are the outcomes of the evaluation?* There is a view that access to donor-focused evaluation reports is limited, and that feedback to recipient organisations is delayed, as is consequent action (Rebien 1996, p.63). In contrast, it is felt that the results of participatory evaluations are provided as feedback to stakeholders immediately and used in a corrective manner. In addition, stakeholders should be in a stronger position to define future strategies and take effective action, especially if the capacity building function has been fulfilled.

*Who asks the questions?* Traditionally, donor agencies have initiated, directed and controlled the evaluation process, via their employment of external evaluators. Indeed, it is a process that recipient organisations and beneficiaries have at times felt threatened by,

perceiving that it was their performance that was subject to evaluation, as reiterated in the DAC (OECD 1997c) NGO impact study. PE potentially reverses the process, enabling local actors to determine the questions asked, notably concerning the relevance, effectiveness and impact of donor programmes, and to engage in information gathering, reflection and analysis in order to address those very questions. The outcome of such a process is one of learning, the strengthening of local capacity, and the potential for social action, again congruent with processes of democratisation.

The ideal-type comparison of Narayan-Parker is useful in highlighting differences between conventional and participatory evaluation. In practice, it is acknowledged that there is more of a continuum between the two perspectives, with elements of one type found in the other: both qualitative and quantitative research methods are used in both types of evaluation, though with a different emphasis; participatory approaches may use external experts as facilitators; conventional evaluations increasingly include a participatory dimension. Nevertheless, from the democracy and governance evaluation reports studied here, it is evident that they can almost all be characterised as conventional evaluations.

### *5.3.3 Participatory Evaluation and DG assistance: congruence and compatibility*

Although PE is not new, it has not been applied in the field of democracy and governance assistance, as this study has demonstrated. These findings are corroborated in Estrella and Gaventa's (1998) review of participatory monitoring and evaluation (PM&E). Their selection of 20 case examples to illustrate the types of projects and programme using PM&E did not include a single one from the political aid sector. It is asserted here, however, that a participatory approach to evaluation is both appropriate and beneficial in evaluating political aid, for the following reasons:

- PE addresses some of the challenges to evaluating DG assistance;
- The principles of PE relate closely to the principles of democracy;
- The process of PE is akin to the process of democratisation itself.

These perceived advantages are examined in more detail in turn.

### *Participatory evaluation addresses some of the challenges to evaluating DG assistance*

A number of problematic issues have been identified in the studies examined here, particularly with regard to impact assessment. These have included:

- How to assess the relevance of external actors' efforts in the overall political context of the opportunities for and obstacles to democratic development?
- How to untangle the multiplicity of actors and factors involved in democratisation trends and to single out the contribution of external assistance?
- How to assess the impact of donor activities at micro and meso levels and to construct the linkages in order to evaluate impact at the macro level of regime change?

- How to disentangle external – internal relations and to ascertain the extent to which perceived donor success is dependent on local support, or, conversely, the extent to which donor efforts are undermined by countervailing, internal forces?
- How to determine whether there have been unintended, negative effects on the democratisation process through external activities?

It is proposed here that these challenging questions can be most effectively addressed and answered through reflection and analysis by well-informed national actors, engaged in a collective process of dialogue and negotiation, partly informed by a locally-produced political context study. This is precisely the format of a participatory evaluation, in contrast to a conventional approach where external consultants are faced with answering such complex questions on the basis of a short field trip.

*The principles of participatory evaluation relate closely to the principles of democracy*

As outlined above, Estrella and Gaventa (1998, pp.17-27) have defined the four characteristic features of PE as: participation, learning, negotiation and flexibility. All four are closely associated with the principles of democracy. Participation itself is clearly central to democratic processes. The democratic principles of popular control and political equality (Beetham 1999) are realised precisely through political participation, minimally in electoral processes, as well as more substantially through a variety of democratic practices, both to influence policy-making and to hold government to account, from local level upwards to national level. Regarding the other principles of PE, there is a clear resonance with democratic values. Democracy is precisely about resolving differences and taking political decisions through negotiation, dialogue and compromise. In democratic processes and practices, the need for flexibility and the degree of learning are both held at a premium.

It is therefore asserted that a participatory approach to evaluation is especially appropriate to the DG field, given that its own defining features are closely associated with those of the very democratic processes that are being examined and assessed.

*The process of participatory evaluation is akin to the process of democratisation itself*

There is general consensus that democratisation is essentially an internal process. Especially from an actor-oriented theoretical perspective, democracy is constructed and crafted (and resisted and undermined) by various coalitions of domestic actors and interest groups, inclusive of elite groups. Frequently, it is local knowledge of rapidly changing regime circumstances and the alliances between national, pro-democratic actors that are central to processes of democratic transition and consolidation. Despite the contemporary (and historically unprecedented) phenomena of democracy promotion from outside, there is agreement that internal actors and activities are key to democratisation, and that the contribution of external actors, while not necessarily insignificant, remains limited and marginal. Notwithstanding such relative consensus, one problem identified with the democracy assistance activities of development agencies is the limited involvement of recipient organisations, both government and non-government, in the

design and implementation of projects and programmes (Brouwer 2000, pp.30-2). Such shortcomings are underlined by the application of a standard model of democratisation and menu of democracy assistance activities, with limited consideration of the particular country context (Carothers 1997).

PE provides the key input of local knowledge and analysis that is essential to an evaluation that provides a truly critical examination of external activities, in contrast to those studies that are congratulatory and even obsequious in tone (for instance, Sutherland 1999). In this way, the very process of a participatory evaluation becomes akin to the process of democratisation itself (at least from an actor-oriented perspective). The application of reflexivity in research and dialogue between actors are crucial to both processes. It is through reflection and dialogue that actors gain a critical awareness of the success and limitations of their efforts in relation to current social realities, and through which further action for change is stimulated. It is also through dialogue and negotiation that different interest positions can be resolved, with common value positions constructed across different social positions and identities. In this particular sphere, a participatory evaluation entails collective reflection by and dialogue between a range of well-informed domestic actors in order to appraise the efforts of external actors. The strengths and limitations of donor activities are examined in the context of the prospects for and constraints upon sustained democratisation in the particular country context. The anticipated outcome of such negotiation and consensus-seeking is a learning process for all. For donor agencies, their past and current efforts are subjected to critical reflection and appraisal, enabling objectives to be revisited and strategies refined, informed by internal perspectives. For participant evaluators, knowledge of processes of political change is enhanced, in turn informing and strengthening local action for democratic change. Thus, the act of evaluation becomes an act of democratisation.

#### *5.3.4 Undertaking a Participatory Evaluation of Democracy and Governance Assistance*

As discussed above, PE is widely practised in various development fields, though not as yet in the democracy and governance sector. How can PE be applied in the DG sector? What can be learnt from the experience of PE in other fields? While considerable experience of PE has been built up, especially in the past decade, it has been characterised by variation in practice, partly due to the emphasis on flexibility and adaptation to the specific context. Nonetheless, some common guidelines can be referred to, especially concerning the main stages or phases of the evaluation process.

Four key stages have been identified as follows: planning, data gathering, data analysis, reporting and dissemination of findings (Estrella and Gaventa 1998, p.28). These are taken as a framework within which to examine how a participatory evaluation of political aid could be undertaken. Added here to these four stages is a preliminary or pre-planning stage.

##### *Preliminary stage: selection of evaluators*

Estrella and Gaventa make passing reference to this prior step by stating that “the relevant stakeholders groups initially need to be identified and selected” (ibid.), but with

no further discussion or problematisation. Two crucial questions need to be posed, however: who selects the participants and on what criteria?

It is proposed that a PE of the democracy promotion efforts of external actors (either singly or preferably collectively) would be undertaken by a substantial group of key, well-informed national actors. The notion of involving ‘all, relevant stakeholders’ is a common one in PE in general. While ‘stakeholding’ is taken as a key criterion here in determining selection, the definition is broadened to include those who are affected both directly and indirectly by the interventions under examination. Indirect stakeholders are those groups and individuals who are *not* recipients of the assistance, i.e. with no *direct stake* as such, but who have a critical engagement with democratic processes in their country and thus are affected by the political interventions of external agencies. Thus not only would the evaluation group include representatives of recipient and beneficiary organisations, but also it would include representatives of categories of non-recipients (Brouwer 2000, pp.10-11). Therefore this evaluation group would incorporate a multiplicity of actors representing government and non-government organisations, recipients and non-recipients, academics and activists, as well as donor representatives.

This question of who participates is not an uncomplicated one. While a distinctive feature of PE is the inclusion of a wider range of stakeholders, “there still remains great ambiguity in defining who stakeholders are, who should be involved, and to what extent or depth they can or want to be involved” (Estrella 2000, p.10, citing Whitmore 1998). Such ambiguity will doubtless arise in applying PE to the DG field. As in PE more widely, however, it is an issue to be addressed and, with experience, to be resolved. The fact that it is a tricky issue is not a reason for discarding a participatory approach, and thus denying the significance and importance of involving internal actors in the evaluation of external interventions. Additionally, bringing together a group of local pro-democratic actors is not unique. The experiences of, for example, holding National Conferences in Francophone African countries in the early 1990s (see International IDEA 1998), and of conducting democratic assessments through ‘national democratic dialogue’, for example in Indonesia (International IDEA 2000), are valuable precedents.

### *Planning*

Having established the group of participatory evaluators, planning can commence. An initial forum can enable agenda-setting, with participants voicing their particular concerns with and interests in the programme, and negotiating the key issues for research and investigation. It is anticipated that the issue of gender and women’s equality will be amongst those raised. The main two specific tasks are establishing objectives and selecting indicators (Estrella and Gaventa 1998, pp.28-9), with a large initial workshop as the most likely forum.

Establishing objectives amounts to determining the terms of reference of the evaluation, that is establishing the parameters of the evaluation and the criteria that it will address. Criteria could include the relevance of the DG programme to the priority needs of the country, the nature of impact assessment to be undertaken, the degree of participation in the programme process, and so on.

Performance indicators can then be selected, either qualitative and/or quantitative. Guijt (2000, p.204) describes this as one of the most difficult steps in setting up a participatory evaluation, though not an impossible one. It is important in her view that indicators are not pre-determined, but identified and negotiated by primary stakeholders. This should ensure greater context-specificity. A useful question to ask is, 'What information would best tell you whether the objective had been reached?' A limited number of indicators will need to be agreed through negotiation, but with discussion likely in smaller thematic groups. Guijt further notes that achieving consensus on both objectives and indicators becomes less straightforward the greater the number of groups of stakeholders involved (2000, p.205), with implications for the DG field. It is also recalled that Kapoor (1996, p.14) discussed the advantages (and limitations) of participatory indicators, formulated through 'group workshops' lasting 3-4 days with 30-40 people. Guidance could also be sought from the 'candidate' indicators generated by donor agencies, especially the qualitative indicators forthcoming from USAID.

Another feature of this critical planning stage is a clear division of labour and allocation of tasks: what is to be analysed and by whom? Task-oriented working groups are a likely outcome, inclusive of a co-ordinator for each one. Discussion and agreement on documentation is also vital at this stage. What will be documented, the whole process or just the data and the findings? What form will it take: a written report and/or a video recording, for instance? Where will it be stored and to whom will it be disseminated? (Guijt 2000, pp.214-5).

#### *Data collection*

Research commences in earnest in this phase in working groups. A variety of tools and techniques are available and an initial question is to determine which are the ones to use for data gathering in the particular context. Guidance can be sought from the handbooks and manuals on PM&E (see Estrella and Gaventa 1998, Appendix 2). Many of the tools are drawn from participatory development methods more generally, notably PRA (Participatory Rural Appraisal), and one challenge is to select and devise methods that are appropriate to the political sphere. Another is to ensure that gender sensitive techniques are included, enabling gender analysis to be undertaken. One possible issue concerns the time constraints of some stakeholders / participants which may restrict their ability to collect data. This is not insuperable, however, with delegation possible, and such individuals should at least participate in the planning and analysis stages.

