Evaluating Democracy Support

Methods and Experiences

Evaluating Democracy Support

Methods and Experiences

Editor: Peter Burnell

Contributors:

Harry Blair
Héctor Chayer
Sandra Elena
Hanne Lund Madsen
Natalia Mirimanova
Patrick D. Molutsi
Margaret J. Sarles
Fredrik Uggla
Michael Wodzicki

Preface

Democracy support has grown dramatically in the past two decades, and so has interest in the methods and techniques of evaluating democracy support.

It is often asserted that evaluation of democracy support differs from the evaluation of other areas of development cooperation. In particular, it has been noted that the former field faces problems that relate to the diverse conceptions and definitions of democracy and democratization; the complex nature of democratization processes; and the difficulty of attributing changes at the national political level to individual projects. Such difficulties form the theoretical setting for the chapters of the present volume.

This book is based on the proceedings of a workshop on Methods and Experiences of Evaluating Democracy Support, organized by the International Institute for Democracy and Electoral Assistance (International IDEA) and the Swedish International Development Cooperation Agency (Sida), and held in April 2006. The main aim of the workshop was to explore ways in which existing methods and techniques of evaluating democracy support deal with challenges of causality and attribution. Clearly, this book is nowhere near providing answers to these questions, nor is it intended to. Rather, IDEA and Sida seek to share the main deliberations of the workshop, to stimulate further debates on the subject of evaluating democracy support and the challenges it faces, and—most importantly—to contribute to any new conceptualizations of methods and techniques for evaluating democracy support.

The workshop also aimed to bring together three different communities, using International IDEA's convening power and Sida's expertise—the community of evaluators, the community of democracy programme designers and planners, and the community of implementers and practitioners. It produced a rich debate and a meeting place for very different perspectives.

We take this opportunity to thank Professor Peter Burnell for the excellent work he has done in editing this book, as well as contributing the introductory chapter, which sets the publication in context. We thank Eve Johansson, whose professional input improved the readability of this publication tremendously. We also thank Keboitse Machangana, Advisor for Democracy Analysis and Assessment at International IDEA, and Fredrik Uggla of the Department for Evaluation and Audit at Sida for working tirelessly to bring this book to fruition.

Vidar Helgesen

International IDEA

Eva Lithman

Sida

Stockholm
August 2007

Contents

Chapter 3 71
Programme theory evaluation and democracy promotion: reviewing a
sample of Sida-supported projects Fredrik Uggla

Chapter 6 155
Evaluating a democracy support evaluation: the Rights & Democracy
ten-year taking stock exercise Michael Wodzicki

Chapter 7 171
Gauging civil society advocacy: charting pluralist pathways Harry Blair

Chapter 8 195

Evaluation of the utility of community-level democracy support for conflict
resolution: the Community Action Investment Programme in Tajikistan
Natalia Mirimanova

Chapter 9 217
The evaluation of democracy support programmes: an agenda for
future debate Patrick D. Molutsi

Acronyms and abbreviations

AKF	Aga Khan Foundation
BUCO	Building Unity for Continuing Coconut Industry Reform (the Philippines)
CAD	Canadian dollars
CAIP	Community Action Investment Programme (Tajikistan)
CIDA	Canadian International Development Agency
COCOFED	Coconut Producers Federation of the Philippines
COIR	Coconut Industry Reform Movement (Philippines)
CSO	civil society organization
Danida	Danish International Development Agency
EDSA	Epifanio de los Santos (Avenue) (Manila)
EMB	electoral management body
EU	European Union
FIAN	FoodFirst Information and Action Network
FORES	Foro de Estudios sobre la Administración de Justicia (Forum for Studies on Judicial Administration)
GDP	gross domestic product
GTZ	Deutsche Gesellschaft für Technische Zusammenarbeit (German Technical Cooperation Agency)
HDI	Human Development Index
IDB	Inter-American Development Bank
IDEA	International Institute for Democracy and Electoral Assistance
IMF	International Monetary Fund
JRI	Justice Reliability Index
JSCA	Justice Studies Center for the Americas
LFA	logframe analysis
m	metre
MSDSP	Mountain Societies Development Support Programme
MSSD	most similar system design
MTF	Multisectoral Task Force (the Philippines)
NAS	National Academy of Sciences (USA)
NBA	Narmada Bachao Andolan (Save Narmada Movement) (India)
NCSC	National Center for State Courts (Argentina)
NGO	non-governmental organization
NHRAP	national human rights action plan
NIMD	Institute for Multiparty Democracy
OECD	Organisation for Economic Co-operation and Development
OHCHR	Office of the High Commissioner for Human Rights (UN)
PCIJ	Philippine Center for Investigative Journalism
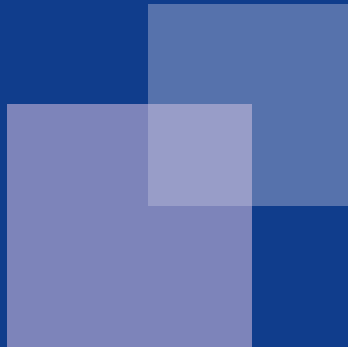
PEU	Programme Evaluation Unit
PKSMMN	Pambansang Koalisyon ng Magsasaka at Manggagawa sa Niyugan (coalition of NGOs representing small coconut farmers) (Philippines)
PRSP	poverty reduction strategy paper
PROJUM	Programa de Juzgado Modelo (Pilot Court Reform Programme) (Argentina)
PTE	programme theory evaluation
R&D	Rights and Democracy
R&R	resettlement and rehabilitation
RBA	human rights-based approach
ROL	rule of law
SADEV	Swedish Agency for Development Evaluation
Sida	Swedish International Development Cooperation Agency
SORA	Strategic and Operational Research Agenda (USAID)
SSRC	Social Science Research Council (USA)
TI	Transparency International
UK	United Kingdom
UN	United Nations
UNDG	United Nations Development Group
UNDP	United Nations Development Programme
UNHCR	United Nations High Commissioner for Refugees
UNICEF	United Nations Children's Fund
USAID	United States Agency for International Development
USD	US dollar
UTO	United Tajik Opposition
VDF	Village Development Fund (Tajikistan)
VDPP	Village Development Planning Process (Tajikistan)
VO	Village Organization (Tajikistan)

Chapter 1

Methods and experiences of evaluating democracy support: a moving frontier

Chapter 1

Peter Burnell

Methods and experiences of evaluating democracy support: a moving frontier

In the early 21st century we live in an age of evaluation, performance indicators, league tables and the like. This can be said almost without regard to domain or kind of activity, country, or, indeed, organization or type of organization, whether governmental or non-governmental. Assessments of the state of democracy in different countries and comparisons of the same between democracies or in a single country at different points in time have now become commonplace. There is even a *Handbook on Democracy Assessment* from the International Institute for Democracy and Electoral Assistance (International IDEA 2002). But that is not all.

Attempting to assess the progress that democracy has made in a particular country or region or in the world as a whole is one thing; trying to estimate the bearing that international factors in general have had on that progress or lack of progress is completely different. Yet neither endeavour is the same as assessing the record of international democracy support or evaluating the performance of organizations for whom that activity features prominently among their activities, perhaps as their sole or main activity.

Democracy support is an international activity that involves an increasing number of institutions—indeed a growing number of different types of institution, some of them more specialized than others. On one side or the other they engage the majority of the world's countries in what in historical terms is a relatively new activity. Here the story parts company with what is known about the larger business of international development cooperation, or what is sometimes called development assistance or foreign aid—something that has evolved over many more years. Development assistance has lengthy experience of trying to assess the performance of development aid interventions. For a long time development economists have tussled with complex and at times seemingly insurmountable issues concerning how to evaluate the performance of aid. Many of the findings, in respect of both

evaluation methodology and the actual results from assessing aid, are freely available. (On methods see for example the Swedish International Development Cooperation Agency (Sida) evaluation manual, *Looking Back, Moving Forward* (Molund and Schill 2004).)

There is a clear contrast here with democracy support. For, as Sweden's new government-funded Swedish Agency for Development Evaluation (SADEV) says, knowledge of the results of efforts to strengthen democracy and human rights in other countries is 'limited'. Even in the United States research on the impact of democracy support and the conditions under which it can be most effective has been said to be 'lagging behind' the increased funding commitments to democracy support (United States Agency for International Development 2005: 3). SADEV's own interest in democracy aims to develop methods for improved planning, follow-up and evaluation of democracy programmes. Moreover, this ambition to devise more rigorous ways of doing evaluations is not confined to Sweden, let alone SADEV. On the contrary, by 2007 similar aspirations were being expressed on a more global stage in relation to the national, multinational and multilateral endeavours that make up international democracy support. Notable examples include a workshop on Measuring the Impact of Democracy and Governance Assistance, co-sponsored by the United States Agency for International Development (USAID) and the Netherlands' Clingendael Institute in The Hague, March 2005 (the discussions were reported in United States Agency for International Development 2005, and Green and Kohl 2007). Then there was a follow-up workshop entitled Methods and Experiences of Evaluating Democracy Support, sponsored by IDEA and Sida and held in Stockholm in April 2006 (which provides the basis for this book). Representatives from democracy support agencies and academia participated in both events.

This book is a tribute to the ambition and aspiration to improve the way in which we go about evaluating democracy support. This opening chapter helps sets the scene. It broaches such questions as why evaluate, what should be evaluated and how? It notes some recent efforts to get to grips with the methodological challenges of doing evaluations in this context. And above all it introduces some recent and experimental contributions to the debate, which form the heart of this book, before finally concluding with some remarks on the direction that democracy support assessments might take in the future.

What is evaluation?

Evaluation has been defined as the system for and objective assessment of an ongoing or completed project, programme or policy, and its design, implementation and results. The aim is to determine how relevant the objectives have been and how far they have been fulfilled, and to assess the efficiency, effectiveness, impact and sustainability of programmes. Evaluation also refers to the process of determining the worth or significance of an activity, policy or programme (Molund and Schill 2004: 106).

Marginally different versions cluster around this definition, such as the more extended and more normative idea that evaluation goes beyond just assessing results or providing a performance measure: 'Evaluations seek explanations, and account for why and how things happen, and also arrive at value statements' (Forss 2002: 3). A striking example in this book is Mirimanova's account of conflict-prone Tajikistan (chapter 8). There, insofar as external democracy support can be judged to have been successful, this seems to have been largely because local people viewed democratic progress as instrumental to gaining greater access to international economic support, vital for the reconstruction of the physical infrastructure. And yet even there inequalities arising from the distribution of the new resources, which reflect the pre-existing power imbalances, appear to threaten to reopen conflict. They could make one of democracy's main defining values—political equality—that much harder to achieve. This could affect how we value the overall impact. Clearly, then, evaluation can be a complex business. But it is no less clear that there are good reasons for trying to evaluate.

Why evaluate?

Evaluations can serve different purposes. The choice of purpose can influence the design of the methodology, who carries out the evaluation and the spirit in which it is conducted, as well as what happens to the results. Conversely, the purpose will depend in part on the nature of the organization and on who calls for or commissions the study. The source of the demand can lie within the organization or come from outside. The political pressure to evaluate democracy support coming from the funding side is said to be driving a good deal of the present interest. The main reasons for doing evaluations (apart from the fact that it may be legal requirement) are, however, reasonably well known.

   The reasons are: to ensure proper bookkeeping, that is, accountancy-type audits; to serve the aim of achieving efficiency or value for money; to facilitate accountability to the political masters and taxpayers who sanction or authorize democracy support in their name, which is a case that looks incontrovertible for any organization that claims to stand for democratic principles; to enable lessons to be learned from experience and make improved and more effective practice possible; and to offer a form of security against the kind of ill-judged 'political meddling' that displaces the formal goals of an organization or takes decisions on operational details away from the hands of able and experienced practitioners. Finally, evaluations can be called to inquire into very specific qualities such as establishing the level of gender awareness and gender sensitivity (or conversely bias) exhibited by projects or programmes such as those in democracy support. Evaluating their environmental credentials is a comparable illustration taken from the world of more conventional initiatives in economic aid.

   All these reasons apply as well in the case of international development cooperation, where there is a long history of commitment to evaluation. It is sometimes said that

governmental bodies are more aware of their accountability, which tends to make them risk-averse and reluctant to innovate, in contrast to non-governmental and private or autonomous organizations that tend to be more persuaded by the learning effect. However, this dichotomy is much too crude. All organizations should aspire to do better. Evidence both from the chapters in this book and from evaluations that have been reported elsewhere suggests that almost all democracy support organizations, irrespective of their status or source of funds, experience difficulty in transferring the benefits of knowledge gleaned from evaluations into policy and strategic review. In any case the most prominent autonomous actors in the field of democracy support derive the greater part of their income from official sources. In some cases this comes in the form of an annual grant and in other cases their activities are paid for on a contractual basis, perhaps after a process of competitive tendering. Either way they may be required to submit themselves to formal independent evaluation from time to time as a condition for renewal of the financial support. In these situations the evaluators may report direct to the funding body. In the case of the United Nations Development Programme (UNDP) Evaluation Office, which is neither a government nor a private body but a multilateral intergovernmental organization, Cole et al. (Danish Ministry of Foreign Affairs 2006: 32) found that much greater emphasis was being placed by all concerned on trying to aim directly for the learning benefits of evaluation. By comparison too little attention was being paid to accountability, that is to say 'the systematic assessment of both *expected and achieved* development results, the impact of assistance and the performance of the parties involved' (emphasis in the original).

Evaluation and participation

One further and rather special reason for evaluating democracy support is to use it as an exercise in exemplifying and transferring democratic values, or the principles that democracy purports to stand for and represent. The act of evaluation itself becomes an exercise in democracy support, in addition to whatever purpose it might have for improving the support activities that are under evaluation. This claim resembles the kind of thinking that informs IDEA's approach to democracy assessment, namely the option for the citizens of virtually any society to make their own self-assessment (International IDEA 2002). In turn this means a participatory approach to evaluation, which Crawford (2003b) among others has made a strong case for. Madsen (in chapter 5) also inclines towards this approach in her account of 'process rights' and 'process evaluation'—the desirability of making democracy support not just participatory but accountable and non-discriminatory as well.

Indeed evaluation procedures can themselves be evaluated and compared for their participatory content. Take for instance programme theory evaluation (PTE), the subject of chapter 3, which does not lend itself so readily to a broadly based participatory approach. A broadlybased approach would at a minimum include all

would-be beneficiaries, although it is easy to see how higher echelons in the partner organizations might be brought into the process of doing a PTE. Even so, one of PTE's more attractive features in some eyes might be that it seems to throw the onus of responsibility for shortcomings in democracy support on to those who design and initiate the programmes, rather than the partners, especially those in foreign countries whose task is to help implement the programmes in the field. Ironically, the institution that might need to change its thinking and behaviour in the light of the findings from PTE becomes the supporting agency, such as Sida, rather than, or as well as, the targeted actors and would-be beneficiaries in the field.

Of course participatory evaluation is not a new idea: it is well established in international development (see Molund and Schill 2004: 19–20). And it is important to debate who exactly should participate. The notion of stakeholders, or agencies, organizations, groups or individuals who have a direct or indirect interest in the intervention or its evaluation, is relevant here. Thus for instance Sida's evaluation manual makes a distinction between participatory evaluations and participatory evaluation methods. Further distinctions that might be made are between primary stakeholders, that is to say target groups who benefit from an intervention together with those who may have been adversely affected, and all those persons who feel they should have been included as beneficiaries but are excluded from its effects. In Sida's view, the best way to promote participatory evaluation is to strengthen the element of participation in the preceding stages of the intervention process (Molund and Schill 2004: 20). That means when the project or programme goals, and the criteria by which performance will be judged, are being determined. There will then be consequences for the kind of evaluation questions that it is sensible to ask.

However, whatever specific choices are made over how to operationalize the participatory ethos, at one level the general reasoning remains the same. If participatory approaches are deemed intrinsically desirable in international development cooperation (that is, development by definition must be a participatory process, as well as participation being functional for information-gathering and attaining a project or programme's other objectives), then the political argument must be even more compelling in regard to efforts to support democratization. In reality the evaluations do tell us that the quality of relations between democracy supporters and their partners in the field can be critical to both impact and effectiveness. They inform us that establishing the right time to curtail support is crucial: too soon and the partners may not be able to continue the work; too late and a culture of dependence with all its failings can easily take hold. Both sides have responsibility for getting relations right; using evaluations to heap either credit or blame on just one side would be irresponsible.

There are objections to and reservations about participatory approaches too. These have been well summarized elsewhere (Green and Kohl 2007: 154–6). And participatory evaluation is probably more talked about than practised even among organizations that are predisposed in principle to be sympathetic to the idea.

Nevertheless, there are some examples of limited forms of participatory evaluation being practised in the field of democracy support: for instance, the Clingendael Institute included local partners in the data collection and analysis when assessing democracy assistance in a number of post-conflict societies (see de Zeeuw and Kumar 2006). At the same time the enthusiasts also recognize that building local capacity in doing evaluations can be resource-intensive, and that a long-term view should be taken of the possible benefits. For all the resources at its disposal it seems that even the European Commission has yet to take seriously investment in domestic systems for monitoring and evaluating progress in regard to democracy and governance and related assistance. USAID's efforts offer an interesting comparison, for, although the USAID surveys of local people's attitudes towards and perceptions of the state of democracy in their country do not amount to consulting them on democracy assistance and its efficacy, it seems that these surveys are increasingly informing the indicators that USAID uses to measure the progress of its democracy programmes in these countries (see chapter 2).

Evaluation and avoiding failure

A final reason for evaluating democracy support—one that is so obvious that it is rarely mentioned—is to try to avoid the harm that can be done by mounting support activities that are badly advised or go horribly wrong. Waste of financial resources is not the issue here. Instead it is the consequences for the hopes, freedoms, and sometimes the very lives of peoples who look to international support, only to feel let down or misled, or, in the most damning scenarios, find themselves victimized by their oppressors. These last are the power-holders who react to ill-thought-out democracy interventions by taking reprisals against fellow citizens, in particular against people who have voiced their support for democratic reform and cooperated with foreign support actors. Whether they are simply emboldened or conversely feel more threatened by unsuccessful attempts at external democracy support, the negative consequences of such retaliation may be the same. Where democracy support must share the responsibility for bringing about a collapse of order or an increase in sub-state violence—inter-communal warfare for instance—then the harm done could be even greater.

Even where on balance democracy does register an advance, there may be human casualties or at a minimum some costs of adjustment experienced along the way. These are reasons enough for aiming to get democracy support right, which in turn means trying to establish sound methods of evaluation. There are no grounds for believing that the maxim 'do no harm', expressed so often in regard to international humanitarian assistance, does not apply equally seriously to democracy and human rights support—or, indeed to the methods of evaluation as well. An unfortunate evaluation experience can sap morale and alienate partners; methods that display patronizing attitudes or convey elitist and exclusionary norms would be just as inappropriate.

Evaluation then is burdened with many high expectations. That places a heavy responsibility on those who would design the evaluation methodologies. There may well be tensions if not outright incompatibilities between the different rationales for doing evaluation. This makes the challenge that much more difficult. One of the lessons learned about democracy support (for example, from the inquiry into the Westminster Foundation for Democracy conducted by River Path Associates, 2005) is that organizations should avoid taking on too many responsibilities. They must not try to pursue more objectives than their limited resources will support; similarly, the guidelines given to evaluators should specify a clear and achievable sense of purpose.

Furthermore, no rounded consideration of the reasons for evaluating democracy support would be complete without at least some reference to the counter-case. There can be bad, mischievous, or irrelevant reasons for commissioning evaluations in any walk of life. Examples include a mere gesture or nod to fashion, or the intention of interposing delays in decision making. That an evaluation might be commissioned in the expectation that the findings will only obfuscate rather than give weight to an otherwise compelling case for change is at least a possibility. It would be surprising if some of the more politically motivated demands for more evaluation were not grounded in an interest in finding reasons to discontinue the activity. That could be due to a sense that democracy support undermines the pursuit of more prized national and other objectives. Or it might simply be part of the usual rough and tumble that accompanies competitive scrambles over taxation and public spending levels and resource allocation. Equally, where there is strong political pressure to show results from democracy support—and that means positive results—there is an incentive for the decision to evaluate to concentrate on a selection of activities or areas where the likelihood is that there will a good story to tell. At worst the incentive structure may be such as to give reason to 'massage' the findings accordingly, or release them on a selective basis only.

The timing of evaluations and the duration selected for evidence-gathering can be important as well. Too early and there is a possibility that some of the effects of democracy support will not be registered. The finding by Finkel et al. (2006) that the positive effects of USAID support for democracy and governance included a lagged dimension and tend to be cumulative are worth noting here. Too late and the opportunity to learn from experience and make relevant improvements could be lost. Evaluations can inhibit experimentation: the perceived risks that failure might involve may weigh too heavily as a result. There can be other adverse effects if they are handled clumsily. For instance there is the possibility that evaluation will sow discord precisely where mutual trust and a shared sense of endeavour are most needed—among all the partners to democracy support. Bossuyt et al. (2006) found that their efforts to establish a general picture of European Commission support to governance even aroused suspicions among the European Union (EU) country delegations in the chosen case study countries, not to mention the governments of those countries. And yet commenting on the performance of those delegations and governments was never the purpose of the exercise.

Evaluations can be done in-house, embedded in the organizational culture even, or alternatively delegated to an external contractor: in either case the way in which the terms of reference are worded is likely to be critical to the exercise or its outcome. In the second case the virtue of impartiality or objectivity would seem to be assured, but it might be easier to dismiss unwelcome findings on the grounds that the outsiders were poorly informed or not sufficiently in tune with the organization's strengths. At the same time professional evaluators from the outside might claim to concentrate expertise and experience far more than a democracy actor by itself can manage. In contrast, in-house evaluations will only divert personnel from getting on with doing the job they are employed to do—initiate democracy support. Nevertheless, confirming shared ownership of the evaluation findings and their policy implications could be problematic, for various reasons. So in the world of evaluation there is no single procedural model. In Sida's case, for instance, the evaluation department has a semi-autonomous position within the organization. It began investigating how far Sida's support for democracy and human rights could be evaluated only as recently as 1997.

What can be evaluated?

Before going any further it is useful to make a distinction between the evaluation of democracy promotion in terms of its own democratic objectives (intrinsic evaluation) and in terms of extrinsic evaluation. The latter is concerned with how far democracy promotion serves the various policy rationales, drivers or motivations that underlie the foreign policy decision to support democracy abroad. For example, democracy support can be assessed by how far its achievements really do serve to bring about a more peaceful world, or help combat the threat of international terrorism, or facilitate economic and social development in the developing world. These extrinsic yardsticks, while enormously important and perhaps increasingly relevant as the securitization of democracy promotion and democracy support policy gains ground in some policy quarters, are not the subject of this book. However, we come back to some of the implications for evaluation at the end of the chapter.

Similarly, democracy support and its effects on democratization could also be examined for their consequences for other notable aspects of the political condition within the countries that are selected for support. And these may affect those countries' political development more generally, and the international order too, either directly or indirectly. This relates to such large and important projects as nation-building and state-building. Again, these kinds of externalities, whether favourable or unfavourable, are not the main focus here, with the important exception of Natalia Mirimanova's chapter (chapter 8) in this book on the contribution made by democracy support to conflict mitigation in Tajikistan. Given the high degree of interdependence that can exist between different processes of political change, the point should be made that it may be not just unrealistic but also undesirable to adopt a very narrow understanding

and assessment of the effects of international democracy support.

In terms of the intrinsic evaluation of democracy promotion, then, almost anything can be evaluated. That includes evaluations; the methods used to do evaluations; the process of evaluation; and the evaluators themselves. Assessing the quality of evaluations and devising an assessment framework for this purpose are thought to be a fairly recent innovation in the multilateral and bilateral development agencies (Danish Ministry of Foreign Affairs 2006: 40). Denmark's Ministry of Foreign Affairs conducted a peer assessment of evaluation in multilateral organizations, specifically the UNDP's Central Evaluation Office, in December 2005 (Danish Ministry of Foreign Affairs 2006). It is this office which has the responsibility for evaluating the UNDP's large and growing democratic governance programme, currently valued at around 1.4 billion US dollars (USD) and extending to over 130 countries.

The Danish report found that the Evaluation Office 'enjoys an acceptable level of independence and which produces evaluations that are credible, valid and useful for learning and strategy formation in the organization. At the same time, its potential for helping strengthen accountability and performance assessment is being underexploited, both for the purpose of accountability and as an essential basis for learning' (Danish Ministry of Foreign Affairs 2006: 4). It will be interesting to see whether the Evaluation Office will warrant comments like these when it turns its attention to the UNDP's support for democratic governance. What future evaluations of the office might tell us about the demands of evaluating democratic governance programmes as compared with evaluating other more traditional forms of UNDP development cooperation will definitely be worthy of interest. The Danish report's finding that little has been done up to now to ensure the involvement and ownership of partner country stakeholders in the evaluation process should also be revisited, once the UNDP Evaluation Office proceeds to investigate the UNDP's democratic governance programmes.

In international democracy support there are a number of potential candidates for evaluation. As of now some are more in the evaluators' sights than others. And it should be noted that, while an individual or stand-alone evaluation can be very informative, assessments that are done on a comparative basis are potentially more revealing, irrespective of the basic unit of analysis such as a project or country that the assessment takes.

In the evaluators' sights

First, there are individual projects that occur in one country and address a specific sector or sub-sector, such as human rights non-governmental organizations (NGOs). The candidates may be selected on a more or less random basis or according to such criteria as their symbolic significance, for example, early flagship projects that have the longest history of operation, or by size.

Second, evaluation can focus on *programmes*, which include all projects in a

particular sector or sub-sector either in one country or in several countries. Kumar's (2006) comparative assessment of support for independent media in several countries is a recent example from the literature, although its coverage is of US support only and it does not claim to be comprehensive even then. The German Technical Cooperation Agency (Gesellschaft für Technische Zusammenarbeit, GTZ) now pursues a middle way between evaluating projects and overall outcomes by focusing on intermediate objectives, for example to improve the rule of law. Carothers (2006c) offers an extended compendium of collected experience in reviewing rule-of-law promotion.

Third, the principle of selection can be a *partner country or countries*, selected perhaps because they are leading partners or for special historical or political reasons, the aim then being to cover the full range of projects and programmes mounted in that country or countries. The attempts by Michael McFaul and others to examine the contribution that external support made to the origins of the so-called 'Orange Revolution' in Ukraine in December 2004 are illustrative (see, e.g., Åslund and McFaul 2006). Among other things they demonstrate how difficult it is to pin down the influence of external factors, for it is mainly in the interaction with internal forces and local actors that political outcomes come to be determined.

Fourth, an *institution* responsible for democracy support can be evaluated *as a whole*, with evidence drawn perhaps selectively from the entire range of its activities or country involvements. Recent examples include the evaluations of the Netherlands Institute for Multiparty Democracy (NIMD) (European Centre for Development Policy Management 2005) and the Westminster Foundation for Democracy in the United Kingdom (River Path Associates 2005). The *institutional partners* in countries where democracy initiatives are being supported may also be the subject of evaluation, and in this context the arguments surrounding participatory evaluation become especially salient. By implication, even if this is not specifically called for in the terms of reference, evaluations of democracy support organizations may include comment on the selection mechanism and actual choice of foreign partners. In principle that could extend to relations with collaborating support partners, where the provision of support is organized on a joint or collective basis.

Fifth, and more challenging and considered less often, is the possibility of evaluating the choice of *methods, approaches, or instruments* that are used to promote and defend democracy abroad. In practice that is most likely to mean democracy support—non-coercive and concessionary initiatives, otherwise known as democracy assistance or democracy aid. However, in principle it could include all the ways in which international activities are undertaken with democracy promotion as their primary objective. That includes diplomatic initiatives, foreign aid incentives, the use of trade, investment and other sanctions, covert intervention and even coercive threats or outright military involvement, in short, what is often called 'hard power' as well as 'soft power' techniques. Thus for instance Schraeder (2002) explores the 'spectrum of violence' in which a variety of 'interventionist tools' have been employed in democracy's name, drawing on a five-year joint European–North American research

project that was funded by the Finnish International Development Agency. The findings reported (Schraeder 2002) highlight the constraints on effective democracy promotion. But the time for this ambitious multi-country comparison to be repeated and brought up to date is now rapidly approaching, notwithstanding the recent attempt by Youngs (2006) to summarize European experience only.

Taken all together, then, the overall commitment displayed by just one government or intergovernmental actor, or by an international organization like the United Nations, the West or even the so-called international community, to promoting democracy by all means is an obvious candidate for assessment. Who could not be interested in establishing some plausible overall verdict, say a figure scored out of ten? Or perhaps two figures—one representing level of commitment, and another achievement, success or failure? Such an assessment could be framed in terms of the strategy for supporting democracy (see Burnell 2005; Piccone and Youngs 2006). A complex and challenging exercise of this nature could not be undertaken lightly. It is just the sort of area where research must draw on the work of many analysts and only after the research design has first come up with a conceptual and methodological framework that is adequate to the task (Burnell 2007/8). That is still some way off.

Finally, an even grander extension would encompass not just active democracy promotion, where intentionality and sense of purpose are among the defining properties, but also what might be called passive democracy promotion. That refers to all the ways in which external actors generally, or even just the governments of established democracies in particular, impact on democracy and democratization inside the prospective, new and emerging democracies, for good or for ill, regardless of whether the likely consequences for democracy were intended, considered or desired (Burnell 2006).

From among this ascending list of candidates for evaluation, it is on the first-mentioned and least ambitious candidates that the majority of actual attempts to evaluate have tended to concentrate, at least in the democracy support organizations themselves. The more grandiose and most challenging possibilities have been left to academia to muse on and grapple with, not least because the subject matter is so highly political, although even here until recently the literature has been remarkably thin. A notable exception is the European Union's use of conditionality to promote human rights and democratic reform in states from Central and Eastern Europe seeking accession to full membership of the EU, in accordance with the Copenhagen criteria of 1993. A growing number of detailed studies by scholars in both Europe and North America have homed in on this (Kelley 2004 and Vachudova 2005 are two outstanding examples). By and large they agree that the EU's experience to date has been remarkably successful, while giving rather varied accounts of the main reasons why. It is very unlikely that this record will be repeated in the future, once EU enlargement has come to an end. In the meantime, as the EU itself seeks to reinvent its post-enlargement strategy for promoting democracy abroad, approaches to evaluating EU efforts in the years to come will have to do more than copy even the relatively well developed lines of inquiry that have contributed enlightenment in the recent past.

Results-based and programme theory evaluations

In the world of democracy support, projects and programmes tend to be assessed for their effectiveness, which involves identifying both the outputs and the proximate outcomes, and, rather more ambitiously, their impact. Both are results-based measures.

Effectiveness refers to the extent to which support for democracy achieves its own goals and objectives. Clearly, careful attention to the way in which these are specified, that is to say clear and precise specification at the planning and design stage, is crucial if performance is to be monitored later and for the results to be assessed. Surprisingly, project goals have not always been formulated in ways that allow evaluation (Swedish International Development Cooperation Agency 2000: 3). This would seem to provide an area where improvement should easily be possible, even though in some cases there may be good political reasons why goals and objectives are ambiguous, or not all are stated, or they are sometimes left a little fuzzy.

Impacts may be experienced both directly and indirectly; they can be negative as well as positive, and either intended or unintended. In democracy support, impact assessment tends to mean the wider consequences for democracy and democratization, including those that might emerge only in the medium to longer term. Of course it also includes those which emerge sooner but whose sustainability should be the main point of interest. Several evaluation studies note that impact is more significant than effectiveness and call for more thought to be given to ways of assessing impact. At the same time the difficulties, such as problems over attribution, have been well rehearsed (see, e.g., United States Agency for International Development 2005: 11–14). To illustrate, a Sida inquiry (Swedish International Development Cooperation Agency 2000: 3) into the evaluability of democracy and human rights projects found that for most of the projects examined it would be difficult to evaluate impact by means of a goal-oriented approach based on logframe analysis. This is in spite of the fact that the point of logframe is to specify goals, purpose, outputs and activities in ways that enable results to be identified at every level. There are other limiting factors too, because eliminating the effects of 'noise in the system' from the analysis, in other words the influence of all other factors, is deeply problematic.

While both impact and effectiveness dwell on results, a rather different if still quite experimental approach to evaluation, explained and examined in this book by Fredrik Uggla (chapter 3), inquires first into the assumptions and the reasoning—the internal 'logics'—that inform programmes of democracy support. Programme theory evaluation may not capture consequences in the field, let alone identify and help solve problems that arise in the course of implementation. Nevertheless, it could feed into policy on democracy support at the policy planning and policy making stage of the decision cycle, in other words policy appraisal, especially if used in conjunction with findings from more results-based assessments. After all, it is only some external reference point that can tell us whether the theory's internal assumptions are truly

realistic, once any obvious incoherence or contradictions among the assumptions have been stripped away.

Evaluation: lessons of experience

There is a substantial history of evaluation in the field of international development cooperation. Democracy support should not have been obliged to reinvent the wheel in all respects, but there has been a fairly shallow learning curve in respect of identifying and addressing the problems nonetheless—practitioners have yet to come to terms fully both with some methodological difficulties that are familiar from the evaluation of development assistance and with the difficulties that resonate more sharply or have peculiar resonance in respect of democracy support.

However, while evaluation is supposed to tell us something about democracy support, there is also some merit in reversing the question and asking what have the evaluations of democracy support and their findings told us about evaluation methodology, in particular any weaknesses or shortcomings? To illustrate the point, Forss (2002: 49) reports that the large Danish International Development Agency (Danida) evaluation of Danish support for human rights and democratization in 1990–8 raised 'fundamental and challenging issues, notably concerning methodology and impact evaluation', 'raising more questions than answers'. The lessons from evaluation for democracy support and the lessons about evaluation are analytically distinct, although in the practice of democracy support and its assessment the two should be considered inseparable.

Prominent issues revealed by attempts to evaluate democracy support to date can be briefly summarized in terms of the 'what', that is, the object of democracy support; the 'how', that is, how to collect evidence and interpret its significance as well as decisions on what actually counts as evidence; and on when to make the investigations.

Measuring democratic progress

In regard to the 'what', the general object of democracy support seems less easy to define than is, say, the object of international development assistance. There, not only is economic growth an idea that is reasonably clear, precise and, most significant of all, quantifiable, but there are widely understood notions of what is meant by economic development, social development and human development. At least there are some internationally sourced and recognized indicators for measuring economic growth. In comparison, democracy is an 'essentially contested concept' (Gallie 1956). Democratization is an even more blurry idea: attempts to distinguish democratic transition, democratic consolidation, sustainable democracy and so on do not settle matters but often serve only to add further confusion. The same is true of such normative distinctions as those between liberal and electoral democracy, market democracy and

social democracy, and the like. Contested discourse surrounds the exact relationship of democratization to human rights and to the rule of law in particular, although few commentators would dispute Madsen's claim (in chapter 5) that human rights supply a fundamental pillar. Beyond even the matter of definitions and their relationship to one another, however, lies an even bigger, more contested and wide-ranging debate over how to explain democratic change. Efforts to understand what makes it happen, how and why it happens, and what prevents or reverses the process have generated a vast literature that sends many different and in some cases conflicting signals to those actors who would wish to make a difference in practice.

For the purposes of evaluation, capturing the democracy effects of democracy support in the form of a meaningful, usable and perhaps above all agreed set of indictors is a vexed issue. Take for example just one of the more widely used sets of indicators, the Freedom House annual country ratings for political and civil liberty. These are often used as proxy for the level of democracy. Notwithstanding their convenience and popularity (the one helps explain the other), a number of academic analysts and more policy-focused commentators have voiced serious reservations about the methodology. See for example the entry for 'Freedom House Annual Survey of Freedom' contained in the UNDP Oslo Centre's Governance Indicators: A User's Guide (United Nations Development Programme, no date).

All this complicates the business of impact assessment. There, the setting of benchmarks for democratic progress can fall foul of disagreements not just over the meaning of democracy and democratization but also over what it is appropriate to expect in the context of the particular circumstances of the country concerned. For instance, what consideration should be given to the amount of resistance to change put up by the people in power? This is important as some evaluations purport to find that government ownership of political reforms is a major influence on the success of external efforts to provide support. Bossuyt et al. (2006), for instance, both make this claim themselves and refer to similar findings from other European Commission studies. And should evaluators take account of the resistance that is due to the suspicions, fears and uncertainties about change among ordinary people, perhaps the majority of society, or the economic conditions and whether there has been a previous history of failed attempts to introduce democracy? What might seem like a rather modest advance for democracy in one country could represent a giant leap forward in another country where the situation had initially looked much less promising.

In comparison with democracy, we might think that human rights and also some features of 'good governance' (and thereby 'democratic governance' too) would be more amenable to definition. After all, certain rights are named and spelled out in the form of national or international declarations, bills or conventions of the kind that most governments have ratified or signed. Governance touches on variables that in some cases are more of an administrative and managerial than a value-laden or political nature. But of course the objectives of projects to support human rights and governance might not bear out these properties in practice. It seems that in reality the

indicators for human rights remain confused, or consensus on what they should be remains elusive (see chapter 5).

The International IDEA Handbook on Democracy Assessment (International IDEA 2002) might be considered a breakthrough in terms of offering a universal template by which to assess progress towards democracy. Yet, although the model has now been applied in several countries, there are no examples of attempts to measure the performance of democracy support against the yardsticks that the template provides. As Hanne Lund Madsen pointed out at the IDEA/Sida Stockholm workshop, it would be interesting to consider how the democracy assessment framework could help structure the evaluation of democracy assistance. A very useful exercise would be to pilot just such a study. However, even that would not immediately resolve some of the outstanding methodological difficulties. The overriding issues here concern first the rival claims of qualitative and quantitative approaches, and, second, problems to do with attribution and the assignment of 'effects' to specific causes or democracy support interventions.

Quantitative and qualitative methods

A substantial body of qualitative evidence about democracy support drawn from interviews with stakeholders, consultation of documentary sources, case studies, and the evaluators' own observations already exists, and much of it is publicly available, some of it in printed and some in electronic form. Thomas Carothers, who has been researching the United States' democracy support in particular (although not exclusively), has been in the forefront here, with eight books and many more articles to his credit (e.g. Carothers 1996, 1999, 2004, 2006b, 2006c; Carothers and Ottaway 2005; Ottaway and Carothers 2000). He has been called the world's leading authority on democracy promotion. His assessments have been largely critical of the way democracy support has been pursued, although not so damning as to lead him to believe the activity is fatally flawed. His well-informed and well-judged advice is highly sought after within the democracy support community.

The findings of several other writers, including quite a few from European countries, can be found in a fairly narrow range of academic journals, of which Democratization is the single largest source (a small selection is Blair 2000 and 2004, on support for decentralization and civil society respectively; and Scott and Steele 2005, on the United States National Endowment for Democracy) and in collected volumes (for example Erdmann 2006 on Germany's Stiftungen and their help to political parties in particular). Official and other formal reports provide another and growing source of qualitative assessments (examples are River Path Associates 2005, on the Westminster Foundation for Democracy; European Centre for Development Policy Management 2005, on the Netherlands Institute for Multiparty Democracy; and Bossuyt et al. 2006, on European Commission support for good governance in third countries). The last spilled over into democracy and human rights issues. This

illustrates the difficulties of drawing tight lines around the meaning of democracy, and shows that even evaluations without democracy in the title or terms of reference may end up telling us something interesting about democracy support. Bossuyt et al. (2006: 25–8) in fact employed a combination of methods—archival sources, interviews, case studies, and focus groups and large-scale questionnaires too. The report also helpfully contains short descriptions of the weaknesses of each method. For instance, the circulation of staff inside democracy promotion organizations and transfers outside will impinge on the available institutional memory, and, as United States Agency for International Development (2005) also notes, this may impede good data recovery.

In contrast, USAID has led the way in respect of the aggregate assessment of its entire support programme for democracy and governance in 195 countries, of which over 120 were actual 'recipients' (see chapter 2 by Margaret J. Sarles). This exercise subjected all programmes from 1990 through to 2003 to independent quantitative evaluation (Finkel et al. 2006). This USAID evaluation, which uses both Freedom House and Polity IV data to measure democratic progress, leads us to believe that large quantitative evaluations are technically feasible so long as adequate data can be found and presented in a form that is related to the outputs, outcomes and impact of democracy support (cf. the situation Cole et al. say they found in the UNDP (Danish Ministry of Foreign Affairs 2006: 37)). As chapter 2 reveals, just assembling the data may be no mean feat in itself.

The USAID study also claims to find that democracy assistance can be—that is to say has been—positive, although the overall impact has been blunted by the meagreness of the resources hitherto put at its disposal. Interestingly, the effect was found to be negative for human rights, which the report speculates might be due to a positive influence that support had on the availability and publishing of information about human rights abuses. This illustrates well that it is not just the figures but how we try to make sense of them that is so important. Such studies as the USAID one also serve to caution us against rushing to quick and simple inferences, for example, the idea that there might be a guaranteed democracy dividend for each and every country where democracy support is increased by a given amount. Obviously such a leap cannot be made on the basis of Finkel et al.'s (2006) findings. Endorsing an observation freely made by Sarles, that correlations do not themselves amount to a full explanation, the really vital issue for democracy support actors is the question why democracy support expenditures might have such effects and under what conditions. No matter how technically sophisticated the methodology, the way in which we interpret the findings and moreover the political uses to which they are put (or whether they are ignored) are what make all the difference—between a constructive experience of evaluation and one whose results prove to be irrelevant or damaging to the activity of democracy support.

Assigning consequences

The problem of assigning effects is magnified the further the chosen object departs from a tightly defined project or programme and takes in sector-wide and multi-sector support initiatives as well. The problem increases as evaluators attempt to aggregate the effects of multiple initiatives or, conversely, try to disentangle the effects that can be assigned to just one intervention from a context where several democracy support actors and multiple, perhaps mutually reinforcing (or cancelling), initiatives have been involved. For instance, the possibility of there being unintended cross-sectoral influences at work cannot be excluded. To illustrate, it has sometimes been suggested that external support to civil society has been detrimental to political parties and the development of a competitive party system. It threatens to have this effect by drawing away able leaders and resources, and encouraging the formation of a false idea of 'civil society good, political parties bad'. Moving away from trying to demonstrate effectiveness to establishing impact compounds the difficulties enormously, most notably when trying to move from propositions about the micro level to meso and macro effects.

In terms of 'when', the right moment to look for evidence of what democracy support has or has not achieved may not be self-evident. The timing or duration could well differ as between different projects, programmes, sectors and countries. This only complicates the business of trying to make up some aggregate picture for the purpose of reaching some general inferences about support as a whole. The longer the gestation period or the wait, the harder it could be to reconstruct the data and the greater the chance that any lessons learned will soon be out of date. Conversely, evaluations that are 'quick' (if not necessarily 'dirty') may be unable to capture the full picture and carry the real risk of distorting the portfolio of democracy support in the direction of activities that are believed to hold the most promise of producing early (favourable) results. By general consent in most cases democracy-building is a long-term project and friends in the international community must expect to have to make an appropriately long-term commitment.

Finally, there is the old refrain that you cannot know the counterfactual. A truly convincing verdict on the success or otherwise of democracy support must have good reasons for saying what would have happened in the absence of democracy support. Thus Finkel et al. (2006) resort to modelling the 'normal' projected trend in democratization as a way of trying to establish the difference USAID democracy support made. It employs specific statistical techniques to tackle the endogeneity issue, which is a term that refers to the possibility of reverse causality, in other words where democratic progress or its absence are responsible for pulling in the democracy support.

Just as with (for instance) the issue of quantitative versus qualitative methods, however, the counterfactual problem is not unique to evaluating democracy support. The same can be said in respect of the principles that underlie case selection in

comparative analysis, that is whether to choose the most different or the most similar cases to compare. And, as with the other issues raised, here it is unreasonable to expect inquiries into democracy support to solve problems and settle disputes over methods that have bedevilled social science for generations. It is equally unreasonable to counsel against evaluating democracy support on the grounds that current evaluation methods attract criticism or that the current refinements and all future developments are bound to be less than perfect.

A look forward

This is not the place to foreshadow in detail the contents of the chapters that follow. Rather, just some of the more notable features of the collection are introduced here, against a background of drawing attention to the implications for the future of democracy support evaluation.

Introducing the chapters

First, the methods introduced in the chapters all share an experimental quality, but one that is grounded in the authors' own practical experience and does not just reflect 'ivory tower' thinking. The majority embrace a commitment to qualitative methods, but a high-profile attempt at more quantitative assessment is also strongly represented. It seems incontrovertible that some features of democracy programmes are more susceptible to meaningful quantification and related forms of assessment than are others, just as some forms of democracy support may well be more amenable to evaluation than others. However, knowing how to integrate in a meaningful way the findings from the best of both quantitative and qualitative approaches not only remains one of the most intriguing conundrums; it is also one of the most worthwhile objectives for evaluators to aim at. There is more to be gained by trying to address this issue constructively than by portraying quantitative and qualitative approaches as rivals, a battlefield where analysts feel compelled to take one side and decry the other. In reality there are alternative ways of approaching evaluation even within both the quantitative and the qualitative approaches.

Even so, it is worth emphasizing that decision making on what pro-democracy activities to support, and how, should not be determined purely by which ones can be subjected to statistical measurements of performance afterwards. Even their evaluability more generally should not be the sole determining or overriding criterion. Just as evaluators should employ methods that can capture the unintended (and possibly negative) consequences of democracy support activities, so they should be alert to the possibility of unintended and even undesirable consequences of being too zealous about evaluations and of the evaluation methodology choices they make or approaches they take. Thus there is something to be said for the eclecticism that Sandra Elena and Hector Chayer argue for in chapter 4.

A second shared characteristic of the chapters is that they contain critical commentary expressed either directly or by implication on weaknesses in the state of the art of evaluating democracy support. In practice evaluation itself may not be an entirely happy experience: it can lead to some uncomfortable findings; organizations have to learn to live with the possibility that criticism of themselves or their modus operandi could follow. At the same time none of the authors is shy of noting the limits to the alternative proposals they are putting forward, and all recognize the need to develop their own ideas further.

Third, although evaluation is a common reference point, the chapters offer some quite diverse perspectives in respect of which aspect of democracy support they focus on. None are concerned with accountancy-style audits or purely financial—that is to say, 'value-for-money'—types of evaluation. But they do range over different sectors of democracy assistance, such as rule-of-law assistance and civil society aid. Inevitably, in a book of this size the coverage cannot be comprehensive.

Thus, some important components of democracy support are not highlighted. Examples include the aim of strengthening political parties and competitive party systems in emerging democracies, which is an area that has recently started to attract more attention (see Carothers 2006b); support to electoral procedures (see Bjornlund 2004; Lean 2007); and governance assistance and, more specifically, the strengthening of legislatures. Attention to obtaining the right balance in civil–military relations and full civilian control of the entire panoply of military, paramilitary and intelligence agencies—the 'security community'—in states facing serious internal threats of political violence is an under-researched subject generally. These are all areas of democracy support where attention to evaluation is warranted, notwithstanding the excellent work that has already been done by Carothers and others. Civil–military relations in particular can easily fall between the two stools of international endeavours to promote democracy on the one side and national or international security on the other. Much scope remains for exploring improved ways both of coordinating the support efforts and of determining their efficacy.

However, while there are whole sectors of democracy support as well as types of individual project or programme within sectors that would benefit from greater attention (a more systematic and comprehensive checklist approach to sectoral evaluations would aim to do this), we should not lose sight of the bigger picture either. As is noted above, democracy support is but one element of a much more varied and wide-ranging set of approaches, instruments or tools all of which denote international democracy protection and promotion of one form or another. On the one hand, comparing democracy support against these other approaches—with which they are often used in tandem—would be as fruitful for future strategic thinking on the diffusion of democracy as comparing the effects of alternative sectoral programmes in democracy assistance. This could mean making a double shift, from the ex post evaluation of democracy support to the ex ante appraisal of democracy promotion *tout court* (Burnell 2007/8). On the other hand, the intellectual challenge this poses,

just like that of assessing the sum total of international influences on the prospects for advancing democracy, must lie outside the confines of these chapters.