#### *Data analysis*

Data processing and analysis is undertaken by all evaluation participants, not an elite few, as is traditionally the case. Initially, participants remain in thematic working groups, with critical reflection along the lines of investigation posed at the planning stage, such as the successes and shortcomings of the programme, understanding the impacts of activities, including unintended effects. Findings on the various evaluation questions can then be presented for further discussion at a whole group workshop.

#### *Reporting and dissemination*



This final task entails clear documentation of the processes undertaken, the information gathered, the knowledge gained and lessons learned from the evaluation, stated in a manner that is clear and accessible to all possible users. Action taken as a consequence of the evaluation can be varied and by a range of actors, as discussed above. While specific lessons for external actors will be a key outcome, the participatory nature of the evaluation means that donor agencies do not hold exclusive rights over the product. Rather it becomes a tool for use by all actors engaged in efforts to promote democratic change. In this way the evaluation process can become an integral part of democratisation itself.

#### *The role of a facilitator*

An experienced facilitator of participatory processes is often invaluable and critical to the success of a participatory evaluation. The multiple roles and skills required of the facilitator are outlined in both Estrella and Gaventa (1998, p.43) and UNDP (1997, Part Four). A key role is to act as a catalyst or stimulator, encouraging participatory and sharing of ideas by all, directing and guiding the process to constructive outcomes without controlling it. One crucial skill pertains to negotiation and conflict resolution, perhaps especially important in a field where political differences could well arise. Unfortunately, Estrella and Gaventa (1998, p.43) note that little consideration has been given to this issue in the literature.

#### *5.3.5 Issues and Challenges*

The previous section offered guidance on conducting a participatory DG evaluation. This is unlikely to be a straightforward and trouble-free undertaking, however, and possible challenges and difficulties are anticipated here. As a relatively new and experimental approach, PE is itself engaging with a number of issues and concerns [Estrella and Gaventa 1998, pp.37-48; Estrella et al. (eds.) 2001]. Three such challenges are of particular pertinence to evaluation of political programmes, and are examined here.

#### *Issues of Power*

It is recognised that the practice of PE is underlain by “complex social dynamics and power relations” (Estrella and Gaventa 1998, p.37), particularly as regards processes of negotiation. A key question concerns influence and control over the PE process. There are issues of power both within and between the range of social groups engaged in the evaluation process, given unequal social relations and differences in social status. Between groups in a DG evaluation, differential power could arise between actors located at different institutional levels, for example, between officials of ‘high-level’ government agencies and civil society activists. Within social groups, social divisions can lead to differential influence, by gender for instance. Recognition and awareness of issues of power are clearly essential, with the facilitator potentially playing an important role. A key dimension concerns the planning stage and the question of who influences the selection of evaluation criteria and indicators (ibid., p.38, citing Rubin 1995, p.39). In the case of a Nepal-UK community forestry project, initial domination by powerful and vocal interest groups was redressed somewhat by opportunities being given to other groups to

articulate their needs and interests (Gaventa and Blauert 2000, p.234, citing Hamilton et al. 2000). Power, of course, is a very political concept, and democratisation entails the achievement of greater popular control over both governmental and economic power. PE itself is closely related to such processes, aiming to redress power imbalances, confirming again its suitability to evaluating political interventions.

### *Issues of Conflict*

Following Guba and Lincoln (1989), collaboration and consensus are key aspects of a participatory approach. Yet, the articulation of different views, concerns and interests can also lead to disagreement and conflict, especially where such differences are seen as embedded in existing social institutions of inequality, with the potential to paralyse the PE process (Gaventa and Blauert 2000, p.234). Guba and Lincoln's lack of attention to the potential for conflict in negotiation processes would appear to have been replicated in much participatory evaluation, with the acknowledgement that "conflict and mechanisms for resolving conflict is seldom explicitly discussed in the literature" (Estrella and Gaventa 1998, p.39). Perhaps tellingly, one exception where conflict is recognised and discussed is a case study of decentralisation and legal reform in Bolivia (ibid., citing Alcocer et al. 1997, pp.7-8). It is posited here that the potential for conflict is probably greater in evaluations of political matters. First, politics in general and democratisation in particular involves competing interests and struggles over power, likely to be reproduced to some extent in the evaluation context. Second, the involvement of multiple stakeholders from different institutional levels increases the diversity of views and interests. Seeking resolution to different stakeholder interests is not an uncomplicated or undemanding task. Frequently, there are real, existing differences that emerge out of social inequalities. They are not going to be simply negotiated away, as imagined in Guba and Lincoln's (1989) approach. Nevertheless, the process of participatory evaluation does provide a forum for democratic dialogue that is precisely aimed at addressing such differences, with the potential for compromise and agreement on how to tackle key issues within a democratic framework.

### *Issues of inclusion and exclusion*

The questions of 'who participates' and 'who selects the participants' have been raised already. Such concerns are pertinent to all participatory evaluations, though with particular relevance in the DG sphere, given that democratisation is about *inclusive* political participation. Additionally, these issues are closely related to those of power, given that power relations can determine who participates and under what circumstances (Campilan 2000, p.195). The fact that certain parties are allowed to participate, and others not, is a strong indicator of who has control over a process of PE (ibid.). As regards the range of possible participants, Campilan distinguishes three broad possibilities:

- the local people/community;
- a partnership between project beneficiaries and the usual external monitoring and evaluation specialists/experts;

- a wider group of stakeholders who are directly and indirectly involved in or affected by development interventions;

(ibid.).

The first is clearly most applicable to a participatory evaluation of a local, community-level project. In relation to evaluating DG assistance, the second category would be viable and practicable. However, this research favours the final option. Its value is seen in drawing together a wider variety of well-informed perspectives from both government and non-government sources, not restricted to those who have a direct (financial) stake in the programme. This has greater potential to provide a more balanced and wide-ranging examination and evaluation of donor activities, as well as developing a collaborative strategy for future DG support.

However, this does not provide complete answers to the tough questions of who participates and how participants are selected. Two further linked proposals are made. One is to veer towards inclusivity, though with the recognition that there are practical limits to numbers. Importantly, actors *outside* the programme itself, without a direct stake in it, are important to include. The second is for selection to be based on democratic criteria, that is, on evidence of a positive engagement in the democratic process as a whole, with a manifest commitment to frank, open and transparent discussion, with respect for other viewpoints. By such criteria, those groups unlikely to make a positive contribution would exclude themselves.

### *Concluding comments*

The advantages of a participatory approach to evaluating democracy and governance assistance are manifest. An extended role is given to well-informed local actors, with the evaluation benefiting from their expert knowledge of their societies. Donors are provided with findings that can potentially improve the relevance, effectiveness, impact and sustainability of their democracy and governance support. A basis for more genuine 'ownership' and 'partnership' has been established. Such findings also feedback to local organisations, at governmental and non-governmental levels, through their representatives who were engaged in the evaluation study. Thereby the learning process is transmitted into the wider society, contributing to those pressures and movements for democratic change, albeit in a minor way, upon which not only the success of external efforts are dependent, but also the prospects for sustained democratisation as a whole. In this way, it is claimed that the process of evaluation becomes an integral part of democratisation itself.

## Chapter 6.

### CONCLUSION: LESSONS LEARNED AND THE WAY FORWARD

This report commenced with a statement by Mark Robinson from a previous ESCOR study that “donors lack a systematic approach to evaluation [of democracy assistance and political aid] and there is no generally accepted methodology” (1996: ii). Five years later, what developments can be reported?

One key change has been in the number of evaluations in this field undertaken by donor agencies, far surpassing our expectations. The survey unearthed over one hundred evaluation reports and other related studies. Yet, the studies are characterised by variability and diversity, with no evidence of any consensus on how to evaluate activities in this area. In this respect, Robinson’s statement retains its validity. Nonetheless, the quantity of studies undertaken in the last five years does indicate that greater consideration has been given to addressing the challenges of conducting evaluations in this field, notably by those donors highlighted in Chapter 2.

Both USAID and CIDA have devoted considerable attention to DG evaluation issues for a number of years. Unfortunately, both have been forced into the agency-wide straitjacket of a result-based approach. Such systems have little to do with genuine evaluation objectives, focused on compliance with annual performance management targets. USAID’s version, entitled ‘managing for results’, has been subjected to a formidable critique by Thomas Carothers, notably of its quantitative basis, ‘the inadequacy of numbers’, and the distorting effect on new programmes, designed to produce quantifiable results, in which “The evaluation tail begins to wag the program dog” (1999, p.294). It would appear that such criticism has had an impact, given USAID’s recent shift towards a more qualitative approach in two respects: the development of qualitative indicators for measuring progress towards DG objectives for MFR purposes (as yet unpublished); and the qualitative methodology for country impact studies (USAID 2000k), with publication of the three pilot case-studies expected in 2002. The latter is the most considered attempt to explore the difficult question of impact evaluation at the macro level, that is the impact that DG programmes have had on democratisation in the countries concerned. In the Canadian context, some pertinent discussions have also occurred, aside from the ‘results-based management’ straitjacket. In particular the contributions from Ilan Kapoor (1996 and in IDRC 1999) have highlighted the benefits of a participatory approach to governance evaluation, including the development of participatory indicators.

A number of studies have commented on the problems of applying conventional evaluation methodology to DG assistance, most notably Danida (2000) and UNDP (1998a). Danida’s nine-volume report provides the most comprehensive, single

evaluation study, one that is particularly critical of the appropriateness of a traditional, quantitative approach. Yet, it is unable to offer an alternative methodology, with insufficient elaboration of its own methodological approach, described as ‘inter-subjective validation’, and concluding with mere encouragement of “other methods” of evaluation (2000, vol.1, p.68). Perhaps this sums up the current dilemma. Awareness is expressed of the difficulties of evaluation in this field and of the shortcomings of a conventional approach, yet no agency has developed an alternative methodology in any detail.

This research sought to stress the inappropriate nature of logframe-related evaluation in the political aid field. Aside from the inadequate nature of quantitative indicators and the ‘softness’ of data in a sphere in which institutional relationships and culture are the subject of reform, one key criticism has concerned the inability of a logframe approach to impart or to take into account the political context in which DG projects and programmes are embedded. LFA is inward-orientated, inverting evaluation towards achievement (or not) of pre-determined project objectives, whereas evaluation of DG assistance requires an outward orientation, able to capture the dynamic political context of which it is part. It is argued that the application of LFA and other quantitative-based methods in this field demonstrate the inadequacies of a technical approach to evaluation of political processes. One related outcome is the dearth of critical analysis of the role of external actors, with evaluation solely of donor agencies’ own objectives, compounded by relationships in which consultant evaluators and beneficiaries of assistance are beholden to the funding agencies.

The challenges in evaluating DG assistance remain, however, “more pronounced than in the assessment of other types of development assistance” (Danida 2000, vol.1, p.10). In ‘re-thinking’ evaluation methodology in this field, as advised by UNDP (1998a, p.27), this research turned to the alternative approach of participatory evaluation. It is argued here that a participatory approach to DG evaluation is both appropriate and beneficial. The suitability of a participatory approach is claimed on the basis of congruence and compatibility, sharing many of the characteristics and principles of democratisation itself. Further, the potential for a synergistic relationship between participatory evaluation and democratisation is asserted. Participatory evaluation entails democratic dialogue amongst a range of well-informed domestic actors in order to appraise the efforts of external actors. The anticipated outcome is a learning process for all. For donor agencies, their past and current efforts are subject to constructive criticism, enabling objectives to be revisited and strategies refined, better informed by internal perspectives. For participant evaluators, knowledge of processes of domestic political change is further enhanced, building local capacity and in turn strengthening local action for democratic reform. Thus, the act of evaluation becomes an act of democratisation.

Such benefits require a *genuinely* participatory approach, however, one that was virtually absent, with two possible exceptions, from the large number of evaluation reports examined. Indeed, donor agencies’ ambivalence to participatory evaluation was noted by the DAC’s synthesis report on ‘participatory development and good governance’: “Despite participatory rhetoric, there is very little evidence of genuine participation in evaluation” (OECD 1997a, p.23). Findings here confirm such an analysis, with evidence of participatory techniques used in a limited way only, consulting and ‘listening to’

stakeholders, for instance, yet with a reluctance to bring stakeholder participants fully into the evaluation process with greater powers of initiative and control. The clear danger is of participatory evaluation being interpreted in a nominal and instrumental manner, like that of the concept of participation itself. As Rebien has stated, “participatory evaluation runs the risk of being reduced to a mere catchword with no real substance” (1996, p.84). A key question remains the extent to which higher level institutions - bilateral and multilateral donor agencies and recipient government agencies - are willing to make the requisite cultural changes, that is a wider sharing of power and control. Despite concepts like ‘participation’, ‘partnership’, ‘ownership’ and ‘empowerment’ all becoming central to donor agency rhetoric, there is considerable scepticism concerning their willingness to actually share decision-making power.

Nonetheless, if such concepts are to be realised in practice, what better starting point than in the sphere of democracy promotion, given the vulnerability of Western government agencies to two-related criticisms? The first is that their actions involve an external imposition on internal processes of democratisation, undemocratic in itself. Indeed, the critique can be that Western agencies are intent to impose an external model of democracy, perceived as a limited, procedural form, largely devoid of a social reform content (Gills et al. 1993, p.5; Abrahamsen 2000). Undemocratic practices are clearly contradictory, but, as importantly, debates about democracy should not be foreclosed. One aim of a participatory approach to evaluation is to create a mechanism through which local actors are able to discuss what democracy means in their particular socio-economic context. Concepts of democracy advocated may themselves be more substantive and participatory, differing from the prevailing donor orthodoxy. The second criticism concerns north-south relations and the asymmetry of power between donor and recipient organisations. The lack of inclusion of recipient perspectives in the design and implementation of DG programmes indicates at times a ‘donors know best’ attitude (Brouwer 2000, pp.30-2), perhaps stemming from a perception by donors of their countries as ‘established democracies’, relatively oblivious of the substantial democratic deficits within their own polities. Evaluation is about learning lessons from past efforts, and participatory evaluation entails listening to and learning from a range of well-informed local actors on the contribution of external agencies. In turn, such lessons can be integrated into revised strategies and programmes that become more effectively influenced and guided by internal forces. In this manner, democracy promotion can learn from, and subordinate itself to, processes of internal ‘authorship’.

# Appendix One

## Further donor evaluation studies

Evaluation reports and related information from a further six bilateral agencies and two multilateral agency reports are examined here. In alphabetical order, the bilateral agencies are those of Australia, Ireland, Japan, The Netherlands, Norway and the United Kingdom. The multilateral reports are a World Bank review of civil service reform and a joint evaluation of aid effectiveness in Mali, jointly sponsored by the Club du Sahel, the Development Co-operation Directorate (DCD) of the OECD and UNDP. The bilateral agency reports are mainly project evaluations, representing their evaluation activity to date in the democracy, human rights and good governance field.

### *Australia: AusAid*

Governance has recently been highlighted as an area of key importance within Australian development co-operation. It was the first sector listed by the Minister in his Budget Statement for 2000-2001, and is seen as a fundamental building block for sustainable development. Governance currently accounts for 15% of the Australian aid budget, making it the third largest area of intervention (AusAID 2001, p.2). Given that substantial sums of money are now flowing into the governance sector, AusAID has drawn up plans for what appears to be its first review of the sector focused on a small number of projects (p.2). The review will apparently seek to ask a number of questions about the way Australia carries out its governance programme. The following areas are identified as particularly problematic:

- governance projects being less well prepared than other projects, as measured by the Performance Information and Assessment Section's 1999 "Rapid Review of Project Quality at Entry";
- "linkages between the governance improvement and the central issues of poverty reduction, human rights issues and gender [being] weak in most projects";
- "the use of transparent and participatory planning mechanisms such as the log frame approach [is] uncommon";

(p.6).

It is also the intention to look at impact and issues of sustainability (pp.13-14), especially in the light of earlier findings that governance projects tend to be driven by what the agency can provide rather than what beneficiaries require (ibid. p6).

### ***Ireland Aid***

To date, evaluations of Irish assistance to democracy and human rights have been limited to one project evaluation in Tanzania and a slightly wider review of four democracy-related projects in Mozambique, co-financed with and implemented by UNDP.

#### ***Report on the Final Part of the Programme on Constitutionalism for the Tanzanian Judiciary (1999)***

This can be described as an ex-post project evaluation, focusing on immediate project outcomes, providing generally positive findings in what is a very cursory report. There is no discussion of methodology or methods and no lessons on evaluating projects in this area can be inferred from this example.

#### ***Human rights and democratisation in Mozambique: A three year strategy (Ireland Aid 2000)***

This study reviews “those projects concerning human rights and democratisation supported by Ireland Aid [in Mozambique] since 1997”, that is four projects in all, two Ireland Aid projects (community radio and prisons) and two UNDP-managed projects (elections and legislative strengthening), to which Ireland Aid contributed funds. It provides an ex-post evaluation of immediate outcomes, though oriented to developing future strategy.

The implementation and outcomes of these projects are critically assessed. Findings are generally positive, with the exception of the UNDP legislative strengthening project. Ireland Aid’s own projects are described as “generally successful” and “much needed” (pp.6-7). Useful comments on current projects and suggestions for expanding Ireland Aid’s involvement into other areas are made, (for example, rights of children and widows), linking the Ireland Aid strategy into the general human rights and democracy situation in Mozambique. The study is brief and informative, but limited by Ireland Aid’s low levels of activity. Sectoral interventions appear to be limited to a support programme for three community radio stations and one prison, with no indication of what proportion of the country, or prisoners, benefits from these activities. This lack of perspective means that it is hard to gauge the reach of activities, and the extent of the issues that the agency is trying to tackle.

### ***Japan: Ministry of Foreign Affairs and Japanese International Co-operation Agency***

No evaluation studies have yet been prepared on Japanese democracy and good governance assistance because JICA’s programme is relatively small and has only been evaluated informally so far (e.g. through discussions on the final day of a training workshop) (JICA 2001 – *personal correspondence*). There are however a number of related documents, including a report by the Aid Study Committee on *Participatory Development and Good Governance* (JICA 1995), and Ministry of Foreign Affairs website documents.



The Aid Study Committee report on participatory development and good governance provides an interesting (and somewhat unusual) exploration of the linkages between these two concepts. Presumably this related to the ongoing discussions at that time within the DAC on this specific topic, although the DAC reports (OECD 1997a & 1997b) did not address the actual relationship between participatory development and good governance. A useful contribution is made to debates on the concept of 'good governance' with the distinction between 'normative' and 'functional' aspects. The former concerns 'the ideal orientation of a state', that is, the extent to which it is democratically oriented in terms of its legitimacy, accountability, securing of human rights, civilian control, local autonomy and devolution of power. Whereas the latter relates to 'the ideal functioning of a government', that is, the extent to which it has the requisite political and administrative structures, mechanisms and capacity to function effectively and efficiently (JICA 1995). Two points pertinent to this research are as follows. One expresses caution regarding the role of external actors in democratic processes, stating that aid in this field should be based on the 'beneficiaries' initiatives' (ibid., Chapter 2, section 2.2.2). This arises from an intent to respect differences in culture and values regarding political democracy and not to impose Western democracy and human rights as universal criteria for judgement. The historical, social and cultural diversity of developing countries is recognised and, thus, not only will the form of democracy adopted vary, but also a country's process of democratisation will be realised by its people at its own pace and in its own manner (ibid.). The other relevant comment concerns evaluation as a key future issue. The need for the "development of new criteria and methods of evaluation" is stated, remarking that objectives are qualitative and difficult to quantify (ibid., Chapter 3, 2.1.4). In this context, the monitoring and evaluation of results by local NGOs is mentioned (ibid.), suggestive of a participatory process, which corresponds with the above emphasis on 'beneficiary initiatives'. The extent to which such stated principles and intentions have been put into practice in subsequent years is another matter, of course.

Documents from the Ministry of Foreign Affairs current website provide some additional information on democratisation policy within Japanese ODA. The document entitled *Partnership for Democratic Development* reaffirms the commitment to support democratisation. Two points stand out. One is the strong linkage to market economics, with democratisation and the transition to a market economy invariably linked and discussed as a single objective. The other is the emphasis on 'partnership', with a declared intent to support self-help efforts, based on consultation and agreement with the developing country itself. Another document entitled *JICA Projects for Support of Democratisation* presents a rather different reality. Actual aid to democratic development appears fairly limited in terms of activities and volume, concentrated in technical assistance, inclusive of training courses on 'Japan's experiences and practices' and the provision of Japanese technical expertise. This contrasts with notions of 'partnership' and 'beneficiary initiatives' discussed above.

A major *Report on the Reform of Japan's ODA Evaluation System* was published in March 2000. Significant planned reforms are outlined, though democracy and governance assistance is not specifically mentioned. Three levels of evaluation are indicated: by policy, by programme and by project. Additionally, an intent is stated to undertake evaluations on "fields and programmes of growing importance" yet insufficiently

evaluated (JICA 2000, section 3:2). Whether this includes democracy and governance programmes is not specified. An Evaluation Manual is to be developed.

It would appear that both programming and evaluation of Japanese democracy and governance assistance remains at a relatively early stage, but with further developments anticipated in the near future.

### *The Netherlands*

It is stated that human rights has been an element of the Dutch development co-operation programme since the 1970s, and continues to be a high priority ((Helmich and Borghese 1998, Annex I). Since the end of the Cold War, more emphasis has also been placed on issues of democratisation and good governance, in common with other donors, with the distinction between projects in these areas becoming blurred (ibid.). The human rights emphasis is evident by the Dutch government undertaking responsibility for the human rights theme within the PDGG Steering Committee of the DAC 'Expert Group on Aid Evaluation' (OECD 1997a). Much of the activity in this area is carried out by the four development NGOs, (known as co-financing organisations), to whom much Dutch aid is channelled. They appear to have conducted their own evaluations ((Helmich and Borghese 1998, p.157), but with no detailed information available.

### *Review of the Netherlands development programme for the Palestinian Territories (1999)*

Of recent evaluations undertaken of Dutch development co-operation programmes, only one contains a significant governance section, the review of support to the Palestinian Territories (Dutch Ministry of Foreign Affairs 1999). This is a country programme evaluation with the objective "to assess the effectiveness of the Dutch development programme" in support of "balanced socio-economic development and a functioning civil society" (p.vi & p.vii). Methods comprise the common procedure of a desk review of all project documentation, followed by a study visit to those sectors and projects selected for more detailed evaluation. Additionally, two separate chapters in the report provide contextual information on political developments in the Palestinian Territories and the role of Dutch co-operation. Participation in the production of the report appears minimal, though a meeting was held with the Palestinian Authorities to discuss the draft findings.

Under governance, assistance appears to have been mainly budget support for the Palestinian Authority (PA) and some NGOs. There is no real attempt to evaluate the impact of the governance programme as a whole, or to measure change in this field. Rather, a series of discrete project evaluations is undertaken, focusing on implementation and results. Evaluation criteria are the fairly standard set of policy relevant, effectiveness, efficiency and sustainability, though conspicuously not impact. Findings for selected governance projects are generally positive and the emphasis on institution building, both government agencies and NGOs, is generally deemed successful. The reliability of such findings is somewhat undermined, however, by the following curious conclusion:

“Under governance, budget support for the Ministries of the PA, the Police Force, and the Palestinian Local Authorities, has contributed to the sustainability of essential *civil society* institutions. Support for human rights NGOs has been effective in sensitising the PA on human rights matters” (p.xiv, emphasis added).

### ***Norway: Norwegian Agency for International Development (NORAD) and Ministry of Foreign Affairs***

Two evaluation reports are examined. The first is a review of the linkage of Norwegian aid to issues of human rights and democracy in the 1990s, inclusive of both political conditionality and positive assistance. The second is a more specific evaluation of support to human rights NGOs in Pakistan.