A fourth point that can be made about the chapters as a set is that they offer different but complementary material in respect of the main launch point of their inquiry, and not only in terms of the particular institution or support sector(s) they concentrate on. This point merits some elaboration. Take first chapter 2 by Margaret J. Sarles, which sets out to describe USAID's not inconsiderable experience of trying to get to grips with the challenge of evaluating democracy support, culminating in the state of play at the time of writing—as represented in the report *Effects of US Foreign Assistance on Democracy Building: Results of a Cross-National Quantitative Study. Final Report* (Finkel et al. 2006). The ambitious nature of this attempt at evaluation and related initiatives—exceptional in terms of scale, breadth and universal coverage—make it highly appropriate that this account opens the set of substantive chapters. Contrasting with this account, then, is the chapter by Elena and Chayer (chapter 4), which tackles the question of how to establish outcomes and effects in regard to just one individual sector or field, namely rule-of-law programmes. Chapter 4 is based on the experience of an Argentinian NGO, the Forum for Studies on Judicial Administration (Foro de Estudios sobre la Administractión de Justicia, FORES), in its work in that single country, Argentina.

Writing also from an institutional perspective, Michael Wodzicki (chapter 6) inquires into the lessons to be learned by evaluating a process of evaluation already carried out into the performance of another organization, namely Canada's International Centre for Human Rights and Democratic Development (Rights & Democracy). This chapter relates the experience of this independent organization, which, following its establishment by act of Parliament in 1988, has utilized a human rights perspective on democracy and democratization to advance support to civil society abroad. This chapter explicitly broaches the issue of how to evaluate evaluation methodologies in action, which is an underlying theme running through the book as a whole.

Introducing the theme that makes human rights a major component of democracy, Hanne Lund Madsen (in chapter 5) examines the evaluation of human rights more generally but with specific reference to a rights-based approach to both programming and evaluation. This approach follows from the recommendation of the Office of the United Nations High Commissioner for Human Rights that a 'rights-based approach' to development integrate 'the norms, standards and principles of the international human rights system into the plans, policies and processes of programme development' (Office of the High Commissioner for Human Rights 2003: 1). Madsen's offer of a 'rights-based approach navigator', while it is a contribution to meeting this requirement, also shows how detailed the focus must be in order to explore the two sides—rights-holder and duty-bearer (or electorate and representation).

Meanwhile in regard to democracy support more generally Fredrik Uggla in chapter 3 shows how we can evaluate the assumptions and the reasoning that lie

behind the design of democracy support initiatives, and why this could enhance the art of evaluation. This innovative approach, which Sida is currently investigating, offers a sharp contrast to both the extensive quantitative modelling and the more qualitative fieldwork-based approach to investigation that is described in the case of USAID. And, although PTE might not uncover all the shortcomings that can arise during the course of programme implementation—something that Sarles' reference to USAID consulting 'voices from the field' seems to try to tap—it can certainly help prevent programmes from being adopted that are internally flawed or fundamentally misguided from the outset.

In chapter 7 Harry Blair takes another major sector of democracy support as his focus and asks what it means to establish the effects of civil society advocacy. This question and the novel way he goes about providing an answer can be seen as essential preliminaries to trying to establish how far international support to civil society organizations itself makes a contribution to democracy. Methodologies for the latter must be contingent on an adequate conceptualization of the former. Blair offers detailed case studies from India and the Philippines to illustrate his theoretical approach. The analysis draws out an important distinction between the consequences of advocacy for the civil society organizations and their objectives, on the one hand, and for the wider political system on the other. If the two do not coincide, as at times seems possible, then democracy support organizations are faced with some difficult practical choices in regard to their choice and treatment of foreign partners.

Blair's account also adopts a largely statist approach, that is to say the target of civil society advocacy is presented as government institutions. Further research could usefully address whether the analysis can be extended to the larger set of governance institutions. These are the institutions, some of them non-governmental and some of them inter- and supra-governmental, that observers writing from a globalization perspective argue are increasingly central to the complex of power relations that exists at local, national and global levels. In fact this is an issue not just for evaluation methodology but for democracy and human rights support and democracy promotion more generally. The challenges to democratization that lie beyond the nation state— the challenges that globalization poses to democracy inside countries and the challenge of democratizing the institutions of global governance—have barely registered on the radar of international democracy promotion (see Burnell (ed.) 2006, chapters 1, 2). Unsurprisingly, to date the evaluation discourse has had virtually nothing to offer on this.

In chapter 8 Natalia Mirimanova examines the consequences of community-level democracy support for conflict resolution in the interesting case of Tajikistan. The question how to evaluate is not the sole focus of this chapter. Nevertheless the study raises important issues for evaluation methodology in situations where there are multiple goals, such as peace, development and democracy. Implementing these and other goals of international involvement can encounter conflict between them as well as some synergies and relations of mutual support. How external support affects

democracy and the impact of democratization on conflict resolution are analytically separate issues. In practice they may be related in different ways—one can imagine a matrix of more democracy and less conflict; more democracy and more conflict; less democracy and more conflict; less democracy and less conflict; no change in one or either. Different conflicts can have different causes, and the potential for democracy support to make a difference may also vary from one situation to another. To weigh and summarize the consequences in a single composite verdict may be an impossible task, perhaps confirming that there are limits to what can reasonably be expected from evaluations and from the design of evaluation methods, no matter how refined.

In these situations different kinds of international organizations are involved on the ground and exercising their own distinctive mandates, quite legitimately pursuing their own agenda—peacemaking or peacekeeping, humanitarian relief, economic reconstruction, capacity-building in governance, fashioning democracy, and so on. Here the question of who should evaluate is more than usually tied up with the issues to do with the terms of reference and what should be the main criterion or object of evaluation. On this question, evaluation methodology itself cannot provide answers, but perhaps faces some of the biggest challenges of all. And, given that such situations, whether we call them conflict-prone, post-conflict, 'states under stress' or 'complex environments', are not uncommon and show no sign of disappearing, additional experimental work on the evaluation of international interventions in such circumstances merits a high priority.

That said, even in more stable and peaceful situations there is much to be gained by considering evaluation methodologies for democracy support in tandem with considering comparable methodologies for engaging in more 'normal' development work. That means the participatory strategies for development more generally, and developmental initiatives in female empowerment and promoting gender equity specifically. While the reality seems to be that, both in development discourse and in development practice, development objectives and democratization considerations are often treated separately, the potential gains to be achieved by more holistic thinking regarding support for development and support for democracy together—and so the implications for evaluating either one or both—cry out for serious investigation. In other words, development practitioners and democracy practitioners should talk to one another more often. And, if the evaluators in both cases are not the same people, then perhaps the different constituencies too should follow likewise. First, however, as Patrick Molutsi argues in the concluding chapter, there is still work to be done in bringing together evaluation frameworks for democracy support with assessment frameworks for the condition of democracy itself—and with input from local people being essential to both of these. No less pertinent, Molutsi urges that progress towards a global index for measuring the impact of democracy assistance should be the short-term objective for democracy support actors, which perhaps must take precedence over other, more all-embracing aspirations.

A fifth comment on the chapters as a whole concerns not what they contain

but something they do not do. While they all generate valuable insights into how not to evaluate and the failings that should be avoided, none of them considers whether some useful insights might also be gained by comparing cases where democracy support has been provided against cases where it has not been provided. Instances of the latter could be drawn both from countries where there has been democratic advance and from countries where there has not. Counties where democracy or the momentum to democratize has gone into reverse are also eligible. Reflecting on how to discover whether external support would have made a difference, and how, when, and where, offers yet another area of inquiry relevant to future research on methods and experiences of evaluating democracy support. At the present time, however, this seems like a tall order. In most, if not all, democracy support organizations there is still substantial progress to be made just in terms of organizing and implementing the routine collection of data about actual democracy support interventions, let alone trying to construct data about situations where few or no interventions were planned or have taken place.

Finally, while each chapter offers something different, and none pretends to introduce all the issues that might be raised, when taken all together they help form an overall picture that is quite distinct. For one thing it becomes clear that there is no standard model or single approach that is employed uniformly by democracy support agencies. Indeed, it is not obvious that democracy support organizations either collectively or individually are consistent in terms of whether they evaluate their activities at all, or in the methods employed. Moreover it may be both unrealistic and undesirable to be searching for a universal approach. The 'one size fits all' approach has been the butt of much criticism in debates about approaches to devising solutions for economic development and political reform alike. We should avoid the temptation to seek straightforward comparisons between what in essence may be not just different democracy support activities but intrinsically different approaches to 'doing democracy support'—offered by different kinds of organization and not simply different organizations, each one possessed of its own mandate and endowment of resources or instruments. They all have their own ideas about the main purpose and the outcome to be achieved.

That said, something else that the chapters all agree on is that all democracy and human rights support bodies should think carefully about the analytical framework and issues of methodology not just when designing the evaluations but much earlier in the process too, namely when deciding their support interventions in the first place. Otherwise, attempts to evaluate the support are bound to struggle right from the start. But, if that remark says something only about the early stages in the sequence, then what happens in subsequent stages cannot be ignored either.

What happens after evaluation?

It is evident from the chapters in this book that methods and approaches to evaluation in democracy support are still under construction; much scope remains for review,

reflection and further development: stimulating these is precisely the book's intention. But at the same time there is little point in pondering to excess all the details of evaluation methodology if we do not dwell also on how the lessons learned from evaluation can be taken up by the democracy support agencies and, even better, shared among them. In short, there is the issue of individual and collective institutional learning.

In fact one of the most frequently encountered findings of democracy support evaluations is that the institutional capacity for learning is defective even in organizations that show interest in conducting evaluations and are judged to perform well in other terms too. The NIMD, whose approach to democracy support was commended as innovative and a potential 'best-seller' (European Centre for Development Policy Management 2005), is an example. And more generally Wodzicki (chapter 6) offers some clues from Canada's Rights & Democracy as to why very little action may follow an evaluation. In part the problem seems to be a bureaucratic one. The organizational relationship of those who devise the evaluations and do the evaluations to the other parts of the machinery for democracy support could have a strong bearing on both evaluation design and how—indeed, whether—the findings are used. Use here does not mean the slavish adoption of an evaluation's recommendations, but it does mean a considered response to the findings at an appropriate level in the organization.

The issue of who gets to see the findings could itself be a matter of some contention. For instance, it was learned from the March 2005 Clingendael workshop that respondents to evaluation questionnaires administered by the GTZ are given the opportunity to define how they measure the success or failure of a project rather than feel constrained by questions that have been pre-formulated. This looks like a step in the direction of so-called 'fourth generation' evaluation methodology (the first three 'generations' being measurement, description and judgement). The 'fourth generation' applies a constructivist mindset. In the extreme case it may even call into question the idea of scientifically verifiable reality. At a minimum it appears to take stakeholder concerns into consideration. Yet at the same time it seems that the GTZ does not disclose the evaluation findings to overseas partners, even those whose support is likely to be cut as a result of a disappointing assessment.

The issue of follow-up to evaluations is also bound up in part with political control. How interested are the political decision makers who ultimately determine the policy commitment to democracy support and sanction the resources to do evaluations? Are they sufficiently interested in the research methodologies to want to take advantage of the benefits that sound evaluation procedures and practice might offer? What do they want from evaluations? Do they demand findings only in a form that can be communicated quickly and easily to hard-pressed government leaders or to the mass media, or to the popular constituencies that they turn to for electoral support?

All these and similar questions go well beyond the terms of reference of this collection. But that they are not purely 'academic' questions is confirmed by anecdotal evidence and, for instance, a report commissioned by the European Parliament on

the financial instruments available to the EU for its democracy and human rights activities in third countries. The report, prepared by an independent democracy support organization, determined among other things that a 'choice must be made on whether the EU believes a more precise and systematic evaluation of its democracy and human rights policies is warranted; or whether it wishes to retain the approach that measuring impact beyond the individual project level is not desirable?' (Netherlands Institute for Multiparty Democracy 2005: 21). This is relevant to the question of whether evaluations should be routinized and conducted on a regular basis.

Here the financial, staff time and other costs—costs both to the democracy support organizations and to their partners abroad—should be borne in mind, and not only in regard to small organizations with a modest budget or few dedicated personnel capable of mounting only small-scale democracy initiatives. In their case informal systems of evaluation may make as much sense as the more formal set-piece evaluations—especially if it seems likely that the reports are destined to languish in the files. The USAID evaluation (Finkel et al. 2005) called on very substantial human and technical resources. Tracking how its findings are subsequently used in the organization and by the politicians outside as well will be very instructive. The frequency of evaluations is also pertinent. The USAID study was intended to be just the first round. But a standard recommendation argues the merits of preferring a smaller number of occasional evaluations of high quality over the multiplication of evaluations to the point where they become a major distraction to the organization and prevent it from pursuing its goals.

Putting it another way, while a 'Rolls Royce' approach to evaluation that encourages organizations to employ all possible methods and techniques might seem to offer the best chance of compensating for each method's individual weaknesses (see chapter 4), it may not be the optimal approach. That different methods will suit different situations—'horses for courses'—does not mean that all available methods should be employed for every case. Where a variety of methods are employed there is always a chance that they may produce conflicting results, with one possible consequence being confusion. For instance, what is the obvious course of action to follow where the findings of programme theory evaluation and the conclusions from results-based evaluations clearly contradict one another?

Of course the political context is relevant also to the chances of carrying out experimentation with multi-donor evaluations and, at a more modest level, the pooling of data from evaluations by different organizations, or the compiling of an open-access account of evaluation findings. The report by Bossuyt et al. (2006) for instance listed lessons learned from evaluations by other donors, but did not go on to explore what those evaluations might tell us about how to do or not do evaluations. And on this occasion the compilers did not offer any recommendations of their own.

The above discussion also prompts the question whether in their choice of what and how to evaluate organizations should be guided by trying to establish and confirm their comparative advantage—their niche in the democracy support market. The

answer may well have implications for how they view the evaluations, the methods and the findings in other democracy support organizations. A similar point could be made about situations where different organizations are involved in countries that vary widely in respect of other major political problems and in how these are perceived abroad. These might range from instances of state failure, as in Somalia, or societal disintegration, as in the 1994 genocide in Rwanda, to cases of seemingly smooth democratic consolidation or a surprising and worrying interruption to what had previously looked like a stable new democracy, as in Thailand's 2006 military coup. There may be no consensus among international actors over how they characterize and understand the main 'game in town' and the appropriate response. That has consequences for evaluation.

The fact is that there is no one organization in international democracy support comparable to the World Bank in the field of development lending that has a vested interest and enough resources to take the lead in developing evaluation methods, submitting its own projects and programmes to evaluation, and making the findings publicly available. After all, nowhere among democracy support agencies is there anything like the sizeable financial or economic incentive to seek out the benefits of evaluation that attach to the World Bank, with its multi-billion dollar spending on development assistance, or even the somewhat smaller development aid budgets of most bilateral donors.

Nevertheless, as more and more evaluations are done, perhaps an organization like International IDEA would be an appropriate location to site a public clearing house, a universal knowledge bank. That said there may be limits to the transferability of findings between different contexts and even different institutions, not least where the latter have conflicting views on the merits of different evaluation methods. Moreover, while the obstacles that flow from different reporting systems and timescales and the incompatibilities between different assessment methods may not be insuperable, some constraints are still bound to exist because of the political sensitivities that can surround democracy support policy, and the confidentialities pertaining between partners. For instance, one participant in the Stockholm workshop (Hanne Lunde Madsen) noted in regard to the often 'vague objectives and unclear indicators' of democracy support that 'many initiatives on purpose concealed their real intentions in order to survive'.

There may well be instances where the knowledge gleaned from evaluation ends up in the 'wrong hands'. This is not a reference to the root-and-branch opponents of external democracy support in the established democracies. Instead it means the opponents of democracy in countries where the people in power use the knowledge gained from evaluations to devise strategies of resistance to democracy support. Nevertheless, it seems self-evident that openness and transparency should be defining principles in anything to do with democracy support, including evaluation, so long as it does not put the lives of democracy's supporters at risk. And, while the sharing of findings among support agencies—including findings about methods—in largely

ad hoc and informal ways may be the most that can be expected, there are no strong grounds for saying that this cannot be effective. Indeed the fate that greets the rather remarkable contents of this book will be a good bellwether. And, like donor coordination more generally, perhaps initiatives like this one should be encouraged and resourced more often.

In sum, then, as well as addressing how best to combine qualitative and quantitative approaches, and in addition to evaluating democracy support against other approaches to democracy promotion, one further important item that should be on the research agenda is the question of what happens next, that is, following an evaluation. How do democracy assistance actors respond and how should they respond once a significant evaluation or sequence of evaluations has been conducted? Under what circumstances will the findings be reflected on and the right course of action instigated? How are the results taken up in the wider political debate? Can interested observers in the academic world and think tanks offer useful support in some way—and do they too not have much to learn by trying harder to find out what the democracy support agencies themselves are doing? The suggestion that more research should be carried out into the impact that evaluations at the operational level have on the larger policy decisions seems obvious almost regardless of what is judged to be the main point of having evaluations—for instance, to render democracy support fully accountable; to apply lessons from experience; to democratize the practice of democracy support; or some other goal.

Another question to address here is whether the choice of evaluation method has any systematic impact on the likelihood of institutional learning. In principle this is worth considering even if the connections are likely to be mediated by the quality of the report in which the evaluation findings are presented, and the mechanism or route by which they are conveyed. There is certainly no shortage of exhortation in this area. Bossuyt et al. (2006: 32), for example, called for a 'qualitative jump' in reducing the gap between European Commission policies and implementation practice (on governance they urge the Commission to become a 'learning organization'). Their report provides a compilation of criteria that could be used to assess evaluation reports—meeting needs; relevant scope; defendable design; reliable data; sound analysis; credible findings; validity of the conclusions; usefulness of the recommendations; clearly reported; contextual constraints. Naturally, some of these criteria would benefit from fuller specification, for instance, whose needs?; credible in whose eyes?; useful to what end?; and so on. Other writers on evaluation have formulated the criteria slightly differently, like Forss (2002: 26) who summarized the evaluation industry standards as utility; feasibility; propriety; and accuracy. Of these, Madsen (in this book) for instance places special emphasis on propriety. All in all, the following observation by Cole et al. (Danish Ministry of Foreign Affairs 2006: 41) for all its simplicity is well worth repeating. If institutional learning is to be possible then it helps if the links between analysis, findings and recommendations in evaluation reports are made as explicit as possible and some order of priority is

indicated and if necessary itemized stakeholder by stakeholder, when evaluations throw up an impossibly long list of recommendations. Given the commitment that USAID is now making to evaluation, the amount of impact this has on future US democracy support policy will be a major test case in this regard.

Evaluation in perspective

At the end of the day, views about what democracy is and what democratization looks like will shape how the possible findings from evaluation are conceived in advance and how the actual findings are received in practice—the 'results' that are expected from the process; how the findings are interpreted; the recommendations that are seen to follow; and the action, if any, taken. The political dynamics that even successful democracy support interventions play a part in may be hard to grasp even after the event, evaluations notwithstanding, let alone flag up a train of developments that can be foretold in advance. Quite apart from that, however, and regardless of how advanced the evaluation methods are and no matter how illuminating their findings, they will not by themselves tell us how—or even whether—the findings will be acted on or, indeed whether they should be acted on. This is especially true in regard to such a politically embedded and highly charged subject as democracy support. Even so, this need not be a cause for alarm.

In some measure all the variables just mentioned must be context-specific, on both sides of the democracy support relationship—the democracy support organizations and the societies whose politics are attracting external interest. In 'getting the evaluation methodology right' there is no escaping that there may be some hard choices for instance, overviewing the activity in largely technical terms; or as a procedure essential to making the case for continued funding; or as an exercise in team-building among democracy practitioners on the support side or across both sides in an international partnership; or as something to convey democratization in its own right; an opportunity to promote a specialized form of democratic governance capacity-building. In some situations evaluation's contribution to a solution for acute social conflict or similar malaise that is connected directly or perhaps very indirectly to shortcomings in democracy or human rights may have to be taken into account. In practice it is likely that a number of views about the purpose of evaluation will be present even within the same democracy support organization. Trying to satisfy all of them could well be an evaluation challenge too far. And we should not be surprised if there is some institutional or personal resistance to evaluation, especially if it threatens to bring a loss of control.

The following chapters explore these and many other aspects, as they help further new directions in the evaluation of international democracy support. Forss (2002: 33) even believes that that evaluation is of itself pioneering in nature: 'The fact is, that an evaluator who simply used the models of another evaluator would be accused of plagiarism and would get a bad reputation'. This looks rather extreme. Yet democracy

support itself is an area of human endeavour that is evolving as we speak. It seems likely to remain a moving frontier—or, more accurately, constellation of frontiers—for some considerable time. It is also an inherently political endeavour, both in its object and the driver or motivation. For this reason it would be no disservice to conclude this introduction by saying that methods and experiences of evaluating democracy support should always be kept in perspective. Certainly, the fact that 'there is much that social scientists do not yet know about how democracy grows or is eroded' (Finkel et al. 2005: 59; 84) means there are likely to be more investigations in the future, and this is how it should be. However, while it is 'generally understood that evaluations do not provide final answers, they enlighten the debate, guide decision-making and extend knowledge' (Forss 2002: 3), it is also reasonable that evaluators do not have the final word, especially in democracies.

Thus the conclusions of even the most advanced thinking on evaluation, no matter how favourable the evaluations are to democracy support, cannot by themselves determine whether, where and for how long democracy support should remain a prominent feature of international politics in the 21st century. They offer no guarantee that democracy support efforts will be rewarded with success, let alone that democracy's advance will provide an answer to every need. Indeed, the easy victories for democracy's progress have now been won. Not just opposition to democratic reform but resistance to external democracy support are now more in evidence in many of the countries where little progress has been made. Whether methods for evaluating democracy support should take this scenario into consideration is a moot point, for in the present climate it is not even obvious how democracy promotion strategies and the practice of democracy support should respond. Evaluating the strength of what has been called the 'backlash' against democracy support and seeking to understand the causes should be undertaken as a matter of urgency. The same is no less true of what might be called the 'frontlash', a shrinking away from the idea of offering democracy support that may be detected in some political circles in the established democracies. A distinct lack of enthusiasm has been reported in many European governments and left-of-centre political parties (Mathieson and Youngs 2006).

And yet by the same token we should not assume that reservations about the effectiveness of democracy support, even in the presence of sound techniques, or disagreements over how to evaluate the findings, or evaluations that throw up unfavourable findings, will signal the death knell of democracy support. The Westminster Foundation for Democracy, for instance, has survived a critical evaluation (River Path Associates 2005) although naturally there have been some changes as a consequence. By comparison the verdict that was delivered on the NIMD (European Centre for Development Policy Management 2005) was more favourable, but even it issued a call to pause and reflect. The recommendation to invest in upgrading the NIMD's managerial systems before aiming to renew growth and expand its reach to new countries can be read both as highlighting a limitation and as a vote of confidence in the future.

For the governments of established democracies the international kudos to be gained by being seen to endorse a sincere commitment to democracy support—or at least the adverse reputational consequences of being thought to be hostile—may continue for some time to come, evidence of 'backlash', 'frontlash' or other reactions notwithstanding. But in a rapidly changing world, demonstrating the effectiveness of international democracy support for democratization may not be enough to secure the activity's future indefinitely. This is not so much because democracy support seems likely to become a victim of its own success, or a casualty of failure or of a proven inability to do better. Rather it stems from the premise that we cannot assume that democratization will be seen as a solution to large and potentially hugely threatening problems that could come to crowd in on human beings almost everywhere. And at the very minimum, democracy support is not normally the most significant influence on democracy's progress: this is one finding that nearly all the evaluation studies so far agree upon.

So, while the evaluation of democracy support and research into methodologies both justify more attention, in part because they start from such a relatively low base when compared to, say, the evaluation of international development cooperation, it might prove difficult to justify a quantum leap in allocating resources to these activities. In all probability there will be more evaluations. But without more of the kind of stimulus offered by the exchanges that led to and inform this book, the growth of a culture of doing evaluations, while necessary, may still not be sufficient for there to be a concomitant investment in developing the evaluation methods.

The conclusion that the Stockholm workshop itself arrived at, that more research is required and a rolling programme of further workshops should be arranged, might seem predictable. Better, however, to finish on a more ambitious note, as befits the significance of the desire to see a world populated by democracies and universal respect for human rights. Marching orders for the way ahead should involve exploring ways of assessing and more especially comparing the effects of all the different kinds of external intervention, both positive and negative, and unintended as well as intended influences. Of course democracy support must be included in this, even though it is only one of the many ways whereby external forces and actors interact with the domestic counterparts that in most cases still seem to be the primary determinants of a country's politics and its pattern of political change.

Chapter 2

Evaluating the impact and effectiveness of USAID's democracy and governance programmes

Chapter 2

Margaret J. Sarles*

Evaluating the impact and effectiveness of USAID's democracy and governance programmes

This chapter describes the efforts of the US Agency for International Development (USAID) to examine the impact and effectiveness of its democracy and governance (DG) programmes through the Strategic and Operational Research Agenda (SORA). With annual democracy budgets in the range of 700 million US dollars, even excluding Afghanistan and Iraq, the need to invest in this research is clear. SORA focuses on developing rigorous comparative methods, including country case studies, large-scale quantitative studies, systematized expert interviews, democracy surveys, and specialized comparisons of areas such as the rule of law. While limited by the state of art in academia in measuring processes of democratic development, some progress has been made. A path-breaking quantitative methodology has shown a significant positive relationship between USAID DG assistance and some processes of democratic change, with the highest impact being found in civil society and electoral assistance, and in countries where initial levels of democracy were lowest. Time-series data from democracy surveys are now yielding solid information on changes in democracy that are attributable to USAID programming. The National Academy of Sciences is assisting USAID in refining its operational definitions of democracy, developing better comparative methods for country case studies, and recommending how to combine various methodologies for best results and improve future evaluations.

Introduction to the Strategic and Operational Research Agenda (SORA)

During the wave of democratization that has swept through the world over the past 25 years, the United States Agency for International Development (USAID) has supported processes of democratization in more than 125 countries, extending to every region of the world. As governments moved from authoritarian to more democratic political systems, replacing dictators with elected leaders, institutionalizing democratic rights and procedures into new constitutions and developing new systems of accountable governance, USAID for the first time in its history developed a broad array of programmes to encourage and accelerate those processes.

Often in opposition to many in the development community—sometimes even within the institution—USAID began to support a new field of democratic development alongside health, education, the environment and economic growth. From its initial forays in the 1980s in Latin America focusing on elections and justice, programmes broadened to accompany democratization in the new post-communist regimes of Eastern Europe and extend support to democratic reforms in Africa and Asia. Democracy and governance now represent the second-largest sector of the agency's work, with programmes currently existing in 82 countries. Programmes roughly divide into four categories: Good Governance; Civil Society; Rule of Law; and Elections and Political Processes.

The purpose of this chapter is to describe USAID's efforts to examine the impact and effectiveness of the agency's extensive democracy and governance programmes over the past 20 years, through a number of evaluation and research activities loosely grouped under the Strategic and Operational Research Agenda (SORA). The SORA findings should provide a firm analytical base on which to make decisions regarding the type, mix and sequencing of democracy and governance programmes.

The core of SORA research and evaluation focuses on developing rigorous comparative evaluations of present and past country and programmatic interventions, to capture USAID's field experience and to learn from it. It has expanded over time to include numerous methodologies: SORA activities include the entire process of developing indicators, evaluation, dissemination of findings through training and other means, and adoption of better practices. Ultimately, SORA's success will be measured not by the strength of its findings but by its success in improving future democracy programmes and policies.

SORA began in 2000. As it developed and tested methodologies and theoretical frameworks, the serious nature of the challenges of evaluating democracy and governance interventions became more apparent. We needed to find, or in some cases develop, methodologies for analysing the complex realities of changing political systems and the relationships among different areas of change—a neglected field of political science research. Just as difficult, we needed to be able to measure donor impact on that change and capture it in a rigorous, comparative way so that policy

makers could make decisions based on strong evidence of what was likely to succeed. We needed to understand not only a programme's relationship to small immediate changes but also its relevance to building a sustainable democracy beyond the project level. To answer these needs, SORA has become more ambitious in scope.

At the time of writing some aspects of SORA are complete; some are in progress; many are still under discussion. We have had false starts and greater difficulties than anticipated on many fronts. Perhaps most serious of all, we have had to face the reality that the rigour with which we can examine the impact and effectiveness of our work is limited by the stage of development of academic inquiry that seeks to explain and measure processes of democratic change. Nonetheless, some preliminary results to date give us some confidence that we are on the right track. We have been able to engage some of the best academic and policy-oriented researchers in developing evaluation methodologies that will lead to better decision making on where to invest resources in the future. And we have similarly begun to develop linkages with colleagues in other donor agencies to learn from their evaluation efforts. SORA is now part of an exploding, transforming field in which a large community of academics and practitioners are developing indicators, methods and theory. This chapter is designed to contribute to that transformation.

The account given here first discusses why SORA activities are important to USAID and the US Government. It notes briefly earlier efforts at evaluation and then discusses the main methodologies and findings to date. The methodologies include worldwide quantitative analyses, in-depth expert interviews, country case studies, cross-national surveys and thematic comparative studies. The chapter concludes with a short discussion of plans for the coming year and a summary of what we see as the main challenges yet to be overcome, conceptually and logistically, in developing appropriate evaluation methods and implementing them.

The rationale for SORA

SORA represents a significant evaluation commitment for the US Government over a number of years. Why has USAID been willing to undertake this? We can point to several critical factors.

First, USAID and the State Department have been investing heavily in the field of democratic development for well over 20 years, at an increasing rate (see figure 1). Excluding Afghanistan and Iraq, by fiscal year (FY) 2006, USAID grants to governments and non-governmental institutions totalled nearly 750 million US dollars (USD). The median size of USAID democracy programmes has also increased. From 1998 to 2005, the median size of democracy and governance funding per country rose from 3.5 million to 5.5 million USD per year. Simple fiduciary responsibility requires us to improve our knowledge of how best to support democratization, understand what is likely to work and not work under what circumstances, drop programmes that not work, and maximize the benefits.

Figure 2.1: USAID-managed democracy and governance programmes
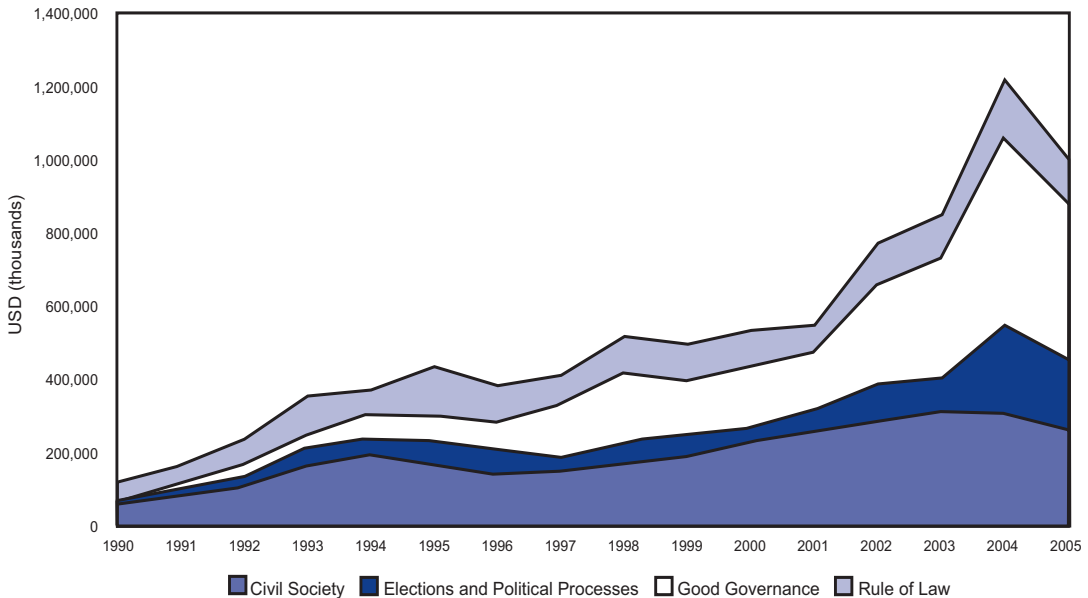
This responsibility is even greater now than before because democracy promotion is a high foreign policy priority for the US Government, and has been over several administrations. The US Government is likely to continue to support the efforts of democratic reformers. As part of a larger US Government effort that includes diplomatic as well as development assistance tools, USAID's efforts in democracy promotion need to be as well targeted and successful as possible.

Second, often what we 'know' is based on insight and anecdote rather than empirical evidence. And it is quite possible that we have been biased on the side of believing in our success. During the prolonged period of worldwide democratic improvement over the past quarter-century, it has not been too difficult for any donor to find strong positive relationships between democracy assistance and improvements in democracy. However, the jury is out as to whether or how that assistance has actually helped democratic reform. As Dr Gerald Hyman, Director of the Office of Democracy and Governance at the time, noted to a group of experts in 2002:

> There remains, both within the community of practitioners and analysts, profound uncertainty about the efficacy of democracy assistance…. We do not really know with any degree of certainty—and based on empirical evidence—what works

and what does not, what works better and what works less well, in any particular context, or in general, for that matter. In the main, we have been left to depend on vague generalizations—slogans even—based as much on hope as experience (Hyman 2002).[i]

Success in democratization over the past quarter-century has had a thousand 'fathers'—actors claiming the credit—including USAID. However, future support for democratic development is likely to be at risk if we continue to rely solely on anecdote and individual success stories. The international arena now is a very different place compared to 25 years ago, and making democratic gains in the future will be more difficult than it was in the past. The unreformed authoritarian governments of today may be the most difficult and the least likely to accept donor support for democratization efforts, certainly within the short time frames US Government planners usually use. With a 'first generation' of improvements in human rights and electoral processes under their belts, many other countries either seem to be finding further reforms increasingly difficult or, in many cases, are actually regressing on a number of democracy indicators. If US Government democracy programmes were thought a success during a period of a 'rising sea', they may be vulnerable to charges of failure in the next period, which may well be more characterized by backsliding.

Third, both the USAID leadership and USAID democracy officers in the field are demanding better information. In 2006, the new director of foreign assistance mandated a set of 'common indicators' of progress across all USAID missions in order to get an improved understanding of country programmes and to be able to report better to Congress and the executive branch. This came on the heels of several years of pressure from policy makers to find out what the US Government (the Department of State and USAID) were actually doing to support democracy, and whether it made a difference. SORA activities are meant to respond to this policy need.

Much of the impetus for improved democracy indicators, and for better knowledge of where to invest scarce field mission resources, however, comes from democracy officers in the field. With some notable exceptions, the compendia of 'quantitative indicators', 'qualitative indicators', 'lessons learned', assessment methodologies, handbooks for action, guidelines, training and other products have often not been based on enough rigorous research. In consequence they usually provide only very general and often untested guidance and insight to the field. As the managers ultimately responsible for whether their programmes have a sustainable impact on democracy and governance, field officers have often incubated some of the most innovative measurement work, particularly democracy surveys. But in such cases their practices are often not widely known nor replicated across USAID democracy programmes.

Fourth, our current evaluation practices are totally inadequate for determining what is likely to lead to sustainable gains in democracy and governance. As part of SORA, USAID commissioned a Social Science Research Council (SSRC) study of the many democracy

evaluations already undertaken. Could we extract from them the comparative data we needed to make policy recommendations on where and how to support democracy and governance? The potential savings in research time and cost would have been significant. The results of the study, unfortunately, were not encouraging.

The researchers identified three major flaws with the evaluations that limited their validity even at the country level, and made cross-country comparisons impossible.

- First, the evaluations did not consistently provide the most basic facts of the project, such as funding levels and the length of implementation time, making comparisons along these two important dimensions impossible.
- Second, the evaluations tended to focus on the immediate outcomes of very specific activities (for example, the number of judges trained) rather than on their link to a broader USAID goal or interest (such as improvements in the rule of law). This made it difficult to assess whether an activity or programme, however 'successful' itself, was relevant to the larger picture of building democracy.
- Third, it was rare for an evaluation to consider whether other factors could have been responsible for the outcome, rather than the USAID intervention: alternative explanations were not explored.

Finally, USAID's democracy promotion efforts have often been judged harshly as 'irrelevant' to the big-picture challenges of democracy. Notwithstanding the existence of USAID programmes, authoritarian tendencies have re-emerged in a number of countries. One charge is that we fund small programmes that may be 'successful' but in fact are irrelevant to the main democracy issues in a country, and have no real impact outside the narrow parameters of the specific programme. Is this a valid complaint? Certainly USAID democracy strategies link their programmes 'up the system' and hope to have an impact on the most critical democracy problems. But to what degree should a specific democracy project, or even an entire USAID democracy and governance programme, be expected to have an independent, measurable impact on the overall democratic development in a country?

The above sets a high and perhaps unreasonable standard of success. Decades ago, USAID stopped measuring the success of its economic development programmes against changes in the recipient countries' gross domestic product (GDP). Rather, we look for middle-level indicators: we measure our anti-malaria programmes in the health sector against changes in malaria statistics, our support for legume research against changes in agricultural productivity. What seems to be lacking in democracy and governance programmes, as opposed to these areas of development, is a set of middle-level indicators that have two characteristics: (a) we can agree that they are linked to important characteristics of democracy; and (b) we can plausibly attribute a change in those indicators to a USAID democracy and governance programme. It seems clear that we need to develop a methodology that is able to detect a reasonable, plausible relationship between particular democracy activities and processes of democratic change.

Earlier efforts: the Center for Development Information and Evaluation and SORA, Stage 1

The current SORA effort is based on earlier efforts to measure the impact of USAID democracy programmes. One of the earliest was a series of comparative 'impact evaluations' undertaken in-house, mostly during the 1990s, by the Center for Development Information and Evaluation (CDIE) in many sectors of development, including two in the area of democracy that focused on programmes for legislative strengthening and democratic local government. For each study, teams of experts evaluated programmes in five or six countries, including by way of field investigations of around three weeks per country. The case studies were synthesized in a final 'lessons learned' document intended for policy makers and general users. Each study was based on a common framework; each required a serious commitment of USAID resources and time.

The results, however, were problematic in the democracy field. In some areas of development, for instance agriculture, USAID had many years' accumulated experience and a strong history of research and evaluation; inquiring into 'impact' was feasible in such cases. In contrast, democratic development was still a very young field, with virtually no academic research behind it. Moreover, because programmes were still not very old, 'impact' studies occasionally included cases where programmes had been in place for less than two years. Certainly there were important insights gained, particularly from the individual case studies. Overall, however, the comparative analysis yielded very little: the lack of available data and experience, the lack of a theoretically sound underpinning, and the very general findings severely limited their usefulness.

In 2001, the Global Center for Democracy began SORA as a pilot comparative evaluation methodology to determine whether it was possible to assess whether USAID democracy programmes had actually changed countries' political systems. The two questions SORA set out to answer were 'In what areas did USAID's democracy programmes positively or negatively influence the political system?' and 'What factors might explain the success or failure of USAID's democracy programmes?'. Academics, independent contractors with expertise in evaluation and democracy, and in-house democracy experts collaborated on this ambitious effort. Two sets of three countries were chosen for sequential evaluation based on hypotheses about where democracy programmes were more likely to have been successful or less successful. The investigation advanced the field of comparative methodology in some significant ways. It included, for example:

- *a common definition of democracy.* This allowed us for the first time to develop common dependent variables, so that we could answer the question 'What should change as a result of our democracy programmes?'. The methodology settled on three 'levels' of expected change—individual ('cultural'), institutional and national;

- *a common methodology, namely 'process tracing'.* Field researchers tried to follow a cause and effect chain from a USAID democracy intervention to its consequences, wherever they were. This comprehensive method let the analysts examine in depth why some programmes succeeded while others failed, as they traced through each process. It also captured unanticipated consequences; and
- *the planning of two kinds of research, to be able to focus on both country and sectoral-level findings—overall country comparisons and comparative sectoral studies.* Six pilot country case studies were planned. In the first year, three relatively successful programmes were chosen, and in the second year three that had been more problematic. At the same time, research teams began to prepare for thematic, cross-national sectoral studies in areas such as legislative strengthening, civil society, political party strengthening, local democratic governance and rule of law in order to compare their effectiveness and learn what kinds of activities were more or less successful under what conditions.

Methodological findings

SORA I was a pilot, to determine whether it is possible to measure the impact of donor democracy programmes at the national level. It found that process tracing was a great advance in linking support for specific activities to broader democracy outcomes, and should be a part of future efforts to trace relationships between assistance and national-level democratization.

However, its limitations also became very apparent. As a method it is biased towards attributing a country's democratic change to a USAID intervention because it does not seriously examine alternative explanations. Furthermore, at least in a short period of time, it is not possible to compare the success of one kind of programme with another, or to get very far in understanding how or why specific interventions seemed to succeed or fail. In a subsequent test of the methodology, researchers tried to compensate for both these weaknesses by focusing on one geographic area and one sector (rule-of-law programmes in Eastern Europe), and beginning their work with open-ended discussions with key national figures to see if their assessments of the causes of change included activities funded by USAID. While these methodological changes improved the evaluations, other time-related and theoretical issues remained. Process tracing became increasingly difficult to use after the first few steps. As one researcher reported, it went 'from a small stream … to a delta, and it became very difficult to follow all the streams'.[2] There were often big gaps in tracing causality that researchers needed to leap over, substituting theoretical assumptions for on-the-ground evidence. In addition, the three-week time frame for in-country study was insufficient for carrying out this kind of causal analysis, which is a painstaking process involving many interviews and can lead the inquiry down some unexpected routes.

Researchers were also forced to make compromises that illustrated potential problems with future evaluations. The careful process of country selection was

compromised when some field missions declined to be selected as pilots. Even more important, researchers occasionally encountered restrictions on their work set by the US embassy or the USAID mission, limiting their access to stakeholders in the country, and thus compromising the validity of the findings. In addition, the 'individual (that is, cultural) change' level of study proved not to be fruitful.

The syntheses also demonstrated a basic conceptual problem in developing policy guidance from the investigation. The process tracing method provided rich detail and often represented excellent scholarship. However, at the end of the day there were too much data, too much richness, and too much nuance for rigorous national-level comparisons. We were drowned in a sea of information. USAID needed to undertake much more synthesis and discussion to make this a success. The work plan had to go from design through dissemination to policy makers and practitioners.

As the evaluation project developed, another major logistical issue also emerged. SORA I presumed that the country-level process tracing would be followed by in-depth sectoral analyses or case studies. Teams were beginning to gear up to undertake cross-national studies in legislative strengthening and other areas. How were they to work in the field in conjunction with subsequent country-based evaluation teams? Would one set of teams be travelling the globe undertaking country studies, while other sets visited the same countries, often asking the same questions and looking at the same issues, to examine one aspect of the programme in greater depth? The sheer expense, the potential duplication of effort and the nightmare logistical problems finally brought a hiatus to this approach.

Substantive findings

SORA I tentatively concluded that most programmes did indeed have an impact in the sector in which they operated, with electoral support programmes the most clearly successful. One of the most intriguing findings seemed to be that USAID's approach to long-term institution-building in democracy and governance might be flawed. The case studies showed a pattern of success when 'institutional development' focused on a single institution, such as elections, and much more ambiguous results for more complex institutions, such as a local government systems or justice systems. There was no hard evidence that such programmes did not work; only that there seemed to be 'ludicrously short' time frames and unrealistic expectations of what could be done. This is a critical area where more evaluation is clearly needed.

A second finding was that programmes in democracy and governance were often too 'technically' oriented and did not sufficiently consider the real incentives of stakeholders and the 'politics' of the situation. Hence, programmes in legislative development often foundered as they became footballs in political party competition. Again, this finding needs further testing through field studies.

SORA, Stage 2

In 2003, USAID went back to the drawing board and commissioned the SSRC to develop a methodological and analytical strategy for evaluating USAID democracy programmes, this time setting up an Advisory Group of academic experts to review the findings.[3] The result was the 'Research Design to Evaluate the Impact of USAID Democracy and Governance Programs' (Bollen, Paxton and Morishima 2003).[4] Their recommendations became the core of the current SORA evaluation effort. Because they may also be useful to other donors, a brief summary of them is provided here.

1. *Focus on the future as well as the past.* Most current projects did not collect the baseline data needed to determine later on what change had occurred and why. Therefore, the processes of data collection and analysis needed at the beginning of every programme should be instituted now so that reliable, comparative evaluations are possible in the future.

2. *Focus on democracy 'activities' rather than a more general sectoral level.* Although the recommendations offered some guidance on how to proceed in sectoral analysis, they also suggested dropping down a level to one where better, more quantitative, measurement was possible. For example, in justice reform, a programme includes many activities such as judicial training, advocacy by non-governmental organizations (NGOs) for justice reform, constitutional reform and so on. Comparing success and failure at this low level of analysis can be more rigorous, although the cost is that it allows fewer conclusions about overall democratic impact. Of course, it is possible to fund the examination of only a small percentage of such activities, so they should be chosen carefully. The SSRC therefore outlined a process of data collection, resource analysis and analysis of expected outputs and outcomes to help determine what activities should be studied.

3. *Do not use only one methodology in evaluating the programmes.* Different methods, and combinations of methods, are appropriate for different sectors and activities. Once it is determined which programmes or activities will be evaluated, consult with an expert group to determine the best mix. The use of combined methods will provide greater rigour, more flexibility, attention to results at different levels, and 'triangulation', that is, using multiple methods and testing the findings arrived at by one method with the findings of another method.

4. *There are basically six kinds of methods that should be considered, and mixed, in undertaking comparative democracy evaluations.* These include randomized experiments; quasi-experiments; surveys of individuals and groups; interviewing and site visits; sector overviews by country experts; and cross-national quantitative research. Each has a slightly different function and provides a different mix of findings at the 'output', 'outcome' and 'impact' levels. Some can be used only in future programmes, others to evaluate past and ongoing activities. The report provides a great deal of detail on how to apply each method. For example, the 'cross-national

longitudinal data collection and analysis' provided specific recommendations on the international democracy data sets that should be compiled, along with specific USAID data, for inclusion in a common data bank for analysis.

5. *Convene a task force* to make the difficult choices on the particular sectors, areas or country programmes to evaluate, once the agency has the data needed for informed decision making.

The SSRC report became USAID's reference document for how to move forward. It is important, therefore, to note that some issues were not covered.

First, the report did not discuss how to define 'success' in terms of democratization. It did not include discussions of how to select dependent variables, or how far up the chain of attribution one could expect to go. Second, by concentrating primarily on the 'activities' within projects the report did not discuss how to measure the interactions of different democracy interventions or how to assess the relative success of different sectors of work that might lead to policy shifts. The report noted that until USAID collected better data it would be difficult to go beyond this level and make any plausible claims of attribution.

USAID responded to the report by initiating activities in several directions. Data collection is now being improved. Large-scale quantitative research commenced. USAID has accepted the importance of evaluation 'triangulation' as described above. And USAID began a new effort with the National Academy of Sciences to supplement the quantitative research with case studies that would overcome the problems that had been experienced in the earlier SORA work.

Setting up a Democracy Database

While it was not initially feasible to develop as intensive a database on USAID programmes as the SSRC recommended, we created a USAID Democracy Database that covered all financing of democracy and governance programmes from 1990 to 2003 (it has since been updated to 2005).[5] Although it is based on official budget figures, developing this database was a mammoth undertaking. USAID's budgets were not designed to trace funding over time or to be useful for comparative research. Budget coding systems changed frequently, so that, for example, for one year, financing for media programmes might be hidden in a general civil society code, and for another year it might have its own coding number. In addition, funds were often obligated to a grantee or to a region, rather than to a country or a country programme, requiring painstaking investigation and budget manipulation. SORA was able to provide significantly better information to USAID on the nature of its democracy funding than could be found in its official statistics; unfortunately, this also meant that the numbers varied slightly from the official budget figures and were not readily accepted. With a new agency system in place in 2007, some of these problems may now be lessened. SORA figures are now compatible with the agency's own budget figures, so

that in future years the Democracy Database may become an official data source.