#### *Aid as a tool for the promotion of human rights and democracy (1997)*

In late 1997, as part of their series of evaluation reports, the Norwegian Ministry of Foreign Affairs published a study entitled *Aid as a tool for the promotion of human rights and democracy: What can Norway do?*, authored by Hilde Selbervik of the Chr. Michelsen Institute. This is a valuable study of Norwegian aid policy and practice in this area in the 1990s, covering both political conditionality as well as the positive measures of democracy and human rights assistance. Despite being termed an evaluation report, however, it is more a general review of policy evolution and implementation, rather than a criteria-based evaluation. The report’s coverage is broad, however, a programme review rather than more narrowly project-based. It does include ‘experiences’ from three country studies (Zimbabwe, Zambia and Tanzania), though these are not comprehensive country evaluations, described as ‘a cursory review’, and do not attempt an impact assessment. Methods include: a review of official documents, including policy statements and strategy documents of other donors; interviews with aid officials in both donor and case study countries; some interviews with the representatives of beneficiary organisations; secondary academic literature.

The report’s initial chapters follow a sequential order. After an introduction, chapter two examines the evolution of Norwegian policy in this field, focusing on the definitions given to the three key concepts of human rights, democracy and good governance. Chapter three explores the main strategy options of conditionality, democracy and human rights assistance, and dialogue, described as ‘pressure, support, and persuasion’. Chapter four is somewhat of a diversion into international relations theory, with an analysis of ‘linkage diplomacy’ and how states influence each other, inclusive of factors that influence effectiveness. The methodology that emerges from such discussions is posited as useful for a second phase evaluation study, but one that would be more pertinent to a follow-up analysis of conditionality rather than democracy assistance. In any event, no follow-up would appear to have been undertaken. Chapter five expands to compare the policies in this field of the ‘like-minded’ donors, whilst the final chapter examines Norwegian support to democratic development in the three countries as well as offering some concluding findings and recommendations. Findings are as follows:

1. The prioritisation of democracy and human rights promotion at policy-level by the Norwegian government is not matched by implementation, neither in the volume of support or in terms of strategy. Much assistance is *ad hoc*, reactive (i.e. application driven) and lacking in coherence.
2. To improve policy implementation, the following are recommended:
  - General guidelines for support (i.e. a strategy document);
  - Country assessments of the human rights and democracy situation (i.e. analysis of the political context in individual countries);
  - Country-specific strategies, identifying appropriate entry points and project areas.
3. At the donor-level, there is a need for:
  - Increased Norwegian capacity and competence in this area, in the Ministry of Foreign Affairs, in NORAD and in the Embassies;
  - Increased donor co-ordination
4. Positive measures are preferable to conditionality, which does not work well, though should not be ruled out.
5. The lack of evaluations in this field is noted, though reference is made to ‘encouraging’ assessments of positive measures.

*Evaluation of Norwegian Support to Human Rights NGOs in Pakistan (2001)*

A recent evaluation study was more specific in its orientation. It examined the work of seven Pakistani human rights NGOs to whom NORAD provides core funding. It is undertaken partly to provide background analysis for a new ‘Position Paper on Human Rights and Democratisation’ to be produced by the Norwegian Embassy in Pakistan (NORAD 2001, p.10). The evaluation also briefly reviews the human rights programmes of other donors in Pakistan. Methods include a document review and an informal survey of organisations, but with interviews as the main source of information, inclusive of human rights NGOs, Pakistani authorities and donor agencies. NORAD is not an implementing agency in this instance, but perceives its role as a facilitator by providing financial support to the partner institutions, themselves relatively autonomous in their activities. Moreover, often these NGOs “do not implement programs themselves but work through a large network of community based organisations at the grass-roots (NORAD 2001, p.2). Therefore the aim of the evaluation is not so much to investigate the inputs and outputs of specific NGO projects in a logframe-type evaluation, but to examine the overall achievements of the organisations. Findings are generally positive concerning the organisations supported by NORAD and the relevance of their activities. It is felt that:

“The organisations supported by NORAD are some of the most important human rights agents in Pakistan. ...They conduct serious human rights work and produce often impressive results in terms of output” (p.2).

Moreover:

“the seven human rights organisations examined in this report have made significant contributions to human rights and democratisation processes in Pakistan, notably (but not exclusively) in the following areas:

- “Human rights issues in Pakistan and internationally;
- “Legal development and reform in order to enforce human rights;
- “Institutional development for the protection of human rights, including national, regional and international networking;
- “Human rights litigation and defense for victims of human rights violations;
- “Contributions to the development of more democratic political institutions, and increased interaction between civic institutions and political authorities”.

(p.4)

The study concludes that “the overall effects of the work of these seven organisations have contributed significantly to the promotion of human rights in Pakistan” (p.27). While there is no attempt to identify and attribute NORAD’s specific contribution, the authors do find that the scale of the organisations’ human rights work has expanded since gaining NORAD support (p.24).

Regarding other donors, it is found that “the Swiss and CIDA have fairly well-developed policies and mechanisms for human rights support in Pakistan, while AusAid and the European Union (including the French Embassy) lack a comprehensive policy (p.34). Recommendations include improved donor co-ordination and more dialogue on human rights issues between donors and government.

The authors comment on the role of impact assessment in their evaluation, noting that impact assessment “implies that we are able to identify the change in people’s life [sic] that resulted from the project.” They continue that such detailed work “was not [within] the scope of the present assignment” (p.25). Usefully, however, the authors do solicit information from partner NGOs on their ‘achievements’ (which, the authors emphasise are the results of combined donor support – p.25). ‘Achievements’ include:

- “Contributed to the establishment of a permanent National Commission on the Status of Women”
- “Liberation of thousands of bonded Haris and prisoners”
- “provision of legal aid, counselling and shelter to vulnerable groups”
- “contributed to shifting the human rights discourse from purely legal rights to practices of cultures and traditions”

(pp.26-27).

These qualitative findings are very broad in their reach, but do show evidence that tangible impact has been achieved. This approach also has the advantage of having an element of self-evaluation, which means that conclusions about inputs and outcomes/ impacts are drawn by those who have the closest knowledge about what they are trying to achieve and whether they have been successful. Of course, this approach does not necessarily have the rigour of an external evaluation, but may produce more valid indicators. Modification by adding an element of peer-review in which the various NGOs

would also comment on each others' work could be trialled as a means of strengthening this approach in future.

On the question of on-going monitoring, the authors note that the Norwegian Embassy takes a very hands off approach, relying on self-reporting under the concept of "recipient responsibility" (p.25). Attempts to encourage NGOs to use the LFA reporting format are noted to have largely failed: "the partner NGOs generally, do not seem to find the LFA a useful reporting instrument" (p.25)

The approach adopted in this report appears to be generally appropriate to the task. The authors' attempts to put Norwegian support into a political context is informative as is the qualitative approach to impact assessment, which gives primacy to the implementing NGOs rather than to the aid agency in defining results and attributing causation.

### ***UK Department for International Development (DfID)***

British development policy was expanded in the early 1990s to incorporate an explicitly political dimension with the introduction of support for 'good government'. Key aspects of good government included sound economic policies, the competence of government and respect for human rights and the rule of law (Crawford 2001, pp.59-61). With the change of government in May 1997, this uniquely British phrase has been replaced with the more ubiquitous 'good governance'.

DfID undertakes an evaluation programme in which the implementation and impact of selected projects in different fields are examined ex-post in order to generate lessons for application to current and future projects. In the good governance field, three project evaluations were conducted during the 1990s on civil service reform and police assistance, plus a wider sectoral evaluation on police projects. These two themes are not altogether surprising, given that they represent areas in which British 'good government' aid has been concentrated, particularly support to police forces (Crawford 2000a, p.46). The reports are examined below in chronological order.

DfID evaluation studies have similar terms of reference and generally follow a common and well structured format. They are ex-post, generally carried out soon after project completion. They are impact evaluations, assessing outcomes against stated objectives, though clearly only immediate impact is discernible. Following the re-orientation of development policy under DfID, all evaluations are required to assess the extent to which a project has explicitly addressed poverty and gender concerns. Cost-benefit analysis was previously routinely included, but less so in recent years: the increasing numbers of social and political development projects lend themselves less readily to economic quantification (Cracknell 2000, p.142). Evaluation teams are frequently a mix of DfID staff and independent consultants, and undertake their task in the following stages:

- initial desk study;
- consultation with key individuals and organisations;
- fieldwork to collect data and interview those above;

- draft report sent to key individuals and organisations for comment;
- draft report submitted, previously to ODA Projects & Evaluation Committee “to discuss and agree the main conclusions and lessons learned” (DfID 1993, Preface), now to DfID’s Director General (Resources) “to note the main conclusions and lessons to be learned” (DfID 1999, Preface).

The overall success of a project is rated on a standard five-point scale from ‘Highly Successful’ (A+) to ‘Unsuccessful’ (D).

Participation is generally limited to fieldwork interviews with key individuals, mainly implementers and beneficiaries of the project, though terms of reference may ask whether the views of government were incorporated in project design and implementation (1993 DfID, p.43). Under methodology in the terms of reference, one study examined here does state that, “A participatory approach will be taken”(DfID 1998, p.45). Participation in this instance remained limited, however, to interviews with ‘key stakeholders’. The objectives of another study undertake to assess “whether or not there is a public perception that the behaviour of the police force as a whole has changed as a result of the assistance programme” (DfID 1999, p.35). Consequently, interviews are broadened to include ‘other interested parties’, including human rights groups. This introduces a potentially useful participatory element, though it is noted that a direct answer is not provided to this question.

#### *Evaluation of ODA Project in Support of Ghana Civil Service Reform Programme (1993)*

This earlier study can be included under the ‘good governance’ heading, but this evaluation is more strictly associated with economic reform rather than political reform. It is an evaluation of an ODA project (1987-92) that was part of the wider Civil Service Reform Programme (CRSP) in Ghana, to which support was mainly provided by the World Bank. The CSRP itself was a component of the structural adjustment programme (Economic Recovery Programme) agreed by the Ghanaian Government with World Bank and IMF. Indeed, one could say that such a civil service reform programme could be implemented irrespective of the political regime, though in fact a democratic transition in Ghana did occur at this time, with democratic elections in 1992. Findings were that the project was only “partially successful”(DfID 1993, p.ii), due to many intended reforms remaining unimplemented. The problem of externally-led reforms without full domestic commitment was highlighted as a cause of slow implementation: “The fact that reforms were essentially externally driven, and lacked the full commitment of some senior members of the Ghana Civil Service, is an important factor in explaining the level of success achieved” (p.iii).

#### *Uganda Police Project Evaluation (1998)*

Undertaken from 1991 to 1998, the Uganda Police Project entailed a variety of activities, inclusive of institutional strengthening, training, provision of equipment. Although rated as ‘partially successful’, assessment of impact was hampered by the indistinct nature of

project objectives and a lack of clearly measurable indicators of achievement (DfID 1998, p.4). The difficulty was not a deficiency of indicators, (there were 30 in the logframe for phase I alone), but that means of verification were not available. The data required was described as “daunting enough for a Western police force with detailed statistical procedures” and consequently the indicators were “never really assessable” (p.37).

The objectives of this good government project are rather surprising, inclusive of economic growth but not human rights. For phase II of the project from 1993 onwards, the objective was stated as “to contribute to the development of law and order..., thus creating an enabling environment for stability and economic growth” (p.8), while “the promotion of human rights was not included among the objectives” (p.28). Despite discovering that the project’s impact on economic matters was “too tenuous to permit firm judgements” (p.4), this objective itself is not questioned. On the contrary, the assumption that improving the quality of policing will improve the prospects for economic development is regarded here as “self-evident” (p.12), thereby conveniently removing any need to explain and support a debatable statement. The omission of a human rights dimension to police aid is criticised in the study on two counts, however. One is simply to re-state that, “Aid support for the police *involves* issues of human rights. These need to be recognised in project design and documentation” (p.29, emphasis added). The other notes the potentially paradoxical nature of police aid, purporting to promote good government: there was a “failure to consider... the potential risk to the reputation of the UK if the UPF [Ugandan Police Force] were to be found guilty of serious abuses of human rights while being supported by British aid” (p.28). As now required of all evaluations, the impact on women is assessed, with the interesting finding that, “A start has been made in creating a sense of gender awareness within the UPF...But all too often gender issues are still seen as exclusively as women’s issues” (p.2). The need to “mainstream gender analysis and planning in project design” is stated as a lesson learned (p.4).