The USAID Democracy Database has already proved its usefulness in measuring programme impact, as demonstrated below. However, it has also had an important serendipitous role as an independent SORA product. For the first time, we have been able easily to explain what we do and have done in every democracy programme over time. We can, for example, show for one country, such as Indonesia when we began a democracy programme, the overall level of funding over time, the major components of the programme (rule of law, human rights, civil society, media, good governance, and elections and political processes) and the amount and proportion each of these has contributed compared to the total. We can also show total levels of democracy funding in comparison to overall USAID funding in a country, and to other areas of development financing. Charts and tables for congressional briefings, for talking points and for supplying background material to the agency leadership and our own office are now standard products that USAID can provide.

A worldwide quantitative study of USAID's democracy impact

Through a university-based competition held by the Academic Liaison Office (ALO) for University Cooperation in Development (a consortium of six presidential higher education associations), USAID commissioned Steven Finkel and Aníbal Pérez-Liñán of the University of Pittsburgh, and Mitchell Seligson of Vanderbilt University, to undertake a longitudinal, worldwide research effort to measure the impact of USAID's democracy and governance programmes in 2005. Dinah Azpuru of Wichita State University provided assistance. The final study, 'Effects of US Foreign Assistance on Democracy Building: Results of a Cross-National Quantitative Study' (Finkel et al. 2006), is a comparative analysis that employs complex growth models not often associated with political science research. It is by far the most scientifically rigorous work we have yet undertaken. Covering the period 1990–2003 (now updated to 2005), it examines the relationship of USAID democracy assistance to changes in national-level indicators for freedom and democracy from the Freedom House and Polity data sets, controlling for alternative explanations.[6] While the results are necessarily general, this approach overcomes some of the problems of bias, attribution and overly nuanced findings that had plagued the earlier case studies. This new study comes closest to giving us sound findings on where and when our democracy programming has had the greatest impact on democratization.[7] In addition, USAID set up an outside academic expert 'review panel' to guide and critique the research at key points.[8]

The most important impact findings can be summarized as follows.[9]

1. Using Freedom House and Polity IV measures of democracy, USAID democracy programmes have had a clear and positive impact on democratization worldwide.

In the words of the authors, 'USAID Democracy and Governance obligations have a significant positive impact on democracy' (Finkel et al. 2006: 3).

2. USAID democracy assistance is not only statistically significant, but can be an important factor in raising a country's democracy levels. The average country eligible for USAID democracy assistance increased its Freedom House score by about five one-hundredths (.05) of a point per year on a 13-point scale over the 14-year period of the study, 1990–2003—a total of a 1.1 point change over the period. Each 1 million USD in USAID democracy assistance increased that value by 50 per cent over what would have been expected that year.[10] With an average programme funding of about 2.07 million USD per year over the 14-year period, therefore, USAID funding doubled the amount of democratic change that an average country would otherwise have been expected to achieve in the respective year.[11]

3. Funding levels are critical, since the level of gains in democracy depend on the investment. The larger the investment, the larger the gain. Median programme size continues to rise; excluding Iraq, it reached 3.32 million USD in 2004. A 10 million USD annual investment would cause a fivefold increase, according to the model—a gain of half a point on the Freedom House scores each year. Funding impact was not conditioned by the size of the country.

4. Two other variables also contributed to democratic development—(a) growth in GDP over the past year, and (b) the 'neighbourhood effect', the level of regional democracy in the past year. These findings are consistent with other research showing that short-term economic performance and 'diffusion processes' from neighbouring states contribute to democratization. In the other direction, political conflict and violence both had a negative short-term impact on a country's level of democracy. Many other possible explanations for the rate of democratization, including measures of economic development, social indicators, political history, and indicators of state failure and the diffusion of democracy, were also considered but not found to be significant.

5. Some areas of funding seemed to be more successful than others. Using the breakdown offered by USAID's democracy data set, the model looked at the individual impact of programmes under the Elections and Political Processes, Rule of Law (including a further breakdown for human rights), Civil Society (including a breakdown for media support), and Good Governance headings. The best investment seemed to be civil society programmes, followed by investments in elections and political processes, in terms of overall impact on democracy. In human rights, a significant negative relationship was found: USAID funding seemed to lead to an increase in human rights abuses.[12] This may be because of better reporting of abuses, although it is also possible that democracy support could lead to a crackdown by an authoritarian government and thus a real increase in abuses.

6. It was not possible to measure whether USAID funding had an impact on the rule

of law, or on governance, primarily because the literature offers no good measures of these concepts. We are now supporting more research in this area. However, the Rule of Law programming did have a lagged effect on overall Freedom House scores.

7. Based on preliminary analysis, there are important differences in impact by region. The effect of USAID democracy assistance has been strongest in Asia, followed by Africa, and lowest in Latin America.

8. Contrary to expectations, democracy programmes seem to have the greatest impact in countries with more difficult political and social settings, where the initial level of democratization is low. We expected that a more middle-level country, with more developed institutions and human capital, could more easily take advantage of funding so that it would have greater impact. In contrast we expected to find that countries struggling in the early stages of democratization, often poorer and with a multitude of other issues occupying the attention of government, would not make as much progress. But the finding was quite the opposite: the lower the initial level of democracy, the greater the gains as a result of USAID programmes. The study also found important lagged effects, suggesting that programmes may take several years to yield results and that the results of programmes may be cumulative. A follow-up quantitative study will look at the cumulative effects.

The methodology was powerful and the results compelling and positive. The growth model used in the study yielded good results, and the concentration on testing for 'alternative explanations' for improvements in the rate of democratic development addressed one of the fundamental criticisms of the SSRC. It also overcame a principal weakness of the process tracing that had been used in the country case studies, which did not consider alternative explanations in a systematic way. This aspect of the methodology probably did more than any other to convince sceptical academics as well as policy makers that USAID's programmes did indeed have a positive effect, even at the national level, on increasing the rate of democratic development.

Of course, the results raised as many questions as they answered. While the model showed that USAID has had a significant overall impact on democracy, it relied on aggregate numbers, and could not tell us which specific countries benefited most and which benefited least from democracy assistance, by how much, or why. And, while the researchers tested for a very wide range of possible explanations for democratic change, on the basis of an exhaustive literature review, the aggregate model itself explained only a part of the process of democratic change observed. Obviously, other factors that have not yet been considered are also important. The study also demonstrated the lack in the literature of precise definitions of operationally commonly used concepts of democracy. The authors have now embarked on a follow-up study to provide greater depth to the study and answer

a number of unresolved issues. This follow-up also includes a specific element to develop a better dependent variable for changes in human rights and governance. The negative relationship demonstrated between USAID democracy assistance and improvement in human rights levels is of great concern to USAID.

Democracy surveys as evaluation tools

Survey research is now emerging in USAID as one of the best and most rigorous evaluation tools for measuring the impact of democracy programmes. Following the end of the civil war and the 1994–6 peace accords in Guatemala, the USAID Mission in Guatemala commissioned Mitchell Seligson to undertake the first national survey of citizens to gauge their support for democracy, the DIMS (Democratic Indicators and Measurement Survey). This survey set the model that has now become the norm throughout USAID missions in Latin America. By 2006, every USAID mission in Latin America and the Caribbean had participated in the surveys, and a similar effort is being scheduled for 2008.[13] USAID-funded national democracy surveys have certainly not been limited to Latin America, however. Most missions in Eastern Europe and Eurasia have carried them out, as have many missions elsewhere in the world, and USAID has also supported Afrobarometer surveys. Outside Latin America, and to some degree Africa, however, surveys have generally not been carried out at regular intervals, nor have the same questions always been asked, which has limited their usefulness as an evaluation tool. This is changing rapidly.

There has traditionally been an assumption that such national-level surveys suffer from several weaknesses that make them unreliable and not very useful as evaluation tools. Because they are national in scope, they might seem inappropriate for looking at project-level change, some of which is below the national level, often focused on a few regions. This national-level orientation has not proved to be a shortcoming, however. Many programmes of civic education, electoral observation and civil society advocacy are nationwide, or the programme objective is to reach people throughout the country. Furthermore, the national surveys are carried out in appropriate indigenous languages (six in Guatemala, for example), allowing programme managers to compare the effectiveness of their programmes by ethnic group, as well as gender, income and employment, and a host of other variables.

Surveys took a quantum step forward as a good evaluation tool with the introduction of over-sampling in the 2004 round, which increased the number of people interviewed in the area of the projects. For the first time, we could measure attitudes, perceptions and behavioural characteristics that we hoped to change, before beginning a programme, as well as monitor change during the life of the project and evaluate the final effects at the end—'before' and 'after'. We could correlate behaviour with other attributes, such as voting behaviour, or trust in government. Additionally, the surveys compared changes in the project area with overall national changes, so that we could compare change in the project areas with change outside the project

area—the 'with' and 'without' scenarios. This essentially sets up a 'semi-experimental' research design, and is the closest we have come to developing indicators that test the actual effects of USAID programmes, to see what change can validly be attributed only to the USAID programme.

Democracy surveys have often been seen as a way of looking at 'democratic culture', a somewhat diffuse concept that seems to have little relationship to the kind of specific dependent variables needed for the evaluation of programmes. In fact, the attitudes, perceptions and behavioural attributes of citizens now help us to define levels of democracy and look for specific changes that can be attributed to USAID programme interventions. The quantitative study carried out by Finkel et al. (2006) focused on Freedom House and Polity IV scores as dependent variables, which often emphasize the strength of democratic institutions and substantive rights. In contrast, the democracy surveys focus on concepts such as support for democracy and political tolerance as measures of the state of democracy. Both have their place and each can reinforce the other.

Specific clusters of questions and indices developed from the surveys are used to monitor and evaluate programmes. In the area of behaviour, for example, the Latin American surveys inquire about whether a respondent has been the victim of a crime, or been asked for a bribe, or participated in various kinds of political activity. All of these measures are used to evaluate USAID programmes. Democracy programmes also may focus on changing attitudes (for example, whether it is worthwhile voting) or perceptions.

Democracy surveys are growing rapidly as a mission-led evaluation tool. In some countries, such as Bolivia, over half of the annual indicators used by the mission to measure the progress of its own democracy programmes are now taken directly from the survey. Similarly, the 2006 democracy survey in Indonesia includes many questions that give the mission the tools needed to monitor changes in the specific areas of democracy they are promoting. Missions in Africa have used the results more as a basis for diagnosis and discussion than as specific evaluation tools, but that is changing; four missions have commissioned over-sampling of Afrobarometer surveys in the last round for the first time. This will provide excellent baseline data to examine change in the geographic areas of the programmes in the coming years.

USAID mission staff in the field and democracy practitioners in Washington alike welcome the surveys, and not only because they provide sound information for evaluation. The surveys are a multi-purpose tool. In fact, many missions consider the most important purpose of surveys to be programmatic, not evaluation—a vehicle for mobilizing support and discussion around democratic reforms. The national legislature, the press and NGOs often discuss and disseminate the results. They are also excellent diagnostic tools. In one recent case, for example, the surveys showed that voter turnout was low not because it was difficult to get to the polls but because women were less inclined to vote. The mission immediately switched its programme from increasing the number of polling places to efforts to improve the turnout of

women. The surveys are normally carried out every two years—in some missions, every year. There is a core definition of democracy based on concepts of legitimacy and political tolerance, used in every survey, as well as a set of core questions. USAID missions, with Professor Seligson and national institutions, have developed 'modules' of questions in particular areas of programming and high political interest, including local government, corruption and crime. This has allowed cross-national comparisons over time. There is also room for missions to add questions of more particular interest to their country.

In Washington, not only are the national results of interest, but the comparative analysis that develops from aggregating the results is valuable. Policy makers analyse trends in democracy, look for potential difficulties and opportunities, and assess the relative success of different kinds of programme throughout the region. US Government offices looking at illicit drug production, for instance, or the importance of issues such as the environment or health, can make use of them. Specialized reports on corruption and crime and other analyses of the survey data are often used by the host government and other donors in policy dialogue and reform.

The multi-purpose use of the surveys presents us with a winning evaluation methodology, with many kinds of users, and hence the willingness to invest in them on a regular basis.

Needless to say, democracy surveys have important limitations. They do not capture change in democratic institutions very well, particularly the slow, incremental improvement in transparency, capacity, and other bureaucratic attributes. Nevertheless they do capture a significant portion of USAID activities in a measurable way. They have become one of the most promising new methodologies to be incorporated into monitoring and evaluating USAID programmes.

Expert interviews: 'Voices from the Field'

SORA has piloted one more methodology to be used in conjunction with country and sectoral studies and large-scale quantitative analyses—structured, in-depth interviews of experienced USAID democracy practitioners. This endeavour, dubbed Voices from the Field, reaches for the opposite end of the continuum of research methodology to large-scale quantitative research. Rather than rely on worldwide longitudinal data sets and indicators, it focuses on extracting rich details through interviews with experienced democracy field officers.

The basic concept behind this research is that experienced experts can provide judgements and information that are not available through any other means. Expert interviews can be used to develop hypotheses that can be put to the test in wider settings. They help us discover what USAID in practice defines as the 'success' or 'failure' of a democracy intervention. They also provide insights on the complex interactions among whole classes of possible causes of democratic progress or backsliding. Such interviews also add texture and real examples to illustrate comparative findings. They help 'explain the story'.

Democracy and governance field officers live with the day-to-day reality of working with governments and civil society organizations and acquire deep knowledge of their capacities and commitment. They operate within the framework of US foreign policy priorities, and embassy and mission leadership objectives, and within a bureaucracy that experiences particular funding and personnel constraints. They plan and develop programmes based on in-depth assessments of democracy and governance issues in the country, and are forced to prioritize among alternatives based on their judgement of how these factors fit together. Democracy practitioners have a sense of sequencing and of the complex relationships among elements that are difficult to pick up in any other way. Moreover, they are responsible for establishing indicators to measure and assess progress on a regular basis, and for revising programmes during implementation if problems arise or initial judgements prove faulty. Their understanding of the relative importance of factors that encourage or inhibit successful democracy programming, and the relationships among those factors, thus offers a unique source of data for this evaluation programme. Most have managed programmes they consider successful and most can cite a spectacular failure or two.

Of course, any team of evaluators is likely to consult with the democracy officers in the field. The methodological challenge is to capture their knowledge in a systematic, comparative way. The protocol that USAID has developed and pilot-tested is its effort to do this. It has found respondents to be franker in an oral setting than they might be if asked to put their views in writing. In this research, the number of experts was limited to a group of officers who have served in at least two countries as democracy officers. It may later be expanded to include a sample of foreign nationals who have experience working in democracy offices in missions overseas.

The protocol examines in depth the experts' understanding of a 'democratic success' or 'failure', and what specific factors seem to be most important in determining success or failure. We can first determine if managers measure success primarily at the activity level or if they have a larger vision of democratization. In the latter case, they can usually articulate the desired 'democracy goal' and trace out the relationship between the activities or projects they support and that goal. This train of reasoning provides a rich vein of hypotheses that we can test in other settings. In addition, the protocol requires the experts to specify and prioritize the variables that are most important to programme success or failure.

The Voices from the Field methodology developed four categories under which to capture these variables. Under each category, respondents are asked what factors are the most important to success or failure, and to scale them. The scaling is an essential part of the protocol, forcing prioritization and allowing comparisons.

1. *General characteristics of the country apart from its path of democratization.* This includes factors such as the level of economic development, cultural and social conditions, historical precedent, international orientation, and conditions in the region. These variables were included based in large part on the variables analysed

in the quantitative study, which were widely discussed and examined.

2. *Country conditions related specifically to democracy and governance.* This includes regime type, government commitment to reform, institutional capacity, corruption, civil liberties, political competition, political inclusion, the power of civil society, and similar topics.

3. *The influence of the foreign policy priorities of the United States and other governments.* As a variable conditioning project success and failure, this factor has not been sufficiently studied relative to endogenous factors, considering that democracy promotion is one of the US Government's highest foreign policy priorities. This factor acknowledges that USAID is a US Government funding institution working under the State Department and according to its priorities, and, overseas, within and under the ambassador and the US Embassy. With the 2006 policy reforms in foreign assistance, under which the US State Department and USAID jointly determine even the 'sub-elements' of particular democracy projects, understanding whether and how policy concerns affect field success is important.

4. *Variables internal to the project or programme.* It is possible that such factors are more important in some cases than anything else in determining success or failure. This includes such factors as the level of funding, and its predictability, variability and sequencing. It also includes issues such as the quality of the project design, whether it is implemented through a well-defined contract or a looser grant mechanism, and the experience, capabilities and other characteristics of the implementing party.

The interview process explores these questions through analysing specific cases. Respondents are asked to identify the 'most successful programme' they worked on, and then to explain why they define it as a success—the characteristics of a successful programme. Later in the interview, they are asked to discuss their 'second most successful' programme in similar terms. In both cases, they are asked to work through the four categories of variables, ranking both the categories and the specific factors in each category. Finally, after some level of trust and openness has been established, they are asked to identify the least successful ('turkey') project they ever worked on, define what they mean by failure, and explain what led to it.

The combination of open questions, closed questions and scaling exercises is designed both to elicit the maximum information possible and to provide data that can be used for comparative study and feed into the other, broader, SORA methodologies.

SORA, Stage 3: the National Academy of Sciences and the future

By 2006 USAID had arrived at:

- aggregate analysis showing a positive impact of USAID democracy programmes, with some details about when impact was likely to greatest;
- findings from the SORA Stage 1 pilots on: (a) how to improve case study

methodology and the issues in undertaking future country assessments; (b) how to test for national and institutional level change; and (c) factors found to be important in the pilots for programme success or failure, which needed further testing;

- a democracy database updated annually on where and how USAID spends its democracy resources, which is of great use to the agency generally;
- growing support for and use of democracy surveys as a strong evaluation methodology, particularly with over-sampling;
- the piloting of more sophisticated interview techniques to capture the learning of democracy officers in the field;
- an identification of intellectual areas, such as 'definitions of democracy' (and therefore the measures of success of our programmes), that stubbornly remained under-researched, as well as questions on how to put together, both intellectually and logistically, the tools and findings of the endeavours to date.

The question was now how to put this complex work together to reach our basic goal. In concert with an academic Advisory Group,[14] we drew up a research/action agenda to move towards a method for fieldwork and research that would get us there. Writing this final 'scope of work' was an ambitious task that engaged the SORA team in USAID and the academic community for many months.

Who could carry out such an agenda? USAID needed an institution with autonomy and independence from USAID and with the ability to put together a first-rate intellectual product. The National Academy of Sciences (NAS) took on this challenge. It is now engaged in developing an overall research and analytical design for understanding the impact of democracy programmes on a country's development along key dimensions of democracy and the factors critical to programme success and failure. The NAS will be responsible for developing operational definitions of democracy and governance that can be used in the field. Based on the previous work reported in this chapter, the NAS will develop and test fieldwork research methodologies in the field to assess their validity and cost effectiveness. Finally, it will make recommendations on how the many pieces of SORA research can best be integrated into analytical products, developing an integrated final research and analytical design for consolidating the research elements.

The NAS committee is also focusing to some degree on 'prospective evaluations'— how to set up programmes so that they can best be evaluated, including, when possible, semi-experimental designs. Field missions are already signalling their enthusiasm for help in this area.

SORA has become more than a centralized series of comparative analyses. It has become a learning process, with improvements in evaluations and policy guidance every year. We are incorporating mission-led successes, such as the democracy surveys, and training and supporting missions elsewhere to undertake them. We are providing specific training to the Washington technical experts, as major 'diffusers of the culture', so that as they provide mission assistance, they improve programmes and evaluations of programmes.

There are significant intellectual challenges ahead to improve programming by better understanding what works and what does not work under what circumstances.

First, more bridging between academia and USAID practitioners needs to take place. Too often, information is conveyed only in one direction—from academics to practitioners. Recommendations for change are often too theoretical or too basic, rather like research doctors looking at a patient and basing their findings on what they already know rather than on a clinical, intensive knowledge of the patient in front of them. Interaction with the 'patient' is not only needed, but essential: USAID needs to strongly engaged in learning from its own experiences, and not to assume that academics can carry the task alone.

Second, the state of knowledge of democratic development is still not very advanced, and that impedes evaluation work. Until we do more to develop 'operational definitions' of democracy and understand the relationships among different elements of democratization apart from the impact of donor interventions, it will be difficult to measure the impact of democracy support on democracy or to identify how to use support resources most effectively. USAID has developed its own indicators of democratic change through trial and error; these need to be tested, particularly for long-term effects.

Third, while USAID has committed itself to SORA, it will be called upon to allocate even more resources over time to improving indicators, monitoring and evaluation and to keep up the training and technical assistance that are needed. In a resource-tight environment, programme improvement, evaluation and monitoring are often left behind.

Fourth, USAID has not had a culture of promoting and disseminating its own findings, either internally or, even more difficult, externally. SORA's approach depends on creating a learning cycle, with increasingly useful baseline data, monitoring and evaluation plans, and on dissemination and training. Essentially, at present the balance between programme implementation and analysis of what USAID has done is too heavily weighted towards the implementation side. With the huge investments that are now being made in democratization each year, it is time to re-balance this equation.

Notes

[1] Dr Hyman is in many ways the 'father' of SORA, providing the initial intellectual and administrative support to get it under way, and sustaining USAID commitment to its goals over many years.

[2] Lynn Carter, email correspondence, 1 April 2007. I am grateful to Lynn Carter for her comments and improvements in this section of the chapter.

[3] The Advisory Group included Robert Bates, Thomas Cook, Charles Kurzman, Gail Lecce, Dietrich Rueschemeyer, Mitchell Seligson, Brian Silver and John Tirman.

[4] The report was commissioned by the Social Science Research Council under John Tirman. The SSRC also commissioned a final revised version, written by Andrew Green, which was delivered in January 2004. This summary draws on both documents.

[5] This work was accomplished through the efforts of Andrew Green, as a democracy fellow in the Strategic Planning

and Research Division of USAID's Office of Democracy and Governance.

[6] Bruce Kay, a member of the SORA team in the Democracy Office at the time, was a major proponent and leader in bringing forward this approach.

[7] It may be worth noting that we undertook this quantitative study with serious misgivings about its usefulness. Even if it were to be successful, we did not think it could provide the kind of information needed, for example, to train democracy specialists in how to improve their work. Moreover, we had very serious doubts that USAID's relatively small programmes could move the large-scale democracy indicators of Freedom House and Polity, even if the programmes in reality were successful. We were very aware that we might find no relationship at this level between USAID's democracy work and broad indicators, and that such a finding could undermine support for democracy work, perhaps undeservedly.

[8] The panel included Michael Bratton (Michigan State University), Michael Coppedge (University of Notre Dame) and Pamela Paxton (Ohio State University).

[9] Steven Finkel, Mark Billera and David Black meticulously analysed this summary, making many useful additions and corrections.

[10] Using Polity IV data, the corresponding change is about a 33 per cent increase over the expected level of democratic growth for every 1 million USD.

[11] The model itself does not look at rates of change in Freedom House, but only the effect on the level of Freedom House scores in a given year. While we are now testing more cumulative-effect type models, at the present we cannot talk about doubling 'rates of change', but rather the effect on a given year. (Steve Finkel, email, 13 April 2007.)

[12] While many reasons have been suggested to explain what we hope was an anomaly, more research has now been commissioned on this specific topic.

[13] These surveys and others, as well as reports in Spanish and English based on them, are available through the Latin American Public Opinion Project (LAPOP) at the University of Vanderbilt run by Mitchell Seligson. They are fully described on the website www.vanderbilt.edu/lapop>. The data from the surveys are freely available for analysis, through the website, managed by the Centro Centroamericano de Poblacion (CCR) in Costa Rica, including an online querying system.
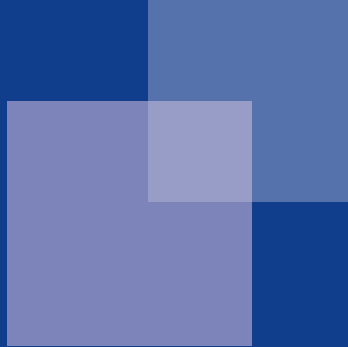
[14] From the donor community and academia, the Advisory Group included Guilan Denoux (Colby College), Larry Garber (New Israel Fund), Martha Gutierrez (Deutsche Gesellschaft für Technische Zusammenarbeit (German Technical Cooperation Agency, GTZ)), Philip Keefer (World Bank), Michael McFaul (Stanford University), Gerardo Munch (University of Southern California), and Mitchell Seligson (Vanderbilt University). Each provided written comments on the plans developed by the USAID SORA team (Mark Billera, David Black, Andrew Green and Margaret Sarles). They met at a workshop organized by Development Associates, Inc. with other USAID participants (Patricia Alexander, Ed Connerly, April Hahn, Kimberly Ludgwig and Keith Schultz) to synthesize the recommendations on the scope of the work for the National Academy of Sciences.

Chapter 3

Programme theory evaluation and democracy promotion: reviewing a sample of Sida-supported projects

Chapter 3

Fredrik Uggla*

Programme theory evaluation and democracy promotion: reviewing a sample of Sida-supported projects

This chapter presents a method for systematizing and evaluating the programme theories of a set of projects in the area of support for democracy. It is argued that this technique allows for the systematic consideration of assumptions and theories, and that such an assessment is particularly crucial in this area of development cooperation. While the method proposed cannot substitute for results-based evaluations, it can constitute a useful complement to such exercises and one that is particularly suited for diagnostic purposes. In order to demonstrate the practical applicability of the technique proposed, 52 Sida-financed projects drawn from Bolivia, Bosnia and Herzegovina, South Africa and Vietnam are subjected to such an evaluation.

Introduction

The Swedish International Development Cooperation Agency (Sida) is currently exploring new methods of evaluating projects and programmes in the sector of support for democracy and human rights. This chapter describes one of these initiatives which consists of a programme theory evaluation.

As is well known, previous attempts, both by Sida and by other development agencies, to evaluate their democracy support have incurred problems relating to the attribution of effects and the use of indicators to measure success. This chapter discusses an approach that attempts to circumvent such problems by focusing on the

---

programme theory underpinning efforts in the area of human rights and democracy, rather than on actual results. Although it cannot substitute for results-based enquiries, such an approach can provide a useful complement which for certain evaluation tasks may actually be of equal importance.

More specifically, the account presented here offers a simple model by which several projects can be brought together for systematic comparison and assessment, which then allows for a discussion about the prevailing practices and general assumptions within a field. Hence, rather than being a model for the evaluation of an individual project, what is proposed below is a model for evaluating and comparing intervention theories and programme logics across a set of projects.

In practical terms, this chapter brings together and examines the programme logic of some 50 Sida-supported projects in the area of democracy, good governance and human rights. The projects have been drawn from four countries that represent very different political contexts—Bolivia, Bosnia and Herzegovina, South Africa and Vietnam. At the same time, the four countries are among the largest recipients of Swedish support for democracy and human rights (as well as of Swedish development cooperation in general). In 2005, the year from which the information below is drawn, Vietnam was the largest recipient of such support, while South Africa was the fourth-largest. Bosnia and Herzegovina and Bolivia occupied places 11 and 12, respectively, on the same list.

The practical application of a programme theory evaluation to this selection of projects yields two kinds of results. In the first place, a number of points are made about the project logics thus discerned, about national variations, commonalities across countries, and the feasibility of the assumptions involved. Second, the findings indicate areas for subsequent study and evaluation. Because programme theory evaluation as applied here is primarily diagnostic, this exercise may be followed up by targeted results-based evaluations and studies directed at specific assumptions and mechanisms that are discerned with the help of this method.

Focusing on programme theory

Evaluability is a key problem in aid for democracy and human rights. In Carothers' view in Aiding Democracy Abroad (1999: 8–10), democracy promoters have tended either to 'under-do' evaluations, carrying them out haphazardly, or using superficial methods, or to overdo them, elaborating complex, rigid methods. Key problems in the area are a lack of suitable indicators and the difficulty of attributing effects to causes—and, hence, deciding what to measure and why.

Programme theory evaluation (also called theory-based evaluation, programme logic evaluation, and, by Owen and Rogers (1999), 'clarificative evaluation') provides a (perhaps defeatist) solution to some of the problems associated with results-based methodologies. In essence, this approach abandons the focus on results in order to study the underlying assumptions and rationales for the programme in question. The

reason for such an undertaking is relatively simple: projects may fail either because of problems related to their implementation, such as too little money or poor guidance and steering, or because the logic on which they were built was wrong in some way—for example, it had an unclear focus or was based on unrealistic assumptions (Pressman and Wildavsky 1979: 191). Programme theory evaluation focuses on this latter set of problems. What this technique considers is thus the theoretical basis for the programme in question, which is evaluated according to such concepts as realism, coherence and relevance. Necessary steps include the reconstruction of underlying theory and the assessment of its constituent parts, as well as their interconnections.

One could argue that the very features that make support in the area of democracy and human rights difficult to monitor and evaluate on the basis of results make assessment of the underlying programme theory essential for the development of successful programmes. As aid officials working in the area are generally unable to observe the actual impact of their efforts, they have to rely on informed guesses and some general assumptions if they are to devise and implement projects in this area. For instance, supporting a neighbourhood association in the name of democracy requires a number of explicit or implicit ideas about the possible impact that such a group might have on political decisions, about the aggregate result of a number of such groups pressing their demands on the state, and about the effects of a vibrant civil society on governance in general (see the discussions by Putnam 1993; and Avritzer 2002). Similarly, supporting a training programme in human rights for police officers requires several assumptions concerning, for instance, the reasons for human rights violations and how people react to training. In such cases and others like them, programme officers who will be unable to observe the eventual outcome of their efforts have to believe in the accuracy of such assumptions in order to justify commencing the projects.

Even though programme theory evaluation cannot aspire to capture the crucial issue of impact, it may serve to question and evaluate such ideas and assumptions. If a programme theory evaluation reveals unrealistic assumptions and unclear theoretical connections then it will have proved its usefulness as an audit technique.

Moreover, programme theory evaluation should be seen as a tool for learning (van der Knaap 2004). It has the ability to expose underlying assumptions, to bring central but tacit elements to the fore, and to critically examine theoretical foundations—all elements that are of critical importance for learning exercises. Moreover, the intended audience for such exercises can be both local project owners and financing agencies such as Sida. To a certain extent, the approach proposed here may thus do away with the exclusive focus on the putative project owners and executors in the field, which is common to evaluations. In fact, in the example below, it is Sida itself that is the focus of evaluation, not its partners.

Furthermore, while impact evaluation can only be undertaken after a programme is completed, and should preferably allow for a time-lag for all the effects to show, programme theory evaluation can be performed at any time during the project cycle.

This means that the learning phase can be integrated with the project implementation process and run alongside it to some extent. By involving intended beneficiaries and local stakeholders in such processes, the evaluation method proposed here can thus open the way for more participatory techniques, and for the concomitant gains of stakeholder involvement and commitment, and mutual understanding.

In some ways, the proposed technique comes close to the so-called logical framework approach (LFA, or 'logframe evaluations') as it shares with that methodology the systematic discernment of goals, actions, and the theoretical connections between them. The LFA is currently used in Sida as a standard tool for ex-ante evaluation of individual projects. In spite of the similarities, however, there are differences between the methods. While the logical framework approach is focused on different levels of goals, in programme theory evaluation the focus is on the mechanisms and actions involved in a project. Furthermore, what is proposed here is a model for aggregating and evaluating project logics across sets of projects, rather than for the consideration of individual projects.

Evaluating programme theory

Notwithstanding the potential benefits, previous experiences of programme theory evaluation have demonstrated a number of shortcomings. Even so, as is spelled out below, the approach has the potential to alleviate some of these problems.

A first problem with programme theory evaluation concerns the reconstruction of programme theories. In most cases, there is no explicit theory that can be distilled from programme documents. This means that evaluators have to begin their work by attempting to piece together such theories if they are to go on to test them later. In this regard, Frans Leeuw (2003) has proposed different methods for reconstructing programme theories. However, while his ideas deal with questions such as where one should look for theoretical statements, and who should be able to give views on the subject, they do not tell us how a scheme of analysis for such statements could be constructed. The absence of a model of analysis, in turn, gives the reconstruction of programme theories an ad hoc character when it comes to deciding what elements of a theory shall be included in the evaluation.

In order to counter such a problem, the present writer employs a fixed model of analysis which, although simple, offers us a guideline as to what elements of the theory should be collected/reconstructed. In addition to making the enquiry more systematic, the application of this model allows for comparisons between the programme theories and project logics of different projects.

A second problem concerns the question of how to judge and evaluate the programme theories that are discerned in the exercise. One obvious possibility would be to rely on juxtaposing the theoretical statements with scientific findings in order to determine how far they are relevant and correct. However, while such a procedure is perfectly reasonable, it is not problem-free. For instance, merely listing scattered evidence that

appears to contradict or confirm assumptions in the programme theory overlooks the fact that most findings in social science are seldom clear-cut and they cannot be applied across all the different contexts (see Haarhuis and Leeuw 2004). Indeed, academic work thrives on contradiction and counter-arguments, which complicates and limits the use of general scholarly findings as a standard for evaluation.

The presentation below offers a different method of analysis that is based on an attempt to discern patterns among a larger set of programme theories. Hence, the objective is not primarily to see whether the theoretical underpinnings of a study concur with social science findings in general, but rather to make explicit certain assumptions or theoretical patterns that recur across a large set of projects. (Of course, such prevalent ideas may subsequently be evaluated according to the extent to which they seem to agree with scholarly findings.)

The combination of these two techniques—using a fixed model of analysis and the comparative assessment of programme theories across a range of different projects—represents an improvement in the use of programme theory evaluation. As is demonstrated below, it allows for comprehensive treatment of a sector or a thematic area. By doing so, it can serve as a diagnostic tool and enhance discussions and learning exercises.

Discerning programme theory

Sida's portfolio of projects in the area of democracy and human rights amounts to hundreds of projects. To construct a manageable sample, four countries were selected for field studies—Bolivia, Bosnia and Herzegovina, South Africa and Vietnam. For these countries, all projects active in the area in April 2005 were initially selected. Subsequently, some had to be excluded due to lack of data or because they lacked relevant components (e.g. evaluative assessments). The final sample included 52 projects.

This sample was subsequently coded and systematized according to a model of analysis chosen to describe the programme theory of each project. The information used was principally drawn from the assessment memoranda that constitute the basis for decision making within Sida, and that are written by Sida staff for each project. Subsequently, interviews were performed with desk officers who had insight into the projects in order to validate the findings gleaned by studying the assessment memoranda.

It should be noted at the outset that there is no such thing as a 'Sida project' properly speaking, as Sida's role is limited to financing initiatives proposed by its partners. Nevertheless, the assessment memoranda do detail the appraisal made by Sida of the projects' relevance and feasibility, and hence the theoretical underpinnings for support. (Of course, there may be a reciprocal effect here also, in the sense that partners may frame and describe their proposals in terms that will appeal to Sida. In such cases, Sida also has an indirect effect on the projects. (For an interesting comparison see Bob 2005.)

Moreover in several cases the entire programme theory cannot easily be discerned,

either because certain parts of it are absent, or because they are not expressed in the documentation from the projects. This may not be as grave a fault as it sounds: to give a typical example, the account given of a project to support the development of a lobbying organization may not include a discussion of whether this organization will register an impact on public policy. However, for present purposes, natural or implied steps were assumed even when they were not expressed in the project documentation (in the example above, it would be assumed that the organization would target public authorities who would listen and alter their behaviour, for instance). As is noted above, such reconstructions were subsequently submitted to desk officers for validation.

The programme theories of all contributions were summarized and reconstructed in order to allow for comparison. This was done by applying a simple model of analysis that allows for the systematization of projects according to a common format. This model of analysis fundamentally depends on two chains—one of actors and one of actions. (The model is somewhat similar to the description of 'development pathways' discussed by Poate et al. 2000: 14ff. See also Jarstad 2005.)

The model of analysis

In international development cooperation, a donor contribution normally passes through one or several intermediaries before reaching the actual target population, which in most cases means people living in conditions of poverty. These intermediaries or actors could be seen as links in a chain between the donor and the target population.

It may seem beside the point to include the chain of actors in an assessment of the programme theory, but in fact it is vital to connect actions/mechanisms to actors, as the chain of actors reflects a number of assumptions about the correspondence between different actors' preferences. In effect, this is a chain of delegation, which in practice can face a whole series of problems to do with co-optation, goal displacement and the like.

In the most general of terms, it is possible to discern some common models of chains that respond to different needs and goals. For instance, if the goal is simply provision of a good, the chain may look as follows:

Provision chain: [Providers – executors – beneficiaries]

In a slightly more complicated model, the goal is not provision, but rather that a target group (for instance, a group of state bureaucrats) will start acting in a different way vis-à-vis the intended beneficiaries. For that to happen, another group may have to inform, teach, or support the target group. In this case the model may look like this:

Change chain: [Providers – executors – target – beneficiaries]

In some projects with a political component, it is possible to find an even more elaborate chain, which includes different sets of target groups with one attempting to

influence the other, while being changed (for instance, trained) itself. The following is an example:

Pressure chain: [Providers – executors – target (1) – target (2) – beneficiaries]

The relevant actors and the actions they are supposed to undertake are introduced into a scheme of analysis as in table 3.1, where a fictitious example illustrates a potential programme theory for a programme outlined according to the pressure chain. The model distinguishes between actions that take place within an actor (internal transformations, white fields in the table) and actions that are externally directed towards another actor (external transformations, shaded fields in the table).

Internal transformations are processes, mechanisms and changes that an actor has to undergo in order to connect to the next link in the chain of actions. The list of transformations includes categories such as 'Absorption of information/training', 'Absorption of arguments', 'Change in behaviour or attitude' and 'Internal reform'.

External interventions are actions that an actor conducts that are directed at another actor/other actors in the chain. Obviously, there is quite a long list of possibilities here. It includes categories that relate to education and information (e.g. 'capacity building', 'training', 'information campaigns'), material support ('financial contribution', 'material provisions'), and pressure ('advocacy', 'lobbying', 'litigation'). Other categories capture mechanisms that are not so clear-cut but are nevertheless crucial. Examples include 'demonstration effect' or 'spreading effects' that attempt to capture two possible mechanisms by which effects on the target group are supposed to reach the broader population.

For both these general sets of mechanisms, the present writer applied the coding according to an open model, in which new categories were added as they appeared.

Table 3.1: Programme theory model of analysis: a hypothetical example

Actor	
	Actor's external intervention
Donor/Sida	
	Financial contribution (to actor X)
Actor X	
	Training of actor Y
TG1: Actor Y	Learning, change of behaviour
	Application of new knowledge in treatment of actor Z, e.g. lobbying
TG2: Actor Z	Susceptibility to influence by actor Z. Change in behaviour
	Target group large/representative enough to allow for significant impact on society
Society	

Key: TG = target group. White fields in the table indicate internal transformations. Shaded fields indicate external transformations—actions that are externally directed towards another actor.

Thus, the programme theory of each of the projects in the sample was reconstructed to fit into this model of analysis. By converting projects to a common format, comparison and aggregation between different projects become possible. In the tables below, most of these aggregative measures are simple counts of the number of times a certain feature (an actor, a specific mechanism and so on) appears among the projects. Because of the nature of most projects, however, such counts can take very different forms. It is very common, for instance, for one particular project to feature several different target groups, and hence several different mechanisms. Below, most counts have been made according to whether a particular feature appears in a project. For instance, it is found that the internal mechanism of 'change in behaviour' is expected to occur in 21 of the projects, whereas 17 projects include elements of the external mechanisms of lobbying and litigation.

The countries studied

The four countries studied differ greatly in their political context.

- Bolivia is the longest-standing democracy of the four, and during recent decades has undertaken a number of institutional reforms, including decentralization and the creation of new institutions, such as a human rights ombudsman. These institutional reforms have not, however, prevented growing political instability, particularly in the form of potent social mobilization.
- Bosnia and Herzegovina represents a post-conflict case in which the most acute political task is to construct a viable state out of the institutions and regional autonomies created under the Dayton peace accord (1995), while promoting reconciliation of the main ethno-religious groups in the country. In this, the proximity of the country to the European Union gives it a special position.
- South Africa must be considered one of the most successful cases of African countries that democratized in the 1990s. Fears of violent conflict emerging out of the post-apartheid situation have largely subsided, as political and economic development has continued apace. Even so, however, the country faces a number of pressing problems as it tries to live up to the population's expectations of social and economic improvement.
- Vietnam remains an authoritarian dictatorship. In spite of economic liberalization and tentative steps towards more administrative openness, the Communist Party is firmly in power, and neither political opposition nor a civil society can be said to exist. Instead of democratization there have been a number of piecemeal institutional reforms, but it is uncertain how far these go in the direction of democratization.

Given these differences, an overall question of the analysis becomes the extent to which projects in the area of democracy support are contextualized, that is to say, whether their design responds to the political conditions in which they will operate. It is a common lament that projects for democracy and good governance are insufficiently

related to political conditions in the countries where they are supposed to operate (e.g. Carothers 1999: 338). In this regard, two hypotheses can be advanced. First, according to a convergence hypothesis we would expect to find a typical model of democracy projects that does not vary much between different countries. No matter how stark the country or regional differences, there is an expectation that they would be overcome by some preferred way of working. Second, one could speculate that certain factors such as contextual ones are bound to assert themselves in the design of projects in spite of the prevailing modes of operation, in which case one would expect to see a fundamental divergence between the logics of the projects in the different countries.

The variation between countries will be a recurring theme in the discussion of findings which follows. Such a discussion is not the only task, however. Of potentially greater importance is what this sample of projects may tell us about Sida's work with democracy in general. Thus, although the limited number of cases observed should inspire caution, the following pages will attempt to give a general overview of common patterns and variations in Swedish support to democracy promotion.

Comparing programme theories

Having converted the programme theory of each of the 52 projects into a single format for purposes of comparability, the assessment compares and aggregates these project logics. For reasons of space, the present assessment is illustrative only and cannot pretend to be exhaustive.

In the first place, the actor chain is examined and discussed. As is noted above, when it comes to executing agencies the elements of this chain rely on assumptions concerning the suitability of different actors, the concurrence in goals between different actors in the chain, and so on. For target groups, conversely, the structure of the actor chain demonstrates who is to be subject to the project intervention.

Second, the elements of the intervention chain—the mechanisms involved in the projects—are examined. In this regard, the frequency with which certain mechanisms are used will inform us about the degree to which different actions are judged to be feasible, what assumptions are held about how to best influence target groups, and what internal transformations are required in order for the projects to be effective.

Third, the analysis demonstrates how information from the two chains can be combined to tell us what changes and developments are expected from whom.

Fourth and finally, whereas the previous questions assume that the (reconstructed) programme logic is more or less clear, the material can also yield information about the extent to which that is true. The final part of this section thus examines the extent to which an elaborate programme theory is really present when it comes to ideas about broader impact.

The actor chain

The scheme of analysis applied here distinguishes between actors according to the different roles they fulfil in the project in question. Hence, supporters contribute funds or activities to the project, executors are charged with the actual performance of project activities, targets are actors that are supposed to change or alter their behaviour as a result of the intervention in question, and beneficiaries are the actors that are supposed to draw benefit from such changes.

It should be noted that these categories are not entirely mutually exclusive. For instance, it is relatively common to find that the same actor is included both as executor and as target. An example is where a state agency provides training for its own employees.

A first query, then, relates to what kinds of actor are charged with the different functions. Primarily, this is an issue for the two 'middle' functions of executors and targets. Supporting and beneficiary actors do not exhibit many differences; typically, Sida and other international agencies cover the first, while citizens in general are the beneficiaries in most cases.

In more analytical terms the distinction between executors and targets relates to who will initiate the relevant developments, and who will be targeted by such initiatives. Hence, while one would expect executing actors to hold a view of goals that corresponds to Sida's own (in order for delegation to work), this may not be a correct assumption in regard to the targets. On the contrary, and as is discussed further below, the targets can be seen as the actors that are to be changed through the project (either qualitatively, by fundamentally altering their behaviour, for instance, or quantitatively, by becoming better at what they do).

In order to answer the questions who is charged with producing change and who is the target of such efforts, the actors are divided into four categories: central state authorities; non-central and autonomous state authorities; national non-state actors; and international actors (typically consultants or Swedish authorities involved in 'twinning' exercises).

Table 3.2: Number of projects involving different types of actor in different tasks

	Bolivia c. state	Bolivia dec. state	Bolivia no-state	Bolivia internat.	Vietnam c. state	Vietnam dec. state	Vietnam no-state	Vietnam internat.
Execution	2	3	1	3	6	0	1	6
Targets	4	4	4	N. a.	8	3	6	N. a.

	BiH c. state	BiH dec. state	BiH no-state	BiH internat.	SA c. state	SA dec. state	SA no-state	SA internat.
Execution	3	3	2	13	2	6	8	6
Targets	9	12	15	N. a.	14	14	13	N. a.

Key: BiH = Bosnia and Herzegovina; c. state = central state authorities; dec. state = non-central and autonomous state authorities; no-state = national non-state actors; N.a = not applicable; SA = South Africa.

Contrary to what might have been expected, the central state authorities in all four countries are included not only as targets for interventions but also as executors of projects. This is particularly so in Vietnam, where a majority of projects rely on the central state bureaucracy for their implementation. This contrasts with the more democratically organized countries such as Bolivia, where only two such cases exist (support to enhanced management systems was implemented with the Bolivian vice-presidency). Similarly, in South Africa, central state authorities are rarely executors of projects, but feature frequently as targets of actions.

The high reliance on state agencies as executors of democracy projects in an authoritarian state such as Vietnam represents a paradox. It can be explained, however, by the simple fact that in a country like Vietnam there are very few possibilities of working outside the state. Indeed, there is only one project in Vietnam that relies on national non-state actors to execute the project, and it rests on non-governmental organizations (NGOs) that are supposedly autonomous and other associations in the field of gender.

With regard to targets, this category in a sense manifests assumptions about what actors need to be changed, strengthened or altered in order for democracy to be enhanced. Here again, Vietnam stands out, as most projects—although executed by state agencies—also have as their targets other state agencies. Conversely, in Bosnia and Herzegovina, most projects aim to produce change in society in order to enhance democracy (see table 3.6). Bolivia and South Africa present a more varied picture, with state and non-state actors more or less equally in focus. In this regard, however, the distribution of targets is not very surprising. In both countries the obstacles to advancing towards enhanced democratization can be said to be located in society as well as in the different reaches of the state. In contrast, the distribution of targets

in Bosnia and Herzegovina is influenced by the fact that the legacy of conflict still plagues society and constitutes a primary obstacle to democratic development. In Vietnam, the fact that the primary obstacles to reform are located in the central state is perfectly understandable.

Going into even more detail, it is possible to use the actor chain to separate top–down from bottom–up approaches (see table 3.3). The former involves central state agencies as executors and decentralized agencies or social groups as targets; the latter involves the reverse situation (decentralized agencies and social groups as executors and central state agencies as targets). Accordingly, in the top–down cases, central state agencies try to change the practices of institutions and organizations at lower levels. Bottom–up approaches feature the reverse order of things—assisting more or less autonomous groups to influence the central state agencies.

Table 3.3: Number of projects featuring top–down and bottom–up approaches

	Bolivia	BiH	SA	Viet Nam
Top–top	1	1	1	5
Top–down	1	5	1	3
Bottom–up	2	4	9	1
Bottom–bottom	3	10	13	2

Key: BiH = Bosnia and Herzegovina; SA = South Africa.

It is very possible that all support for democratization requires working with both the state and society. However, the findings here indicate a variation between the countries that is of some interest. Most notable is the reliance in Vietnam on the central state authorities as both implementers and targets. Also, it is evident that in Bosnia and Herzegovina quite a large number of projects feature central state organs in an executive capacity, and groups at lower levels as targets. This is a surprising finding given the decentralized nature of the country, but it corresponds to a large number of projects in which a centrally located institution receives support for strengthening its relationship with local entities.