*Evaluation of the Indonesia National Police Management Training Project, 1983-96 (1999)*

Controversially, Indonesia became the third largest recipient of British aid in the mid-1990s (World Development Movement 1997, p.6), resulting in numerous parliamentary questions and a National Audit Office enquiry (National Audit Office 1996). Unease concerned the repressive and authoritarian nature of President Suharto’s New Order regime, and the motivations of the British government were questioned, with suggestions that such high levels of aid were linked to the promotion of British commercial interests, including arms deals (World Development Movement, 1995, pp.51-2). Probably the single most contentious element of aid to Indonesia was this long-running project of assistance to the Indonesia National Police (INP), given the close association between the INP and the military and the police’s involvement in suppression of the civil and political rights of opponents of President Suharto, including pro-democracy actors (Human Rights Watch 1994 & 1995, both p.158).



The project initially preceded the introduction by the UK government of the aid policy objective of good government, but continued under that classification in the 1990s. The management training focus of the project is evident from its title. Project activities entailed capacity building and training of trainers for INP training institutions, inclusive of study visits to the UK. In phase II from 1990 onwards, the main emphasis was internal organisational development with the creation of an Internal Management Consultancy Unit within the INP. This aimed to establish an in-house resource for consultancy on management skills and organisational development. A wider objective in phase II was to contribute to the development of ‘a community-based approach to policing’ (DfID 1999, p.11).

As stated in the terms of reference, the evaluation objectives were to assess fulfilment of project objectives and overall impact, with a further task to assess “whether or not there is a public perception that the behaviour of the police has changed as a result of the assistance programme” p.35). In fact, findings are mainly specific to the project activities with limited assessment of wider impact. Findings are generally positive with a rating of ‘partially successful’. Difficulties in providing an impact assessment are partly attributed to the lack of baseline data from which to make clear judgements on outcomes. The absence of such data, plus insufficient attention to impact through ongoing monitoring, are regarded as project weaknesses, imposing limitations on ex-post evaluation work (p.3). But it is acknowledged that DfID has only paid greater attention to performance and impact assessment more recently. Caution is also expressed regarding undue expectations of impact, noting that “it is important to maintain a sense of perspective...[with] reasonable expectation[s] of ... an extremely modest investment. It would be wrong to expect that a project of this scale on its own would be able to bring about a fundamental reform of policing in Indonesia” (p.3).

Such carefulness is less in evidence when it comes to discussing the project’s effect on the militaristic style of the INP, said to be implied in the promotion of a community-based approach modelled on the UK. The previously stated difficulties in evaluating impact are less apparent with the sweeping statement that, “The project has succeeded in helping to lay down a number of important building blocks for the future demilitarisation of policing in Indonesia” (p.3). Similarly, regarding the public perception of the police, it is stated that, “The team gained the impression from many sources that the general image of the police has improved somewhat in recent years, albeit from a low base” (p.3), though the nature of the sources and evidence gathered remain unspecified. A degree of credit for such an outcome appears to be attributed implicitly to the project.

Such unguarded statements contribute to the overall valedictory nature of the evaluation, which becomes most evident when questions of human rights are addressed directly.

Unlike the Uganda evaluation, no criticism is made that “the project did not explicitly set out to address the human rights issue” (p.33). On the contrary, it is argued that such a concern “is *implied* in its aim of developing a policing style that is less militaristic and more in line with UK practice” (p.33, emphasis added). Criticisms voiced in the UK of the inappropriateness of providing assistance to those very forces that are responsible for human rights violations is tested by putting this question to Indonesian human rights groups. All groups are reported as agreeing that the “continuing involvement of outside

agencies was desirable and was likely to have a moderating effect on the behaviour of the police” (pp.33-4), though again methods, sources and data collected remain unknown. In this way the project is legitimated and defended by the evaluators. Indeed they go further by the counterfactual statement that “a more assertive stance on human rights *would* have diminished the project’s influence and created resistance to change within the force” (p.4, emphasis added).

The lessons learned concur with this valedictory approach, with the foremost lesson that “it is possible, possibly even desirable, to work with police forces in politically sensitive and difficult circumstances” (p.34).

In this way, the evaluation amounts to a ringing endorsement of the project in the face of criticism on human rights grounds within the UK, in parliament and from NGOs. The assertions concerning the positive impact of the project on the nature of the INP do not appear well founded, however, with a lack of supporting evidence. There seems to be an unbridged gap between the specific project outcomes, the training of individual officers, for instance, and the wider impact that is claimed. A rather different outlook is provided by the report authored by the British MP Ann Clywd, citing evidence of torture in East Timor by police under the command of an officer who had received aid-funded training in the UK, (cited in World Development Movement 1995, p.7).

*Evaluation Of ODA/DfID Support To The Police In Developing Countries: A Synthesis Study* (Biddle, Clegg & Whetton 1999 2 vols.)

The final report in this section stands out for the way in which it combines detailed accounts of DfID programme support for police in developing countries with a careful consideration of the assumptions that lie behind these interventions and the range of problems and issues that face countries under-going wider police reform. In terms of its range and analysis, it is certainly one of the better evaluations we have read for this research and deserves proper attention.

Biddle, Clegg and Whetton collect much of their data from existing “ex-post evaluation studies of three major projects . . . , material collected on a number of other police projects, expenditure data provided by DFID and bibliographic research”, spanning ten years of development assistance for police reform. It is therefore principally a desk study, though the consultants did conduct a number of interviews with “key actors”, including in South Africa, where they visited a number of projects. The main body of the report (vol.1) is an analysis of the broad aims, assumptions, means and issues informing DfID’s policy towards police reform. The second volume consists of a series of supplementary annexes, dealing with costings, etc, and, more importantly, case studies of Indonesia, Namibia, Uganda, Ethiopia, Nepal, Lesotho, South Africa, Commonwealth Caribbean.

The authors find a number of serious flaws with the way in which assistance has been designed and implemented since police support became the responsibility of ODA in 1992. Previously, the FCO had managed request from governments for police assistance on an ad hoc basis. When ODA took over, policing became one aspect of a wider concern to promote ‘good government’: “Projects were being tied more closely to themes such as

openness, accountability, safety and security (and to a lesser extent, gender), with a stronger emphasis on community policing and the need for attitude change amongst the police” (vol.1, p.8). With the creation of DfID, emphasis again changed, albeit slightly, to push the pro-poor elements of good policing practice: “Freedom from crime, safety and security, safeguarding of human rights and access to justice are seen as essential ingredients of a pro-poor approach to economic growth and sustainable development.” (p.9). ODA and DfID both emphasised the importance of community policing in their work, which is regarded “as one of the special contributions that the UK has to offer to overseas police forces” (p.12). [NB. No definition of community policing is provided]. This fits nicely with DfID’s pro-poor and ‘good government’ agenda but is perhaps too idealistic a target to achieve by the means available for the following reasons:

- Firstly, police reform is a highly political area: “Most of the police projects reviewed have taken place within the context of a transition from single-party or more-or-less undemocratic regimes to forms of multi-party democracy. Many of the police forces have a history of being essentially instruments of and for the state. Frequently combating crime and providing safety and security for the citizen took second place to maintaining public order and internal security” (p.11). This has meant that projects have not simply encountered institutional resistance, but require a complete change of culture. This clearly takes time and effort, and, more importantly, is contingent on a supportive environment and parallel reforms in related sectors such as the courts and prisons.
- These environmental factors have clearly weakened the impact of DfID’s institutional strengthening projects, which take as their premise “that the key to achieving change lies in improving the managerial and operational capacities of the police and, at the same time, in trying to re-orientate what might be called police ‘culture’ in the direction of greater openness and service to the community” (p.11). Quite simply, while managerial/ institutional-level changes are a component of police reform, they are not sufficient to create good policing, which depends on a great variety of other factors such as “improvements in the protection of human rights, public safety, access of the poor to justice, accountability, protection of vulnerable groups” (p.11).
- While community policing is often considered a means of making the police more accountable and acceptable to local communities, the authors note that the term ‘community’ has not been sufficiently analysed: “In reality, communities are characterised by a lack of homogeneity and are made up of groups with differing interests and power. Frequently poorer and marginal groups, who suffer disproportionately from the impact of crime, have been excluded from participation in community policing” (p.32).
- Perhaps most significantly of all, the authors observe that supporting the police or even community policing carries significant political implication about the nature of the state and its relationship with its citizens. There are essentially two strands to this argument. Firstly, the authors problematise the relationship between DfID’s overarching goals of supporting good government and economic development and support for the police; “How does the production and maintenance of the social order which is assumed to be a pre-requisite for development depend on the presence of an

organised formal police force?” (p32). Secondly, they draw attention to recent writing that notes:

“Policing can be carried out by different people using different techniques. It may be carried out by professionals employed by the state with a general mandate (the classic modern idea of the police), by specialist private policing companies, by citizens in a voluntary capacity acting within or alongside the state or public police or as vigilantes. Until modern times most policing functions were not carried out by specialised policing institutions. A recent review of the development of the police [Robinson, Scaglione, and Olivero, 1994] concludes that the development of specialised police is linked to economic specialisation and the differential access to resources which occur during the transition from a kinship to a class dominated society. During this process, communal policing forms are gradually converted to state dominated ones which begin to operate as agents of class control as well as their more general social control function.”(Pp.32-33).

What is at stake here are the wider political implications of aid where it undertakes to provide developing states with the ‘requisite’ institutions of the modern western state. Aside from having “largely ignored” existing indigenous means of providing justice and dispute resolution until recently (p.33), DfID policy on support for the police may well create modern, specialised state structures which the state may be hard pushed to adequately finance, and, by implication, control. As a corollary of this, projects appear to be planned with an insufficient appreciation of the size of the task they are trying to perform, and therefore expect to achieve more than is possible (pp.32-33).

Other issues highlighted include:

- A lack of contextual knowledge and basic research in project planning (p.vi);
- “a lack of congruence, throughout the project cycle, between major stakeholders: DFID and FCO in London and locally, partner governments and police forces” (p.v);
- DfID staffing issues, which are poor to the extent that they affect project performance (“TCOs [UK police officers and police trainers], who have little experience of overseas work, have noted inadequate briefing and professional isolation” [p.v]);
- The problems of trying to improve institutional effectiveness by means of training when “[The] Equipment and infrastructure available to some forces are chronically inadequate [and] .... In some cases ... clearly limit the effectiveness of non-material aid.” (p.25).

The second volume of the report presents a series of appendices on expenditure, monitoring and evaluation, and summaries from eight country of regional case studies. One criticism of this report concerns the lack of integration of case study material into the body of the text. Very little reference is made to the case study findings in volume one, which is surprising if one assumes that the case studies played a significant role in informing the authors’ general comments and conclusions. It may be that the decision to focus on the more theoretical and policy issues was the right one (certainly the report is

strong and well-argued as it stands), however further specificity could have been added to findings by closer use of case study material.

The principal lesson to be learned from the case studies is that projects have not had the intended impact because they were drawn up and implemented without a detailed consideration of the environment, or political context, in which they were to operate. Five of the eight case studies have been selected to illustrate the type of findings provided.

#### *Indonesia*

- Good outputs achieved regarding the training of trainers failed to translate into good outcomes when many trainers were reassigned to other duties (vol.2. p.22);
- An internal consultancy unit in the Indonesia National Police was successfully established to facilitate change, but its output and impact are less certain in the light of doubts over its long term future; arising from “lack of continuity of personnel, insufficient budget resources to establish full time post for consultants leading to pressure to attend to other areas of police work, variable support from Division Chiefs and inadequate indicators and monitoring and evaluation.” (vol.2, p.23)
- Some progress was achieved towards establishing community-based policing, but low rates of crime reporting indicating lack of trust in police. Furthermore, progress towards improvements in respect for human rights was significantly hampered by political considerations: “To have placed greater stress on human rights issues would have risked losing the commitment of INP management and ultimately of rejection by the government” (p.25).