South Africa represents the converse situation, in which support overwhelmingly goes to groups and actors that are independent from the central state, and which act either vis-à-vis the organs of the state or vis-à-vis other parts of civil society/decentralized agencies. (A large number of projects contain simultaneous actions vis-à-vis both the 'top' and the 'bottom' level. An example is support to land organizations that simultaneously lobby the government authorities and attempt to educate smallholders.)

Again, one should note that the question of who attempts to change whom relates to a broader question of the initiators and objects of changes in a democratic direction.

In Vietnam, the fact that most projects target central state authorities corresponds to the political situation in the country. However, that initiators are to be found at the same level may be more surprising, as it appears to rely on an assumption that there is a real willingness at that level to pursue democratic reform. Whereas the present assessment cannot, of course, vouch for the correctness of such an assumption, it can at least make the assumption explicit, and thereby facilitate further discussion of how accurate it really is.

In sum, the above analysis helps us appreciate some fundamental differences in how projects in the area of democracy support are conceptualized. Whereas in Vietnam a typical project consists of central state agencies attempting to influence other parts of the bureaucracy at the same level, projects in South Africa are much more likely to include elements of society attempting to influence the state. Conversely, in Bosnia and Herzegovina most projects aim to produce change in society, and the initiators of such projects are found at the level of central authorities, in society itself, and among international consultants too.

Mechanisms

The discussion so far has relied on an analysis of the actor chain. Moving on to the intervention or action chain, attention shifts from who the participants in the project are, to what is supposed to happen within it, and the assumptions that are embodied in such assessments. As is noted above, the scheme of analysis differentiates between internal and external effects and transformations.

External effects represent the attempts at influence between the different actors. In particular, interest here is on the mechanisms that make up the relationship between executing and target groups. Table 3.4 illustrates some of the mechanisms. It is relatively rare for them to be used in isolation; more typically projects tend to include several different methods of influence.

To this effect, the analysis distinguishes between a number of different possible mechanisms, ranging from the provision of thematic expertise through the placement of international experts (e.g. providing Swedish experts to carry out a study of corruption in Vietnam), to support of court litigation against government authorities (e.g. by supporting the Treatment Action Campaign's work to make the South African Government distribute antiretroviral HIV treatment), and capacity training in a number of areas (such as informing Bolivian public servants about new laws and regulations).

Table 3.4: Number of projects that contain different external mechanisms (fractions of total in brackets)

	Total (52 projects)	Bolivia (6 projects)	BiH (17 projects)	SA (19 projects)	Viet Nam (10 projects)
Thematic expertise/ external consultant	15 (.29)	0 (.00)	4 (.27)	7 (.37)	4 (.40)
Information campaigns	19 (.36)	2 (.33)	2 (.20)	9 (.47)	6 (.60)
Capacity building, training	34 (.65)	3 (.50)	11 (.82)	12 (.63)	8 (.80)
Twinning, international exch.	8 (.15)	0 (.00)	2 (.13)	1 (.05)	5 (.50)
Advocacy/lobbying/ litigation	18 (.35)	1 (.17)	4 (.27)	8 (.42)	5 (.50)
Material, financial support	17 (.33)	3 (.50)	5 (.33)	4 (.21)	5 (.50)

Key: BiH = Bosnia and Herzegovina; SA = South Africa.

As can be seen in table 3.4, training dominates as the instrument of choice in democracy support to these four countries. Indeed, two-thirds of all projects contain elements of capacity training. The remainder of the mechanisms considered here are involved in between 30 and 40 per cent of the projects. The special case of international exchange/'twinning' features in only 15 per cent of the initiatives.

The general picture is thus rather eclectic; a variety of mechanisms are employed. Furthermore, no country appears to stand out very much from the average. In practice, this means that the same general mix of policy instruments appear to be used in all four countries. Given the different political circumstances and the fairly large variation in regard to actors, this is rather surprising. It would also seem reasonable to expect that needs differ in the four countries. For instance, while material support may be more called for in one case, information and training could be the primary deficiency in another. Yet if such differences exist they are only weakly reflected in the data. True, capacity building and training appear to be more common in the countries with the least experience of democratic practices, and material provisions are more frequent in the poorer countries. Beyond this, however, there is no clear-cut division with different mechanisms being employed in different countries. Certain elements in table 3.4 are even counter-intuitive in this regard. For instance, 'twinning' and exchange with Swedish counterparts are more commonly used in the countries in which bureaucratic practices could be expected to differ most from Swedish ones—in Vietnam and in Bosnia and Herzegovina. (One should also note that there may be an element of concept stretching in table 3.4. In particular, the fact that five projects in Vietnam involve mechanisms for lobbying and advocacy may seem surprising, but in most of these the mechanisms are of a 'top–top' kind, that is, they involve one part of the state attempting to influence another part. Only

in one Vietnamese project does this particular mechanism consist of social actors trying to influence the state.)

It is often simpler to trace the external effects that are supposed to take place in a project than the internal changes that are supposed to occur. Examples of such mechanisms are the creation of more tolerant attitudes among young people in Bosnia and Herzegovina, and the encouragement of increased openness towards the public among Vietnamese bureaucrats that a media support programme is supposed to create. Despite their often diffuse character, such mechanisms and transformations are no less important for the programme logic to work. The frequency with which some of these mechanisms occur is displayed in table 3.5.

Table 3.5: Number of projects that contain specified internal effects (fractions of totals in brackets)

	Total (52 projects)	Bolivia (6 projects)	BiH (17 projects)	SA (19 projects)	Vietnam (10 projects)
Absorption of info/train.	39 (.75)	3 (.50)	15 (.88)	13 (.68)	8 (.80)
Changes in attitudes	15 (.29)	1 (.17)	5 (.29)	5 (.26)	4 (.40)
Changes in behaviour	19 (.36)	1 (.17)	7 (.41)	7 (.37)	4 (.40)
Internal reforms	22 (.42)	3 (.50)	7 (.41)	9 (.47)	3 (.30)

Key: BiH = Bosnia and Herzegovina; SA = South Africa.

As can be seen in table 3.5, there are some typical expectations concerning what the projects are supposed to contribute to. In keeping with the stress on capacity building and training above, a majority of projects involve assumptions concerning the absorption and application of information. In comparison, other expectations about the internal processes that are supposed to occur are less frequent. This might be interpreted as evidence of a view of change as being primarily a matter of absorbing information.

However, there are also some differences between the countries. In this regard it is instructive to compare Bolivia with the other cases (even though the relatively small number of projects in Bolivia in the sample makes such comparison somewhat uncertain). In Bolivia, what appears to be expected is internal reforms to enhance efficiency and so on, rather than any reorientations with regard to attitudes and behaviour. In the other three cases, it is changes in attitudes and behaviour that are much more frequently stressed. Thus, whereas in Bolivia support goes to ongoing processes of reforming the state, in the other cases the goal seems instead to be to make it perform in a different way, namely, more democratically. In line with what might be predicted, this tendency is also stronger in Bosnia and Herzegovina and Vietnam than in South Africa, where internal reform is more commonly stressed.

Actors and mechanisms combined

Finally, we should ask how far the different assumptions are applied to different actors. How are the two chains combined? Are certain mechanisms only applied to certain actors? Is it possible to find a model according to which different types of actors receive different kinds of incentive to change or develop?

Table 3.6 presents a selection of mechanisms. These include three of the most commonly assumed internal transformations along with three different instruments of external influence. To a certain extent, the latter correspond to the metaphorical 'carrots, sticks and sermons' that constitute the tools of leverage (Bemelmans-Videc et al. 1998). In political terms, these categories translate into material support; information and training; and lobbying and litigation. Each of them can be said to embody implicit and explicit assumptions about the deficiencies that an actor has and, accordingly, what instruments will improve that actor's performance. Hence, material support must reflect a conviction that the existing material conditions are not conducive to democratic governance. Similarly, the provision of information and training reflects an assumption that lack of knowledge is a primary obstacle.

Table 3.6: The number of projects in which specified effects are supposed to occur, below the executive level

	Bolivia. c. state	Bolivia. dec. state	Bolivia. society	Vietnam c. state	Vietnam dec. state	Vietnam society
Changes in attitudes (IT 3)	1			3		
Changes in behaviour (IT 4)	1			3		1
Internal reforms (IT 6)	2	1	1	2		
Material and financial support	1		3	4	2	
Information and training, twinning	1	2	2	4	3	5
Lobbying and litigation, popular participation	1			5		

	BiH c. state	BiH dec. state	BiH society	SA c. state	SA dec. state	SA civil soc.
Changes in attitudes (IT 3)			5	3	1	2
Changes in behaviour (IT 4)	2	1	7	3	2	4
Internal reforms (IT 6)		2	3	3	5	2
Material support		4	1	3		1
Information and training, twinning	3	6	8	9	9	11
Lobbying and litigation, popular participation	4	4		8	6	

Key: BiH = Bosnia and Herzegovina; c. state = central state authorities; dec. state = non-central and autonomous state authorities; no-state = national non-state actors; SA = South Africa.

Table 3.7: Summary of external mechanisms employed

	Central state total	Decentralized state total	Society total
Changes in attitudes (IT 3)	7	1	7
Changes in behaviour (IT 4)	11	3	12
Internal reforms (IT 6)	7	8	6
Material support	8	6	5
Information and training, twinning	17	20	26
Lobbying and litigation, popular participation	18	10	0

Unfortunately, few clear conclusions leap out from tables 3.6 and 3.7. True, information and capacity building ('sermons') appear to be employed most frequently vis-à-vis actors in society, just as lobbying and pressure ('sticks') are used vis-à-vis the state. Indeed, it is striking that actors beyond the state are seldom offered anything but training. In comparison, material support ('carrots') features only rarely. With regard to internal changes, these are expected to occur with equal frequency in state and society.

When country variations are taken into account some differences do become clearer. Some of the patterns noted above reappear here. It is interesting to note, for instance, the difference between Bosnia and Herzegovina, where social actors are the ones who are supposed to alter their preferences and behaviour, and Vietnam, where such changes are expected to occur in the central state. As is discussed above, such differences appear to correspond to the political realities of each country.

Inexplicably, assumptions about the susceptibility of the central state organs to change are most frequent in the two polar opposites, namely Vietnam and South Africa, whereas such assumptions feature less in the cases of Bolivia, and Bosnia and Herzegovina.

Lack of assumptions

The analytical framework employed in this chapter also allows us to gauge any assumptions or links that are missing from the proposed causal chain. While these have been filled in to a certain extent in the discussion above, we could add an analysis of what is rarely, if ever, discussed in Sida's appraisals of the projects it supports.

Generally speaking, much more argument and thinking appear to go into the first steps of the chain, that is, the relationships between supporting and executing agencies, and the primary target levels. Little discussion takes place of how the project is supposed to impact on the population in general. Here the assumptions are seldom explicit.

Of course, the fact that ideas about how impact is to be achieved are seldom explicit should not automatically be interpreted as evidence of a lack of thinking or

lack of theoretical support for a project. After all, there is enough evidence to support tacit assumptions concerning issues such as the importance of a strong civil society for democracy, or the positive effect a human rights ombudsman can have on the rule of law, and so on. But it is also true that such effects are seldom automatic, and to the extent that the absence of discussion about how to make a broader impact indicates an absence of thinking in this regard, this amounts to a problem. Table 3.8 shows the frequency with which some possible broader effects are included and discussed in projects.

Table 3.8: Impact made explicit: the fraction of projects that contain discussions about certain mechanisms related to impact beyond target group level

	Bolivia	BiH	SA	Vietnam
Enhanced service provision	7/8	6/12	10/22	4/5
Demonstration effects	0/1	0/7	0/4	0/3
Change in preference, attitude or behaviour	N.a	4/10	0/3	6/8
Make use of offered opportunities	N.a	2/13	0/8	9/12
Absorption of argument or information.	0/2	2/6	0/7	7/12

Key: BiH = Bosnia and Herzegovina; N.a = not applicable (mechanism not present in the projects considered); SA = South Africa.

As can be seen in table 3.8, there is both thematic and geographic variation in how far projects contain discussions about certain mechanisms related to impact beyond the target group. For instance, projects concerning Vietnam are typically much more developed in this regard, which may be because they are designed for a politically more difficult environment. Conversely, projects in the two more democratic states seldom contain much in the way of explicit thinking about impacts, beyond the effects that can be tied to enhanced service provision.

Interestingly, certain mechanisms are much more frequently discussed and problematized than others. For instance, issues and questions connected to service delivery figure strongly, whereas the capabilities of target groups to absorb arguments and to effect changes in preferences and behaviour are mentioned much less frequently. In particular, there is no example of demonstration effects (that is, of a project having a broader impact by influencing sectors of the population beyond the target group) being discussed in even the most superficial manner. However, only assessing the extent to which mechanisms are explicitly indicated does not allow us to assess patterns

of more profound thinking concerning impact. An even stricter evaluation would be to distinguish the projects that contain an elaborate discussion on their mechanisms of impact. If this criterion is applied, a rather discouraging picture emerges, although one that reveals important differences between countries. Such elaborate discussions about impact appear in seven out of ten projects in Vietnam and in two out of six projects in Bolivia. In the cases of South Africa and Bosnia and Herzegovina they feature even less.)

General findings

Assumptions and arguments

As is repeatedly indicated above, the different mechanisms involved in the projects relate to assumptions about the possibilities and feasibility of effecting certain actions. For instance, using lobbying as a strategy implies assumptions regarding the susceptibility of the targets to such actions, just as 'twinning', or the exchange between a Swedish and a local body, relies on assumptions concerning the transferability of experiences, the power of example, and so on. Similarly, using capacity training as a mechanism must build on the assumption that a lack of capacity is the principal weakness or one of the principal weaknesses to be addressed. The evaluative aspects of the present exercise relate primarily to such assumptions. What follows are examples of how these findings can be assessed and used for evaluation purposes.

1. Assumptions can be discussed on the basis of how realistic they are. In this regard, it is worth noting that certain mechanisms appear more likely to attain their objectives than others. Typically, one would expect changes in attitudes and behaviour to be more difficult to effect than the simple transfer of material provisions, for instance. Even so, it is notable that more projects rely on the former mechanisms than on the latter. While this may be perfectly justified, one could juxtapose the assumptions underlying such a distribution with what is known about how susceptible people are to different forms of influence, for instance.
2. Mechanisms can be related to the context in which they are supposed to work. For instance, it is noteworthy that different forms of training and provision of information are the most commonly employed mechanism in all four countries. But information is only one link in a chain that typically depends on the target actors' ability to digest information, to act accordingly, and subsequently to have an impact on their broader context. Furthermore, such a mechanism necessarily implies the assumption that the primary obstacle in the way of democratic development is a lack of knowledge, rather than, say, something to do with the

distribution of power, or political conflicts, which political analysts may be more inclined to see as the major cause. Such assumptions can be assessed on the basis of what is known about the local context.

3. (Related to the previous point), the relatively frequent use of certain instruments in some countries could serve as a discussion point. This is exemplified in the 'twinning' exercises, which rely on a number of assumptions about the transferability of experiences and ideas across contexts. While there is nothing extraordinary about that, what is surprising is that this mechanism is more commonly employed in the two cases that are possibly the most far removed from the conditions in which the Swedish bureaucracy operates, namely Vietnam and Bosnia and Herzegovina.

4. The absence of discussion of certain links—particularly links related to broader impact—is worrying. If that corresponds to a real absence of thinking about such issues, the potential effectiveness of the projects has to be called into question. Regrettably, these findings resonate with a previous study of a sample of Sida projects in the area of democracy and human rights, which found that '[t]he projects reviewed were very weak in specifying assumptions that would allow the activities to be convincingly linked to the goal' (Poate et al. 2000: 74).

5. The results above can provide an answer to the initial question about convergence and divergence. Unfortunately, the answer is not as clear as one would hope: there are both common and different elements between the countries; rarely are particular mechanisms and actors completely absent. That said, the findings above indicate that, while the variation with regard to actors is quite substantial, the mechanisms employed exhibit much greater similarities across countries. It may be that the choice of partners is more conditioned on the local context, compared to the selection of interventions and actions to be undertaken in the projects.

How are we to use the results?

In a sense, the results reported in this chapter amount to a rough description of the programme theory and project logic of a sample of democracy promotion projects supported by Sida. It should be noted, however, that, apart from general discussions of the kind just undertaken, a number of more rigorous evaluative activities (discussions, targeted evaluations, or academic studies even) could be planned on the basis of these findings. Such activities would serve both the purpose of control and that of learning.

With regard to control, it is possible to subject the findings to an evaluation of the feasibility and realism of the assumptions involved. Thus the assumptions involved may be juxtaposed with what is known about certain mechanisms (Haarhuis and Leeuw 2004; see also Pawson 2002 for an interesting perspective on how to perform such a juxtaposition). For example, the importance of assumptions concerning

training in different forms could be evaluated on the basis of what previous studies and evaluations by, for example, Finkel (2003) and Blair (2003) have found about them. Similarly, assumptions about how certain forms of behaviour spread in a polity should take account of studies of the critical mass that may be necessary to sustain such behaviour (Axelrod 1987).

Alternatively, the findings above could be used to design studies to specifically test certain assumptions. Given that lobbying and public pressure appear to be mechanisms frequently used against the central state authorities in South Africa in particular, a separate study could be commissioned to test the susceptibility of the South African state to such measures.

Of course, the information contained in such studies could also be used for learning purposes. In fact, one of the principal uses of the kind of evaluation presented above is diagnostic. For instance, the finding that over half of the projects surveyed contain elements of capacity building could lead to an investigation into the experiences of such elements, of whether there are Sida-specific factors that are responsible for such a focus, and of the alternatives. In this regard, it is informative to make comparisons both between Sida's experiences of different countries and between the experiences of different bilateral development cooperation agencies.

In sum, it should be stressed that the programme theory evaluation sketched above constitutes only a first step. Further inquiries must be undertaken to turn the findings into operationally useful results. Even so, the example here has shown how the systematic, comparative approach used for analysis can bring forward several important points that merit further investigation.

Conclusion

This chapter has attempted to state the case for using programme theory evaluation techniques in the area of democracy promotion. This method is no panacea for the problems associated with evaluating this area of development cooperation. Importantly, programme theory evaluation can overcome neither the problem of attributing causal influence nor that of defining what will be counted as success in such endeavours. It is unable to say anything about actual conditions, or accordingly about achievements or the obstacles that projects have actually faced. In that sense, the method proposed here can never substitute for studies and evaluations performed on the ground, so to speak. Moreover, aggregating several projects as has been done above includes a risk of concept-stretching or, at any rate, a certain arbitrariness in designing the categories.

Even so, the kind of systematic, comparative evaluation of implicit and explicit project theories and logics demonstrated in this chapter can serve important evaluative purposes. Some of these uses have been demonstrated above; others have simply been hinted at. This concluding section briefly recapitulates.

Programme theory evaluation can help discern weak points and unsustainable

arguments and assumptions in project design. In particular, the approach suggested here allows for the systematization of arguments and assumptions across a range of projects, which in turn serves as a diagnostic tool. Insufficient discussion concerning impact mechanisms was given as an example. And, although the present exercise could not provide a definitive judgement on these issues, it did call into question the use of 'twinning' arrangements in certain contexts, and the problems associated with relying on state authorities for executing democratization projects in settings that were not very democratic.

Although the sample drawn on in the chapter is too small to be representative, some tentative points can be made concerning prevailing modes of action in Sweden's support for democratization. For instance, it has been demonstrated that, while the choice of partners and targets differed between countries, the actions undertaken within the projects exhibited much less variation. In particular, the most salient finding is possibly that training and capacity development appear to be used as a treatment for all ills no matter what the political context.

We have identified a number of points and areas for future discussion, and, perhaps most important, for future studies and evaluations. For instance, given the importance attributed to mechanisms connected to training and to applying different pressure techniques vis-à-vis central state organs, we badly need studies that can give clear indications of the extent to which this and the related assumptions are justified, the conditions under which such interventions are likely to succeed, and the obstacles they may face.
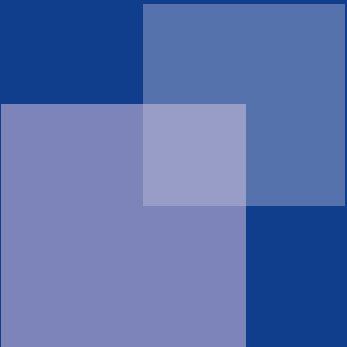
As the last point makes clear, the account of programme theory evaluation described here is but a first step which needs to be followed and complemented by other kinds of studies and evaluations.

Chapter 4

Progress and myths in the evaluation of the rule of law: a toolkit for strengthening democracy

Chapter 4

Sandra Elena and Héctor Chayer

Progress and myths in the evaluation of the rule of law: a toolkit for strengthening democracy

This chapter discusses the conceptual and methodological difficulties associated with the evaluation of rule-of-law programmes and proposes an evaluation methodology designed to overcome these difficulties. The FORES Evaluation Toolkit prescribes five evaluation phases. FORES' experience of evaluating the World Bank's major judicial reform initiative in Argentina demonstrates that each phase in the toolkit provides information that complements and facilitates the interpretation of data from other phases. FORES' experience with a participatory evaluation methodology in the Rio Negro court reform programme demonstrates the cost-effectiveness of the methodology when the beneficiaries of the project implement the evaluation themselves, with training and oversight by local professional programme evaluators. The Judicial Reliability Index, developed by FORES and various partner organizations, offers a powerful way to measure the overall effect the conglomeration of rule-of-law programmes in a country has on the legitimacy of public institutions. This index is also useful in completing the fifth phase of the FORES Evaluation Toolkit. The chapter concludes with recommendations for donors, evaluators and organizations interested in democratic development regarding effective rule-of-law programme evaluation.

Introduction

The aim of this chapter is to deepen the debate surrounding the process of evaluating democracy programmes. Democracy programme evaluations are essential to understanding the outcomes and impact of programmes. Given the range and number of the variables that affect democratic institutions, measuring the effect of individual rule-of-law (ROL) programmes is challenging. This chapter proposes an

evaluation methodology that helps attribute effects to specific programmes in the ROL area. Broad implementation of this evaluation methodology would ensure that consideration is given to programme impacts at the programme design phase, encourage key actors to participate, and systematize ROL evaluations.

The theoretical methodology shared here is drawn from lessons learned through the evaluation experience of FORES—the Foro de Estudios sobre la AdministraciÛn de Justicia (Forum for Studies on Judicial Administration). FORES is an Argentine non-governmental organization (NGO) that has been working in the ROL field in Latin America for the past 30 years. FORES' main areas of expertise are the training of judicial actors, providing technical assistance to Argentine and foreign judiciaries, legal research, the organization of judicial seminars, advocacy and public opinion research, and the evaluation of ROL programmes.

This chapter is a compendium of practical experiences and insights gained in dealing with ROL programmes that the authors have assembled over several years. It is written from the perspective of an NGO with a long-term involvement with local civil society and may therefore differ from the views of donors or bilateral and multilateral agencies, as well as from those of local politicians and authorities. FORES, in its double role as implementer and evaluator of ROL programmes, is in a unique position to understand and analyse successes or failures in the current evaluation methodology, and to make recommendations on how to improve the current state of the art.

The recommendations presented here focus on the evaluation of the outcomes and impact of ROL programmes. They are not intended to address issues related to the design and implementation of evaluation, nor with financial evaluation or the disbursement of budgets. Instead, the focus is on evaluating how specific programmes impact on the strengthening of democracy.

We understand the evaluation of 'outcomes' as the analysis and comparison of the proposed and actual results of a programme. It should be noted that outcome and output are conceptually distinct. Outcome refers to the effects of a programme, while output refers to the specific products delivered by a programme. Meanwhile, the impact of an ROL programme refers to how the programme affects people beyond the group of its direct beneficiaries.