#### *Namibia*

- Poor political circumstances in Namibia hampered the effectiveness of the project, compounded by internal problems with ODA related to change

#### *Uganda*

- The Uganda project was seen as successful in many ways, but a desire by police management to keep officers on the move (characterised as a conscious choice to “avoid corruption, but wasted training” [vol.2, p.36]) seriously diluted any outcomes the training programme made have had for working practices and any impact on beneficiaries.
- This particular project is also singled out for its “chronic problem” with Monitoring and Evaluation (vol.2, p.36), although this is an issue that appears to effect all policing activities.

#### *Ethiopia*

- The tenor of evaluation reports on Ethiopia is generally positive in terms of project outputs: “Most of the activities specified in the logframe have been achieved” (vol.2, p.47); but authors also acknowledge that serious and deep seated problems remain.

- This project (like an earlier one with Indonesian police) was terminated early by the British authorities over concerns about police brutality. The authors register a note of regret, stating that at the time of its cessation there was “still a long way to go before the EPF [Ethiopian Police Force] could be said to have been sufficiently transformed to meet the overall goals and purposes of the project which were to develop an open and accountable police service capable of delivering services in line with community needs” (vol.2, p.49).

### *Nepal*

- The Nepali case study highlights that institutional reform is not something that can be simply delivered in a short space of time, but does rather require long-term commitment from donors and partnership with both government and police authorities/ recipients. Specifically, the authors concluded that “despite the good start”, “five years were not enough to ensure effective and sustainable change” in an environment characterised by weak openness and accountability and where the role of the state is ill-defined and contested by civil society (vol.2, p55). This finding in effect brings us back to earlier conclusions about the inter-relatedness of police reform, judicial reform, and the nature of the relationship between state and society.

In conclusion, Biddle, Clegg and Whetton provide a thorough-going and valuable analysis of the issues facing support to police. They go considerably beyond the remit of the other evaluation studies examined here, and provide a useful account of the complexity of reform, formed in the realisation that the connections between the overarching goals of development assistance (say, promoting good government, human rights or democracy) and the means adopted to do so (by delivering, say, electoral assistance or police reform) are tenuous or weak. The principal reason for this is that insufficient attention is paid to the country context, meaning that there is little knowledge of how institutional, political and cultural specificities may affect an agency’s plans. In part, one can surmise, this may be because recipients are insufficiently involved in project planning, and that development in many ways continues to be a question of meeting the delivering agency’s objectives.

### ***Club du Sahel, OECD Development Co-operation Directorate and UNDP***

#### *Improving the Effectiveness of Aid Systems: The case of Mali (1999)*

Along with the Club du Sahel and UNDP, the Development Co-operation Directorate of the OECD instigated this recent major report into aid effectiveness, based on the case study of Mali (Club du Sahel et al., 1999). Although not at all specific to democracy and governance assistance, important process-oriented lessons can be extracted on aid evaluation methods, notably the significance of the involvement of national actors as a means to improve the assistance offered.

Through the case study of Mali, the objective was an ambitious one: to review the international aid system, assessing its general effectiveness and highlighting both best practices and failings. The methodology comprised an extensive review of a decade of

development assistance (1985-95), conducted with an apparently wide degree of control and participation from Malian stakeholders, both government and non-government. The process by which the review was instigated and implemented was as follows:

- The lead agencies (i.e. the Malian government and the DAC) jointly informed “all the parties concerned” about the review. This group included the headquarters and field office staff of development agencies (Mission Française de Coopération, Dutch Embassy, European Commission delegation, Swiss Co-operation liaison office, Canadian International Development Agency, German Technical Co-operation Agency, UNDP mission, World Bank) and organisations and representatives of civil society (NGOs, professional associations, information and discussion centres), (pp.17-8).
- The Malian government established an Aid Review Monitoring Committee, consisting of officials from the Ministries of Finance and Foreign Affairs, other government departments, the Commissariat au Plan (Planning Commission), and representatives of external aid and co-operation agencies, under the chairmanship of the Commissariat au Plan.
- With assistance from the DCD, the Commissariat au Plan contacted donor headquarters and field offices, and also civil society organisations. The latter are informed of the review, and asked to “participate in future exchanges or even in the organisation of discussion and analysis” (p.18). It appears from the report, however, that only one NGO/CSO actually contributed or was asked to contribute to the process.
- The Monitoring Committee defined the scope of the review, and, with the DCD, established a panel of experts from aid agencies and “Sahelian and West African personalities” to “monitor and advise on the orientation of the process by participating during the main junctures of the review” (p.18).
- Malian consultants were recruited to perform the review.
- The Monitoring Committee and DAC Experts drew up the methodology to be used by the consultants, drawing particularly on:
  - “(i) general documents relating to strategy, policy and evaluation, from the ministries responsible for the sectors in which aid is concentrated, and from aid agencies;
  - (ii) interviews with key Malian officials, members of civil society and donors, for the purpose of obtaining the lessons they have drawn from their experience and their general assessment of the efficacy of the aid system;
  - (iii) analysis of the entire project portfolio over the period 1985-1995 referring to the reports of co-operation agencies;
  - (iv) in-depth analysis of a small selection of significant projects, using interviews, monitoring and evaluation reports and field studies” (p.18).
- These proposals were then refined by the DCD (p.19). The terms of reference underwent a subsequent revision at a workshop at which the “Monitoring Committee,

officials involved in managing aid, representatives of Mali's civil society, the consultants selected for the review, representatives of the local co-operation agencies, the panel of experts appointed for the exercise, a DAC delegation and observers from the Club du Sahel" were all present (p.19).

- A participative element to the review was introduced by a Malian NGO (the Djoliba Centre), aimed at providing "civil society perceptions and proposals". This entailed "direct participation" by means of a survey questionnaire, a public conference, and radio discussions in local languages in three regions of the country (p.20).

The participatory process seemed to fall away somewhat at the end. Consultants' and NGO findings were discussed at a meeting attended by "all the parties involved in the process" held in Paris (p.20), although this comprised only a small Malian delegation, with no mention of civil society involvement. Furthermore, the implications of the meeting for the review were *never fully worked out* because a synthesis meeting to be held in Mali in order "to familiarise Malian officials with the exercise and to build a consensus around the future plan of reform" did not take place (p.21). The report states that "all the parties wished to pass directly to the operational phase after the Paris meeting" (p.21). Yet this is contradicted by the statement that "The failure to hold the [second, Mali-based] meeting ...delayed the emergence of an agreed basis for action, [with] a small number of donors expressing reservations about the provisional document and some of the recommendations formulated during the Paris discussions" (p.21).

The review process continued but in a more ad hoc fashion, focused on "the relevant players" (p.21) from among the Malian government and donor community rather than on the initial, more broadly participative approach. In this ad hoc manner, the conclusions of the Paris meeting were discussed, generating proposals "to the government that it embark on a reform of the aid system in Mali" (p21). This section of the report fails, however, to indicate any mechanisms of consultation and involvement, suggesting that they were limited, at best.

Despite the implementation problems, this study is generally useful for its methodological approach. The report itself strongly endorses the participatory process and the authors believe that it is an example of what can be achieved with commitment to greater national involvement. The study also raises wider issues concerning public participation in development assistance. In particular it was found that although "The relevance of projects and programmes was judged to be very satisfactory...[this] is hard to reconcile with the views collated by the Djoliba Centre [an NGO that canvassed public opinion as part of the review], which found that 'local people had the impression that aid took no account of their priorities'" (p.25). The review continues: "We are bound to ask whether aid objectives seem relevant because project or programme objectives match the priorities and strategies of national players, or because their priorities and strategies have been tailored to the declared objectives of aid" (p.25).



## **World Bank**

The introduction of the concept of ‘good governance’ into development aid policy owes much to the World Bank. The concept first came to prominence in the 1989 World Bank report *Sub-Saharan Africa: From Crisis to Sustainable Growth*, and subsequently has become a key element of World Bank interactions with developing countries in all regions. Public sector management, inclusive of civil service reform, has been a key component of governance activities, strongly associated with structural adjustment programmes. Other elements have included the legal framework for development, and the promotion of accountability and transparency (World Bank 1992). The one large-scale evaluation study in this area has focused on civil service reform, examined below.

### *Civil Service Reform: A Review of World Bank Assistance (1999)*

Conducted by the Bank’s Operations Evaluation Department, this is a desk study review of civil service reform projects, both completed and ongoing. Although broad ranging in scope, the study does not entail fieldwork or attempt to assess impact. The review sought to: evaluate Bank strategy; examine the efficacy of different classes of civil service reform (CSR); identify key factors shaping CSR performance. A large sample of 124 loans (comprising 350 interventions) in 32 countries between 1980-97 is assessed for their strategic relevance (i.e. the fit between policy content and country context) and their efficacy (i.e. the achievement of stated objectives), although the evaluation methodology is not specified.

In examining Bank strategy in this area over the two-decade period, the report analyses three forms of ‘bureaucratic dysfunction’ and three corresponding categories of civil service reform, with a clear sense of chronological development and differential emphasis. The problem diagnosed in the early 1980s was a fiscal one arising from overextended public sectors (‘bloated bureaucracies’). This was coupled with concerns about capacity constraints, especially regarding core economic management functions necessary for structural adjustment. Consequently, the strategy implemented throughout the 1990s was ‘do more with less’. On the one hand the civil service was downsized, while on the other administrative capacity was enhanced. Some lack of success, however, led to the introduction of a *governance* approach in the early 1990s, that is a focus on the wider institutional context that influences civil service reform, with strategy shifting to broader institutional reform, both inside and outside of government. Such institutional reforms “sought to make the state ‘more accountable and open’, rather than simply more efficient” (World Bank 1999, p.8). Measures covered both the revision of intra-public sector rules and regulations to improve transparency and accountability, and a strengthened role for external organisations that monitor and limit arbitrary action (for example, an independent judiciary, oversight bodies like an Auditor General, NGOs, greater public access to information).

The review ‘unbundled’ and categorised civil service reform projects into these three classes of measures, downsizing, capacity building, and institutional reform, then assessed ‘relevance’ and ‘efficacy’ for each intervention. The overall finding was not of satisfactory performance of CSR measures: “The review found that Bank-supported

CSRs were *largely ineffective* in achieving sustainable results in downsizing, capacity building, and institutional reform” (p.iii, emphasis added). Regarding efficacy, “only 33 percent of closed CSR interventions [i.e. completed projects] and 38 percent of ongoing efforts achieved satisfactory outcomes” (p.ii). Downsizing and capacity building efforts failed both to produce permanent reductions in size and to overcome capacity constraints in economic management. While “institutional reforms could not substantially limit arbitrary action by bureaucrats or politicians”, with “no evidence” of ‘ownership’ by civil servants and the adoption of codes of ethics in any meaningful way (p.ii), and the absence of effective checks and balances.

As a desk study, there is no participatory element in the evaluation itself. But the discussion of participation in civil service reform indicates an interesting shift. It is acknowledged that “the use of participation as *an approach to intervention* was largely absent from CSR interventions”, in fact actively “discouraged due to fears of policy capture by organised interests such as public sector unions” (p.12). Yet, the establishment of the governance approach, inclusive of the analysis that it has yet to bear fruit, has entailed a fundamental shift of position. Recommendations here include that “The Bank *should* employ participatory processes to nurture reform constituencies in government, the private sector and civil society” (p.iii, emphasis added). It must be queried, however, whether this represents a genuine approach to participation or a manipulative one.

Questions arise concerning the respective economic and/or political dimensions of civil service reform. On the one hand, the economic element is explicit. For the past 2 decades, the World Bank’s civil service reform measures have been an integral part of structural adjustment programmes, themselves linked to market liberalisation. Frankly stated by the Bank: “civil service reforms [have] served as a primary vehicle for removing institutional impediments to market-led development” (p.i). On the other hand, any political dimension is more concealed. Formerly, civil service reform, or public administration development more generally, could be perceived as aiming to enhance state efficiency and effectiveness whatever the nature of the political regime. Such an orientation also squares with the World Bank’s mandate to be non-political. Yet, whether by intent or by default, a governance approach introduces not merely a political but a democratic dimension. Institutional reform entails the strengthening of checks and balances to executive power, both state-related organisations (for example, audit bodies) and non-state (civil society) organisations. It is difficult to conceive of such organisations operating effectively other than in a democratic context. Yet, a further question remains, are democratic processes valued in themselves, or as a means to the slimmed-down but effective state that facilitates market-led economic development? Whatever its mandate, perhaps the World Bank can genuinely say that it is not its objective to promote democracy.

## Appendix Two

### *Agencies' responses to enquiries*

#### *Positive Response*

Australia	AusAID
Austria	BmaA
Canada	CIDA
Denmark	Danida
European Commission	Directorate for External Affairs
European Commission	Europe Aid
European Commission	Evaluation Office
Finland	Department for International Development Co-operation
Finland	Ministry of Foreign Affairs
France	Department for International Co-operation
Germany	BMZ
Ireland	Ireland Aid
Japan	JICA
Netherlands	Ministry of Foreign Affairs
New Zealand	Ministry of Foreign Affairs
Norway	Ministry of Foreign Affairs
Norway	NORAD
Sweden	SIDA
United Kingdom	DFID Evaluation Unit
United States	USAID
Intergovernmental	DAC
Multilateral	World Bank –PovertyNet
Multilateral	UNDP
Other	Ford Foundation

#### *No response*

Belgium	Ministry of Foreign Affairs
France	Le Groupe de l'Agence française de Développement (AfD)
Germany	GTZ
Greece	Ministry of Foreign Affairs
Italy	Ministry of Foreign Affairs
Japan	Ministry of Foreign Affairs
Luxembourg	Lux-Development
Portugal	Ministry of Foreign Affairs
Spain	Spanish Agency for International Cooperation (AECI)
United Kingdom	DFID Government and Institutions Department
Multilateral	World Bank – OED

# Bibliography

## Books and articles

- Abrahamsen, R. (2000), *Disciplining Democracy: Development Discourse and Good Governance in Africa* (Zed Books, London).
- Beetham, D. (1999), *Democracy and Human Rights* (Polity Press, Cambridge)
- Beetham, D., Bracking S., Kearton I. and Weir S. (2001), *International IDEA's Handbook on Democracy Assessment*, (International IDEA, Stockholm)
- Brouwer, I. (2000), *Bringing Democracy Assistance Recipients In*, (Florence/Berlin, European University Institute/Humboldt-Universität zu Berlin), Research Project: 'Democracy Promotion and Protection in Central & Eastern Europe and the Middle East & North Africa', Discussion Paper No. 3/2000.
- Campilan, D.M. (2000), 'Conceptual tools for tracking change: Emerging issues and challenges' in Estrella, M. et al. (eds.), pp.192-200
- Carothers, T. (1997) 'Democracy Assistance: The Question of Strategy', in *Democratization* vol. 4 no. 3, pp.109-32.
- Carothers, T. (1999) *Aiding Democracy Abroad: The Learning Curve*, (Carnegie Endowment for International Peace, Washington D.C.).
- Cracknell, B.E. (2000), *Evaluating Development Aid: Issues, Problems and Solutions*, (Sage, London).
- Crawford, G. (1995), *Promoting Democracy, Human Rights and Good Governance through Development Aid: A Comparative Study of the Policies of Four Northern Donors*, (Centre for Democratisation Studies, University of Leeds).
- Crawford, G. (2000a) 'Promoting Democratic Governance in the South', in *European Journal of Development Research* Vol.12 No.1
- Crawford, G. (2000b) 'European Union Development Co-operation and the Promotion of Democracy', in P.Burnell (ed.) *Democracy Assistance: International Co-operation for Democratization*, (Frank Cass, London)
- Crawford, G. (2001), *Foreign Aid and Political Reform: A Comparative Analysis of Democracy Assistance and Political Conditionality*, (Palgrave, Basingstoke).
- Estrella, M. et al. (eds.) (2000), *Learning from Change: Issues and experiences in participatory monitoring and evaluation*, (Intermediate Technology Publications, London and International Development Research Centre, Ottawa)
- Estrella, M. (2000), 'Learning from change' in Estrella, M. et al. (eds.), pp.1-14
- Estrella, M. and Gaventa J. (1998), *Who Counts Reality? Participatory Monitoring and Evaluation: A Literature Review*, IDS Working Paper 70, Institute of Development Studies, Brighton.
- Fernandes W. and Tandon R. (eds.) (1981), *Participatory Research and Evaluation: Experiments in Research as a Process of Liberation* (Indian Social Institute, New Delhi).
- Gaventa J. and Blauert J. (2000), 'Learning to change by learning from change: Going to scale with participatory monitoring and evaluation', in Estrella, M. et al. (eds.) (2000), pp.229-243

- Gills, B., Rocamora, J. and Wilson, R. (eds.) (1993) *Low Intensity Democracy: Political Power in the New World Order*, (Pluto Press, London).
- Guba, E.G. and Lincoln Y.S. (1989), *Fourth Generation Evaluation*, (Sage, London)
- Guijt, I. (2000), 'Methodological issues in participatory monitoring and evaluation' in Estrella, M. et al. (eds.), pp.201-16
- Helmich H. with Borghese E. (eds.) (1998), *Human Rights in Development Co-operation*, SIM Special No.22, (Netherlands Institute of Human Rights, Utrecht)
- International Institute for Democracy and Electoral Assistance (IDEA) (1998), *Dialogue for Democratic Development: Renewing the ACP-EU Partnership for the 21<sup>st</sup> Century*, (International IDEA, Stockholm).
- International Institute for Democracy and Electoral Assistance (IDEA) (2000), *Democratization in Indonesia: An Assessment*, Capacity Building Series 9, (International IDEA, Stockholm).
- Leftwich, A. (1996), 'On the Primacy of Politics in Development' in Leftwich, A. (ed.) *Democracy and Development: Theory and Practice*, pp.3-24, (Polity Press, Cambridge)
- Linz J. & Stepan A. (1996), *Problems of Democratic Transition and Consolidation*
- Marsden, D. & Oakley, P. (eds.) (1990), *Evaluating Social Development Projects*, (Oxfam, Oxford)
- Narayan-Parker, D. (1993), *Participatory evaluation: Tools for Managing Change in Water and Sanitation*, World Bank Technical Paper No. 207, (World Bank, Washington D.C.)
- Pawson, R. and Tilley, N. (1997), *Realistic Evaluation*, (Sage, London).
- Rebien, C.C. (1996), *Evaluating Development Assistance in Theory and Practice*, (Avebury, Aldershot).
- Robinson, M. (1996), *Strengthening Civil Society Through Foreign Political Aid*, [ESCOR Research Report R6234], IDS, Brighton
- Schmitter, P.C. and Brouwer I. (1999), *Conceptualizing Research and Evaluating Democracy Promotion and Protection*, Working Paper SPS No. 99/9, European University Institute, Florence.
- Van Rooy A. (1998) (ed.), *Civil Society and the Aid Industry* (Earthscan, London)

## **Evaluation Reports and Related Studies**

- AusAID (1998), *PNG Cluster Evaluation: Institutional Strengthening Projects*, (AusAID, Canberra)
- AusAID (2001), *Evaluation Issues Paper: Improving Governance in Order to Improve Economic Management Draft document*, (AusAID, Canberra)
- Biddle K., Clegg I. & Whetton J. (1999) *Evaluation Of ODA/DfID Support To The Police In Developing Countries: A Synthesis Study*, (University of Wales, Swansea)
- CIDA (1994a), *Accompanying Democracy: Lessons Learned El Salvador 1990-1993*, (CIDA, Ottawa)
- CIDA (1994b), *Lessons Learned in Human Rights and Democratic Development: A study of CIDA's bilateral programming experience*, (CIDA, Ottawa)
- CIDA (1995a), *Democratic Development and Human Rights (DDHR): West Africa*, (CIDA, Ottawa)
- CIDA (1995b), *Evaluation of CIDA's human rights program mechanisms in Sri Lanka*, (CIDA, Ottawa)
- CIDA (1995c), *Choosing the Right Policy Levers: Drawing lessons from the Government of Canada's interventions in South Africa*, (CIDA, Ottawa)

- CIDA (1996), *Democracy and human rights: What are we learning?*, (CIDA, Ottawa)
- CIDA (2000), *CIDA Evaluation Guide*, (CIDA, Ottawa)
- Club du Sahel, OECD Development Co-operation Directorate & UNDP (1999), *Improving the Effectiveness of Aid Systems: The case of Mali*, (OECD DCD, Paris)
- Danish Ministry of Foreign Affairs/Danida (2000), *Evaluation of Danish Support to Promotion of Human Rights and Democratisation 1990-1998: Vol. 1 Synthesis Report*, (Danish Ministry of Foreign Affairs, Copenhagen)
- Danish Ministry of Foreign Affairs/Danida (2000), *Evaluation of Danish Support to Promotion of Human Rights and Democratisation 1990-1998: Vol. 2 Justice, Constitution and Legislation*, (Danish Ministry of Foreign Affairs, Copenhagen)
- Danish Ministry of Foreign Affairs/Danida (2000), *Evaluation of Danish Support to Promotion of Human Rights and Democratisation 1990-1998: Vol. 3 Elections*, (Danish Ministry of Foreign Affairs, Copenhagen)
- Danish Ministry of Foreign Affairs/Danida (2000), *Evaluation of Danish Support to Promotion of Human Rights and Democratisation 1990-1998: Vol. 4 Media*, (Danish Ministry of Foreign Affairs, Copenhagen)
- Danish Ministry of Foreign Affairs/Danida (2000), *Evaluation of Danish Support to Promotion of Human Rights and Democratisation 1990-1998: Vol. 5 Participation and Empowerment*, (Danish Ministry of Foreign Affairs, Copenhagen)
- Danish Ministry of Foreign Affairs/Danida (2000), *Evaluation of Danish Support to Promotion of Human Rights and Democratisation 1990-1998: Vol. 6 Ghana*, (Danish Ministry of Foreign Affairs, Copenhagen)
- Danish Ministry of Foreign Affairs/Danida (2000), *Evaluation of Danish Support to Promotion of Human Rights and Democratisation 1990-1998: Vol. 7 Guatemala*, (Danish Ministry of Foreign Affairs, Copenhagen)
- Danish Ministry of Foreign Affairs/Danida (2000), *Evaluation of Danish Support to Promotion of Human Rights and Democratisation 1990-1998: Vol. 8 Mozambique*, (Danish Ministry of Foreign Affairs, Copenhagen)
- Danish Ministry of Foreign Affairs/Danida (2000), *Evaluation of Danish Support to Promotion of Human Rights and Democratisation 1990-1998: Vol. 9 Nepal*, (Danish Ministry of Foreign Affairs, Copenhagen)
- Department for International Development /Overseas Development Administration (DfID/ODA), UK, (1993), *Evaluation of ODA Project in Support of Ghana Civil Service Reform Programme*, (DfID, London)
- Department for International Development (DfID), UK, (1998), *Uganda Police Project Evaluation*, (DfID, London)
- Department for International Development (DfID), UK, (1999), *Evaluation of the Indonesia National Police Management Training Project, 1983-96*, (DfID), London)
- Dutch Ministry of Foreign Affairs / IOB (1999), *Palestinian Territories: Review of the Netherlands development programme for the Palestinian Territories*, (Dutch Ministry of Foreign Affairs, The Hague)
- European Commission (1995), *Communication to the Council and Parliament on The European Union and the External Dimension of Human Rights Policy: From Rome to Maastricht and Beyond*, [COM (95) 567, 22 November 1995], (European Commission, Brussels).

- European Commission (1997), *Evaluation of the PHARE and TACIS Democracy Programme 1992-1997*, (European Commission, Brussels)
- European Commission (1999a), *Appraisal of the EDF Support to the 1999 General Elections and Evaluation of the 1998 Municipal Evaluation in Mozambique: Final Report*, (European Commission, Brussels)
- European Commission (1999b), *Evaluation of the MEDA Democracy Programme 1996-98*, (European Commission, Brussels)
- European Commission (1999c), *Evaluation of EC Country Strategy: South Africa 1996-1999*, (European Commission, Brussels)
- European Commission (1999d), *Evaluation of EC Country Strategy: The Dominican Republic 1996-2000*, (European Commission, Brussels)
- European Commission (2000a), *Evaluation of EC Country Strategy: Papua New Guinea 1996-1999*, (European Commission, Brussels)
- European Commission (2000b), *External Evaluation of Community Aid Concerning Positive Actions in the Field of Human Rights and Democracy in the ACP countries*, (European Commission, Brussels)
- European Commission (2000c), *Report on the Implementation of Measures Intended to Promote Observance of HR and Democratic Principles in External Relations for 1996-99*, (European Commission, Brussels)
- European Commission (2000d), *Fiche contradictoire*, (European Commission, Brussels)
- European Commission (2001), *The European Union's Role in Promoting Human Rights and Democratisation in Third Countries* [COM(2001) 252 final, 8 May 2001], (European Commission, Brussels)
- European Court of Auditors (2000), *Special report No. 12/2000 on the management by the Commission of European Union support for the development of human rights and democracy in third countries, together with the Commission's replies*, (European Commission, Brussels)
- European Court of Auditors (2001), *Special report No. 21/2000 on the management of the Commission's external aid programmes (in particular on country programming, project preparation and the role of Delegations) together with the Commission's replies* (European Commission, Brussels)
- FINNIDA (1996), *Effects or Impacts? Synthesis Study on Evaluations and Reviews 1988 to Mid 1995*, (University of Helsinki, Helsinki)
- Heinz, W., Lingnau, H. and Waller P. (1995), *Evaluation of EC Positive Measures in Favour of Human Rights and Democracy*, (German Development Institute, Berlin).
- International Development Research Center (IDRC) (1999) *Evaluating Governance Programmes: Report of a Workshop: 8 April 1999*, (IDRC, Ottawa)
- Ireland Aid (1999), *Report on the Final Part of the Programme on Constitutionalism for the Tanzanian Judiciary*, (Ireland Aid, Dublin)
- Ireland Aid (2000), *Human rights and democratisation in Mozambique: A three year strategy*, (Ireland Aid, Dublin)
- Japanese Ministry of Foreign Affairs (1999), *Japan's Support for Democratization*, (Ministry of Foreign Affairs, Tokyo)

- JICA (1995), *Participatory Development and Good Government: Report of the Aid Study Committee*, (JICA, Tokyo)
- JICA (2000), *Report on the Reform of Japan's ODA Evaluation System*, (JICA, Tokyo)
- Kapoor, I. (1996), *Indicators for Programming in Human Rights and Democratic Development: A Preliminary Study*, (CIDA, Ottawa)
- NORAD (2001), *Evaluation of Norwegian Support to Human Rights NGOs in Pakistan*, (NORAD, Oslo)
- Norwegian Ministry of Foreign Affairs (1997a), *Evaluation of Decentralisation and Development* [Evaluation Report 3.97], (Royal Ministry of Foreign Affairs, Oslo).
- Norwegian Ministry of Foreign Affairs (1997b), *Aid as a tool for the promotion of human rights and democracy: What can Norway do?* [Evaluation Report 7.97], (Royal Ministry of Foreign Affairs, Oslo).
- OECD DAC (1997a), *Evaluation of Programmes Promoting Participatory Development and Good Governance: Synthesis Report*, (OECD, Paris)
- OECD DAC (1997b), *Final Report of the Ad Hoc Working Group on Participatory Development and Good Governance, Parts I & II*, (OECD, Paris)
- OECD DAC (1997c), *Searching for impact and methods: NGO evaluation synthesis study: A report prepared for the OECD/DAC Expert Group on Evaluation* (University of Helsinki, Helsinki)
- OECD DAC (1999), *Evaluating Country programmes: Vienna Workshop 1999*, (OECD, Paris)
- Rawkins, P. and Qualman A. (2000). *Democracy and governance programming lessons for CIDA: Ethiopia Case Study*, (CIDA, Ottawa)
- Schacter, M. (2001), *Sector Wide Approaches, Accountability and CIDA: Issues and Recommendations*, (CIDA, Ottawa)
- Sida (1996), *Evaluation and Participation - Some lessons*, (SIDA, Stockholm)
- Sida (1999a), *Diakonia programme for democracy and human rights, the El Salvador case: A qualified monitoring*, (SIDA, Stockholm)
- Sida (1999b), *Access to Justice in Rural Nicaragua: An independent evaluation of the impact of rural court houses*, (SIDA, Stockholm)
- Sida (2000a), *Swedish Support to Local Self Governance in Mongolia*, (SIDA, Stockholm)
- Sida (2000b), *The Evaluability of Democracy and Human Rights Projects* [Sida Studies in Evaluation 00/3], (SIDA, Stockholm)
- Sida (2000c), *Study of the Swedish Philippine NGO Programme: Final Report*, (SIDA, Stockholm)
- Sida (2000d), *Strengthening Democracy on the Atlantic Coast in Nicaragua*, (SIDA, Stockholm)
- SPM Consultants (1996), *Evaluation of the European Special Programme on South Africa*, (European Commission, Brussels)
- Sunderland S.L. (1999), *Supporting Democracy: The South Africa-Canada Program on Governance*, (IDRC, Ottawa )
- UNDP (1995a), *Programme support to governance, democracy and human rights in Cambodia*, (UNDP, New York)
- UNDP (1995b), *Support to the police of the Republic of Mozambique*, (UNDP, New York)
- UNDP (1997a), *Governance for Sustainable Human Development*, (UNDP, New York)



UNDP (1997b), *Who Are the Question-makers? A Participatory Evaluation Handbook*, (UNDP, New York).

UNDP (1998a), *UNDP and Governance: Experiences and Lessons Learned*, (UNDP, New York)

UNDP (1998b), *Evaluation of the Governance Programme for Latin America and the Caribbean*, (UNDP, New York)

UNDP (1998c), *Strengthening the legal capacity in Vietnam (Phase2)*, (UNDP, New York)

UNDP (2000a), *The UNDP role in decentralisation and local governance: A joint UNDP-Government of Germany evaluation*, (UNDP, New York)

UNDP (2000b), *Support to the Strengthening of the Capacity of Tanzania's National Assembly: Draft evaluation report*, (UNDP, New York)

USAID (1994), *Weighing in on the Scales of Justice: Strategic Approaches for Donor-Supported Rule of Law Programs*, (USAID, Washington D.C.)

USAID (1995a), *Constituencies for Reform: Strategic Approaches for Donor-Supported Civic Advocacy Programmes* (USAID/CDIE, Washington D.C.)

USAID (1995b) *Evaluate the Past to Build the Future: Report on the Second Phase of the Evaluation of the Impact of the Centro DEMOS Programme*, (USAID, Washington D.C.)

USAID (1996a), *Legislative Strengthening in El Salvador*, (USAID, Washington D.C.)

USAID (1996b), *Parliamentary Assistance in Nepal*, (USAID, Washington D.C.)

USAID (1996c), *Evaluation of the Haiti communal governance program*, (USAID/OTI, Washington D.C.)

USAID (1996d), *Modernising Bolivia's legislature*, (USAID/CDIE, Washington D.C.)

USAID (1997a), *Assisting legislatures in developing countries: a framework for programme planning and implementation* , (USAID, Washington D.C.)

USAID (1997b), *Democratic Local Governance in Bolivia*, (USAID/CDIE, Washington D.C.)

USAID (1997c), *From Bullets to Ballots: Postwar Cambodia's Struggle with Democracy*, (USAID/CDIE, Washington D.C.)

USAID (1997d), *From Bullets to Ballots: 'Partidocracia' comes to El Salvador*, (USAID/CDIE, Washington D.C.)

USAID (1997e), *From Bullets to Ballots: Electoral Assistance to Post-Conflict Societies*, (USAID/CDIE, Washington D.C.)

USAID (1997g), *Democratic Local Governance in the Ukraine*, (USAID, Washington D.C.)

USAID (1997h), *Democratic Local Governance in the Philippines - Tradition hinders transition*, (USAID/CDIE, Washington D.C.)

USAID (1998a), *Handbook of Democracy and Governance Program Indicators*, (USAID, Washington D.C.)

USAID (1998b), *Spreading power to the periphery - An Assessment of Democratic Local Governance*, (USAID/CDIE, Washington D.C.)

USAID (1998c), *Democratic Local Governance in Honduras*, (USAID/CDIE, Washington D.C.)

USAID (1998d), *Democratic Decentralisation in Mali: Work in progress*, (USAID/CDIE, Washington D.C.)

- USAID (1999a), *1998 USAID Agency Performance Report*, (USAID, Washington D.C.)
- USAID (1999b), *A Handbook on Fighting Corruption*, (USAID, Washington D.C.)
- USAID (1999c), *Program Review FY1998*, (USAID, Washington D.C.)
- USAID (1999d), *USAID Political Party Development Assistance* (Technical Publication Series), (USAID, Washington D.C.)
- USAID (1999e), *USAID Annual Performance Plan for FY 2000*, (USAID, Washington D.C.)
- USAID (1999f), *The Role of Media in Democracy: A Strategic Approach*, (USAID, Washington D.C.)
- USAID (1999g), *Legislative Strengthening In Poland*, (USAID/CDIE, Washington D.C.)
- USAID (2000a), *USAID Agency Performance Report FY 1999*, (USAID, Washington D.C.)
- USAID (2000b), *G/DG Accomplishments FY1999*, (USAID, Washington D.C.)
- USAID (2000c), *Impact Assessment of OTI/Indonesia FY1999 Program*, (USAID/CDIE, Washington D.C.)
- USAID (2000d), *USAID's Experience in Decentralisation and Democratic Local Governance*, (USAID, Washington D.C.)
- USAID (2000e), *Agency Performance Plan for FY 2001*, (USAID, Washington D.C.)
- USAID (2000f), *USAID Handbook on Legislative Strengthening*, (USAID, Washington D.C.)
- USAID (2000g), *Managing Assistance in Support of Political & Electoral Processes*, (USAID, Washington D.C.)
- USAID (2000h), *Decentralization and Democratic Local Governance Programming Handbook*, (USAID, Washington D.C.)
- USAID (2000j), *Conducting a DG Assessment: A Framework for Strategy Development*, (USAID, Washington D.C.)
- USAID (2000k), *Country Case Studies of the Relationship between USAID DG Programs and Political Change: Research Design and Protocol for Year-One Set*, (internal document, final draft September 20, 2000)
- USAID (2001), *Civil Society Assessment Strategy for the Philippines*, (USAID, Washington D.C.)
- World Bank (1992), *Governance and Development*, (World Bank, Washington D.C.).
- World Bank (1999) *Civil Service Reform: A Review of World Bank Assistance*, (World Bank, Washington D.C.)

## **Other Sources**

- National Audit Office (1996), *Aid to Indonesia, Report by the Comptroller and the Auditor General, HC 101 Session 1997-97, 29 November 1996*, (HMSO, London)
- Human Rights Watch (1994) *Human Rights Watch World Report 1995: Events of 1994*, (Human Rights Watch, New York)
- Human Rights Watch (1995) *Human Rights Watch World Report 1996: Events of 1995*, (Human Rights Watch, New York).
- World Development Movement (1995), *Gunrunners Gold*, (World Development Movement, London)

World Development Movement (1997), *Submission to the Committee of Public Accounts on National Audit Office Report on Aid to Indonesia*, (World Development Movement, London)