Different perspectives: evaluation practice in the public sector

The evaluation of public policy programmes is relatively new. In the United States during the 1970s, a growing need for information about the outcomes of public-sector programmes spurred evaluation of education and public health programmes. Presidents John Kennedy and Lyndon B. Johnson encouraged the extension of evaluation to other areas of public policy too. In Europe, evaluation efforts began later due to a different conception of the role played by the government and state.

In general terms, the concept of evaluation arises following a change in ideas

about government and public administration. The traditional understanding of public administration conceptualized it as a set of procedures that should be followed, regardless of the results. Today's conception of public administration, which developed around 1970, is more results-oriented. Under this new paradigm, the 'products' delivered by state agencies are known as 'public services' and the citizens receiving them as 'clients'. Along with this new paradigm, two new ideas became important— programme results; and quality of service (Boix 1992).

In the past 30 years, most Western countries have adopted evaluation procedures for public-sector programmes. Meanwhile, international organizations such as the World Bank, the International Monetary Fund, the European Union, and in Latin America the Inter-American Development Bank (IDB) are important promoters of evaluation. All of them include evaluation clauses in their loan and grant contracts, and have internal evaluation offices. Following this example, other donors such as bilateral aid organizations, international foundations, institutes and universities also request evaluation procedures as a condition for their support.

In most developing countries, the evaluation of public policy was promoted not by local stakeholders but rather by international organizations. This adds a sensitive political dimension to public policy evaluation: it may be seen as a way of supporting 'foreign control' over local institutions.[1]

Despite the rapid growth of evaluation in the democracy field, many theoretical and methodological issues remain unresolved. Widespread use of evaluation procedures in democracy support programmes in Latin America is coupled with a lack of consensus as to what 'evaluation' means. The term 'evaluation' is used in different ways to communicate a variety of meanings, some of which fall outside the understanding held by academics and practitioners in the field.

Other, no less ideological, questions that arise include such basic concepts as 'What are we going to evaluate?'. This simple question invites many different answers. We may evaluate the design and implementation of a programme, or its outcomes; or we may limit ourselves to considering whether the budget was properly disbursed. The answers to the question 'What is an evaluation and what should it include?' can be divided into various typologies of evaluation. In general terms, the typologies include evaluation of needs, of programme design, of programme implementation, of programme reach, of programme outcomes and programme impact, and financial evaluation of the programme.

The question 'How are we going to evaluate?' is subject to even broader interpretation. The methodology for evaluating a programme varies from country to country, programme to programme, and donor to donor. Sometimes evaluation is an opinion based on expert observation. At other times it is the analysis of hard data collected through social research methods.

Answers to the question 'How are we going to evaluate?' can be summarized in four main alternatives. The first is the traditional approach of evaluation by objectives. This advocates following five steps: (a) the specification of objectives;

(b) the specification of a list of objectives in order of importance; (c) the selection of tools for measuring the outcomes; (d) the collection of data; and (e) comparative analysis of the data. A second alternative incorporates applied research to determine the effectiveness of a programme, while also trying to understand: (a) the reasons for success or failure; (b) the programme philosophy; and (c) a redefinition of means to accomplish the goals. A third alternative includes four different kinds of evaluation: (a) context evaluation; (b) input evaluation; (c) process evaluation; and (d) outcome evaluation. This perspective is more systematic and global because it adds the analysis and understanding of needs to the equation. A fourth and newer perspective is client/ user-oriented, and completely changes the focus of evaluation. It takes into account the real impact the programme has on its clients, and evaluates the programmes according to the clients' needs and values.

As to the mechanisms or tools for collecting information for evaluation purposes, a nearly infinite number of variations exist. The most important are: (a) documentary analysis and the use of secondary sources; (b) surveys; (c) focus groups; (d) the collection and analysis of hard data; (e) interviews; (f) *in situ* observation; and (g) committees of experts. The characteristics and pertinence of each are discussed below.

The theoretical problems exist despite, or perhaps because of, the rapid proliferation of democracy programme evaluation. The number of evaluations has increased along with the number of democracy programmes. ROL is a relatively new field in comparison to other democracy-related areas. It started as part of the United States' aid programmes in the mid-1980s and spread throughout countries in receipt of US aid in the 1990s. This field is still expanding regionally, as is the number of topic areas it includes.

ROL programmes work under the assumption that the rule of law is necessary for economic development and for democracy. If a country does not have effective ROL, it does not attract foreign investment and will not be able to finance development. Under these assumptions, economic development is a requirement for a strong democracy, and vice versa.

However, these arguments have been challenged by some notable ROL academics and practitioners. One of the more prominent is Thomas Carothers (2006c: chapters 1 and 2), who argues there is a notable lack of proof that a country must have a settled, well-functioning rule of law in order to attract investments. He points out that China is capable of attracting considerable foreign investment despite its notorious lack of Western-style rule of law (Carothers 2006c: 17–8), and explains that a good number of ROL practitioners share his concern about the lack of knowledge in the ROL field. He claims that, when pressed, practitioners admit that the base of knowledge from which they are operating is startlingly thin, and cites an ROL expert who has worked for many years in Latin America as saying that 'we know how to do a lot of things, but deep down we don't really know what we are doing'.

Carothers also asserts that even in established democracies—those supposed to be emulated by developing countries—a number of shortcomings in the rule of law

exist. These include (a) overloaded courts that delay justice; (b) lack of adequate judicial remedies, particularly for minorities; (c) a criminal justice system that often punishes minorities more severely than members of the majority population; and (d) politicians who abuse the law. As a result, it is more accurate to say that the rule of law and democracy are closely intertwined but major shortcomings in the rule of law often exist within reasonably democratic political systems.

One of the reasons Carothers offers for the lack of information on the subject is the little attention and support that aid organizations give to applied policy research. Aid organizations are more action-oriented and usually consider research a waste of resources.

In fact the concrete effects of ROL programmes in the overall development of the rule of law in a country very often remain uncertain. The lack of lessons learned extends to ROL evaluation as well. Linn Hammergren says that

> while it is commonly acknowledged that evaluation is essential to programme development, this lesson has had little apparent impact on judicial reforms. For the quantity of work that has been done, evaluations are remarkably few, and all too often neither widely consulted nor even available. Everyone reads the evaluation of their own project; almost no one reads those of anyone else's work. This suggests an amazing lack of interest in acquiring information and an incentive system which allows and possibly encourages it, but it is also evident that by intent or mere oversight, evaluations are not easily accessible, even to members of agencies which conducted them. A recent suggestion that major donors share their evaluations is a good sign, but it will be hard to implement if only because they may not know where they have stored them (Hammergren 2002).

Hammergren carried out a series of informal interviews with individuals charged with evaluating programmes for the United Nations Development Programme (UNDP), the United States Agency for International Development (USAID), the IDB and the World Bank; all the interviewees made clear that they did not have access to all the documentation that should have been available. As all the work was commissioned by the respective agencies, Hammergren suspects that this reflects an information storage and retrieval problem, not a conscious effort to keep evaluators in the dark. However, it also demonstrates an inadequate internal usage of the documents: if they were being read and used, then they would have been easier to locate.

The arguments above suggest that very little knowledge has been accumulated about the definition, effects and limits of ROL programmes. Nevertheless, some important lessons have been learned. Today we know that any serious analysis must be country-specific and that talk about ROL lessons learned must reflect upon the social, political, geographic and cultural context in which a programme is implemented. We concede that it may be possible to talk about regional trends or patterns, but

emphasize a local focus. This focus requires cooperation with local groups that have worked in ROL in the country for a long time. These local perspectives to ROL programmes are invaluable.

Usually, big ROL programmes do not take into account the expert opinions of local NGOs working in a country for a long time. This omission leads to design and implementation problems, difficulty in evaluating results, and inability to process lessons learned. Donor agencies should consult local experts or organizations that are deeply rooted in the local communities and have a clear understanding of the ROL situation in the country and close ties to key stakeholders.

The main obstacles to an effective evaluation in the rule-of-law field

The main obstacles to effective evaluation in the ROL field are summarized below. The list is not exhaustive but is intended as a first step in the debate. The findings are taken from FORES' own experience dealing with ROL programmes and their evaluation. Our experience tells us that most of these obstacles are more common than we think.

Conceptual obstacles include:

- *Lack of uniformity in definitions.* The absence of clarity in the evaluation terminology leads to confusion about the objectives of evaluation. The concepts 'outcomes', 'outputs', 'results', 'objectives', 'effective' and 'efficient', among others, are used in inconsistent ways. Even though some good glossaries are available, where all the terms are defined, the definitions are not always adopted by evaluators and donors. It is usual to find significant misuses of the terminology even in specialized academic articles.
- *Lack of uniformity in ROL indicators.* One of the biggest obstacles to effective ROL programme evaluation is the lack of homogeneous indicators that allow comparison of data within a country and between countries. A simple indicator, such as the number of judges per inhabitant, could be understood in different ways, and therefore calculated according to different criteria.

    This confusion is a product of the variations among legal and judicial systems. Common law or civil law frameworks and oral or written processes make such a difference that it is not always appropriate to apply the same indicators. For example, the number of cases pending in a US court means the number of hearings pending that many judges of the court may attend. But in a written system like that of Argentina, cases pending mean files that only the assigned judge can resolve.

    As a result, the information obtained about these indicators is not reliable and cannot be used for meaningful comparative evaluation. If we take into account the high cost of producing statistical data and the impracticability of doing so in

each and every evaluation, the decision to standardize and produce reliable ROL statistics should help the evaluation process.

There have been some high-profile efforts to improve the system in recent years. With encouragement from ROL organizations such as FORES, the Justice Studies Center for the Americas (JSCA) developed a manual including a comprehensive list of indicators, with the objective of collecting, disseminating and standardizing judicial statistics and indicators. We would see considerable progress if countries, NGOs and other organizations adopted this standardized measuring system.

• *Difficulty in identifying the causes of effects in the ROL field.* The ROL field, as in other democracy areas, is dynamic in nature, and has an indefinite number of intervening variables.

   Although some experts have conducted important studies demonstrating causal relationships between particular variables, such as the allocation of resources to infrastructure and information technology, on the one hand, and improved clearance rates and reduced case duration on the other (Buscaglia and Dakolias 1999), there is a lack of consensus among ROL experts and practitioners about the causal relationships connecting with other reforms and their results.

• *Judges' rejection of evaluation.* Most judges perceive evaluation as a threat to their power and as a control mechanism imposed from outside the judiciary. In a country like Argentina, where judges consider themselves beyond the reach of monitoring mechanisms, evaluation turns out to be impracticable in some judicial jurisdictions.

Operational obstacles include:

• *Improper selection of indicators.* This is one of the most common problems when evaluating ROL programmes and it is due mainly to some of the conceptual obstacles mentioned above. Indicators must correspond to the legal and judicial system they aim to evaluate, and with the goals of the project, and must take cultural differences into account.

   For example, in court reform programmes it is common to set time reduction objectives for court decisions. Sometimes programme designers set a fixed rate of reduction, for example, 25 per cent: at the end of the project, court decisions will be produced in a quarter less time than before. This objective presents an array of problems: not all the time involved in producing a judicial decision is attributable to inefficient court management; and there are what ROL experts call dead times, and delays that cannot be controlled by the court or where judges do not want to intervene. Even if we try to separate court delays from other delays, the time reduction may be negligible.

   Similar issues arise in other democracy areas. For example, it is common

to find studies that try to measure a parliament's productivity by counting the number of decisions reached by each parliamentary committee and by the parliament in general. However, the importance of these decisions cannot be determined without classification. In the Argentine Parliament, for example, the majority of decisions are declarations of interest related to relatively trivial things and only a minority concern a small number of important decisions. The total quantity of decisions is therefore not a meaningful indicator.

- *Lack of reliable data.* The evaluation of programme outcomes implies a comparison between the situation ex ante and the situation ex post facto. When there are no hard data at the beginning of a programme, and no hard data are produced as part of the programme, evaluation ex post is impracticable. It is advisable, therefore, to determine the reliability of indicators to be evaluated at the design stage.
- *Lack of beneficiary involvement.* It is common practice for evaluation objectives to be set by the implementing agency without all key actors involved in the project being invited to contribute. The objectives and indicators are selected according to the criteria of the implementer, which may differ substantially from those of the programme user or the beneficiaries' needs. Without a commitment to the original programme objectives, beneficiaries frequently fail to perceive the utility of evaluation, or perceive it as yet another encroachment on their turf. This presents a problem for evaluators because frustration among key actors usually results in a lack of cooperation with the evaluation process. For example, time reduction indicators in court management programmes that are set at the discretion of the implementing agencies, without the judges being consulted, may lead to non-cooperation by the judges. Judges may believe that the indicators are not realistic and are imposed from outside, and simply ignore the evaluation, thus thwarting the entire effort.
- *Lack of real donor interest.* Sometimes donors or implementing agencies perceive the evaluation process as something imposed by their internal policies or by-laws. This too may serve as an obstacle to programme evaluation by rendering the evaluation a mere formality, not a learning experience. In such cases, the results will not be used to improve other programmes, but will be forgotten. Low donor interest also implies that the budget and time allocated for the evaluation will not be adequate, and therefore poor results.
- *The evaluator takes responsibility for poor results, or 'everybody hates the evaluator!'.* One of the toughest moments in the life of a programme is the evaluation stage. Everyone feels challenged, and it is very common to find different parties blaming each other for project failures. Also, key stakeholders may feel that the evaluation process is the moment to express their criticisms, and proceed to do so. These criticisms, if they are not made through the correct channels, can be dangerous if they lead to a pessimistic view of the entire programme. Another

common situation is that bad results in the evaluation give rise to challenges to the evaluation methodology, and even to the evaluators' competence and choice of technique being questioned. The final thing that may occur is that, when confronted with bad results, the client, usually the donor agency, asks the evaluator to change some 'terminology' to make the evaluation less critical, or not to disseminate it. In both cases, an evaluation will not accomplish its goals: the production of lessons learned for future programmes is left incomplete.

- *Impossibility of evaluation due to wrong design.* Some programmes are designed in such a vague way that it is impossible to determine the main objectives, secondary objectives, indicators and courses of action to accomplish stated goals. Another instance of flawed design is the implementation of programmes that were originally designed for other countries, including countries with different judicial structures. An example of this is PROJUM (the Programa de Juzgado Modelo, or Pilot Court Reform Programme, the World Bank's judicial reform programme), initially designed for implementation in Costa Rica and implemented in Argentina, without reflection upon the differences between the two countries' judicial organizations.

- *Lack of implementer knowledge of evaluation.* Some implementers—local government agencies, local NGOs and others—do not have expertise in evaluation or its requirements. They are unable to keep records and collect data to facilitate evaluation. Here the responsibility is shared by designers, implementers, donors and the local government. All parties should decide and include the objectives and activities of an evaluation at the design stage of the programme, and then be sure that the implementers understand the evaluation's purpose and techniques.

- *Uncontrolled exogenous variables.* Attributing effects to a democracy intervention in the ROL sphere is no less difficult than it is in other areas of democracy support. There can be so many intervening variables. Moreover, ROL programmes, just like other democracy programmes, are of a very political nature, and unexpected situations may therefore dramatically affect project results. For example, PROJUM, implemented between 1999 and 2005—suffered from a variety of unexpected political developments that directly impacted on the programme results. These developments included changes in the implementing government agency, misunderstandings between implementers and key stakeholders, lack of political leadership, and the Argentine economic crisis of 2001, which virtually paralysed the programme for more than a year. While some of these exogenous variables could have been foreseen by good expert analysis, others could not.

FORES' evaluation toolkit

Effective ROL evaluation requires a strategy designed to overcome the above barriers to effective evaluation. Here, FORES proposes just such a multifaceted methodology,

capable of assessing the complex interconnections and causes of project results. The components of the FORES 'evaluation toolkit' are:

- institutional evaluation;
- participatory collection, analysis and comparison of hard data;
- collection and analysis of key actors' opinions;
- evaluation of external influences; and
- impact evaluation through analysis of public opinion.

The toolkit includes various components, all of them necessary to achieve an effective overall evaluation. These components must be executed in such a way as to address the problems pointed out above. FORES' toolkit deals with conceptual obstacles by engaging in a professional, multidisciplinary analysis that ensures knowledge of the state of the art and contact with the local legal and judicial system. This is central to avoiding mistakes in evaluation design or indicator selection.

Operational obstacles—such as judicial resistance to evaluation—are addressed in the first stages of the programme design through a participatory methodology. Sufficient involvement of the beneficiaries should ensure more reliable data and the appropriate attribution of effects. The uncontrolled exogenous variables are analysed through an innovative institutional and environmental approach.

Although the ideal scenario includes all the proposed tools in the toolkit, time, money and human resource constraints may require the use of just a selection of them. Our experience suggests that in big ROL programmes the inclusion of all the tools is the only way to understand programme outcomes and impact, and to attribute effects to programme efforts. Institutional evaluation, data analysis and actor opinions are the tools most frequently used in the evaluation process; however, the addition of an external influences evaluation and a public opinion analysis, which are not usually part of the process, make it easier to attribute effects to causes. The analysis of exogenous variables that we propose in the external influence evaluation helps to explain successes and failures that are not attributable to the programme. The public opinion evaluation is the best tool for understanding the medium- and long-term impact of the programme. For all these reasons we ask donors and implementing agencies to include all the components of our toolkit in the evaluation of democracy programmes.

The institutional evaluation

A thorough analysis of the target institution or organization should be performed by an expert team of multidisciplinary professionals. It is important to have a clear picture of the organization and how it works in order to understand its functioning and the complex political processes taking place within it.

With this tool, FORES has learned to identify the causes of particular outcomes

and of the exogenous variables that affect ROL programmes. The tool also helps strengthen the coalition that supports the programme, because major institutional actors—who do not usually get involved with a programme—are incorporated as a relevant source of information. This gives them the opportunity to feel that they are part of the project, and they thereby gain a sense of ownership of the ROL programme.

The tools for performing the institutional evaluation are varied: (a) analysis of documents, by-laws, charts, and all relevant institutional information existing in paper and electronic form; (b) *in situ* review of how the organization works; (c) conversations with organization officials and employees to understand formal and informal roles; and (d) brief analysis of other organizations in close relationship with the target organization.

The institutional evaluation of any organization should be performed at two different points in time—at the beginning of the programme and when all activities have been completed. This allows comparison of the situation before and after programme implementation.

It is advisable to create a data collection instrument that is as objective as possible. As many years may pass from the design of the programme to the final evaluation, an objective instrument, such as a chart, can give some uniformity, similarity and comparability to the data.

The following topics should be studied and analysed:

- the regulatory framework—laws, decrees, by-laws and other internal documents—to determine the kind of agency, mission and objectives, legal attributions and roles, among others;
- main actors or key stakeholders—authorities, officials and employees with their formal duties and informal roles, formal and informal power relationships, and cooperation between or conflict among the actors;
- the internal decision-making process—who is formally in charge of the main decisions and who actually makes them, and whether the formal rules for decision making are followed. In the case of informal decision-making processes, a detailed analysis should be done;
- the strategic plan (if it exists) and main projects; and
- other agencies close to the target agency. Particularly in the case of ROL programmes, it is necessary to understand how the internal processes of the target judicial organization relate to those of other interacting organizations. A common mistake is to attribute delays or inefficiency just to the organization itself when in reality the problem may lie outside as well. For example, courts have often been blamed for delays in the initial stages of a judicial process. However, a deeper analysis shows that these delays are usually due to problems related to the way in which the process is serviced, which depends not on the court but on an external office.

Participatory collection, analysis and comparison of hard data

The collection, analysis and comparison of data are the most traditional components of the ongoing ROL evaluation processes. There is almost no debate about the need for ex ante and ex post assessments to compare the results. There is also consensus about the importance of processing hard data with indicators and indexes.

The problems with evaluation indicators in the ROL field have been described above. They arise from the lack of reliable data or inappropriate selection of indicators at the beginning of the project.

In this sense, FORES considers the evaluation as an ongoing process that should be carefully designed in parallel to the programme. It demands not only careful and up-to-date knowledge but also a deep familiarity with the legal system under analysis. As has been said above, some indicators may work in one country but not in another or, worse, they can be applied in different places but mean different things.

This professional approach must be complemented by the genuine participation of the beneficiaries in the selection, definition and calculation of indicators. Training plays a key role here, because the beneficiaries must learn how to select indicators and define them according to their 'close to the field' knowledge. This training is a real challenge for evaluators because lawyers, judges and judicial employees usually do not have the necessary skills; and it always means an extra effort for the consulting team.

There is no consensus among experts about the indicators that should be used in this field; and there is no tradition of measuring ROL with indicators. FORES' experience in this matter suggests guidelines that assist in the successful collection, analysis and comparison of hard data for evaluating ROL programmes, particularly in the case of the court reform programme in Argentina's Rio Negro Province. They can be summarized as follows.

- Observe strict and high technical standards for the selection, definition and calculation of indicators.
- Get beneficiaries involved in the definition of indicators and in deciding the kind of information to be collected.
- Select information that has a reasonable collection cost; it must be collected through a systematic process, and gathered by the beneficiaries themselves.
- Make the causal relationship between the proposed intervention and expected results explicit during the programme design phase.
- Analyse exogenous variables and uncontrolled alternatives that may cause the indicators to vary, using specific tools such as institutional and external influence evaluation.
- Always perform ex ante and ex post measurements of all indicators; avoid using indicators that cannot be measured after the end of the programme.

Collection and analysis of key actors' opinions

In recent years, most evaluations have included an assessment of key actors' perceptions of the programme results. This has come about as a result of a paradigm change in the public policy perspective—one that is more client- or user-oriented. In this sense, an important indicator of failure or success is the opinion of the actors who deal with the organization or institution.

The methodology for such assessment is usually composed of opinion surveys, in-depth interviews and focus groups, all of which are amply discussed in the wider literature.

The collection and analysis of key actor opinions strengthens their commitment to the programme, and helps identify the causes of effects that may not be evident otherwise. Some high-level actors or groups can also provide hypotheses for the analysis of exogenous and uncontrolled variables affecting the programme.

Surveys and focus groups should be conducted at the beginning and at the end of a programme. Ex ante data are needed for comparison, and it is advisable to use the same collecting questionnaire or tool on each occasion.

When possible, use of a control group ensures more accurate results. A control group is a population with characteristics similar to the one under analysis. For example, if we evaluate the results of a judicial reform project implemented in a court, it would be advisable to choose another similar court and perform the same survey, ex ante and ex post. This is, however, not always possible in ROL programmes as some institutions are unique.

The main topics to take into account for this component of the evaluation include a complete description of the affected population—age, gender, social class, profession and so on. Depending on the situation, it may be advisable to determine quotas; but when the population is homogeneous, this is not so important. The evaluation should also uncover the needs, values and expectations of the main actors, as well as their opinions, based both on their perception and on concrete experiences of dealing with the organization.

Evaluation of external influences

The results of ROL programmes tend to be influenced by situations that are not related to the programmes. External issues should be thoroughly analysed to understand how they interact with a specific project. This is the most difficult component of the evaluation: external influences may be numerous and hard to foresee, particularly in democracy programmes of a political nature. Also, the lack of theoretical consensus among experts, academics and practitioners about the causal relationship between variables and their wider political effects complicates the analysis.

Nonetheless, the external influences analysis is a tool that allows us to understand how external facts affect programme results. This may not be necessary in stable

contexts, but it is particularly important in the uncertain and complex political environments in which most ROL programmes are implemented.

During the evaluation of PROJUM, the use of an external influence evaluation allowed for the external obstacles that caused unexpected delays to be isolated, and this in turn made it possible to produce more useful recommendations. Without this tool, some of the conclusions would have been incorrect.

An external influence evaluation should take into account the following guidelines.

- Identify main external situations (at a local, national and international levels) that may have affected the design and implementation of the programme; it is important to collect leading actors' opinions to ensure that nothing relevant is missing.
- Analyse changes in the main personnel that have taken place in the organizations involved in the programme.
- Review media releases related to the project or related topics.
- Expose and explain all possible intervening external factors that may have affected the programme, and weight them.
- Validate your conclusions as to the impact of external factors through discussion with local experts and leading actors.
- For donors and implementing agencies, always request this component of the evaluation in the evaluation terms and conditions.

The external influence evaluation should not be long and expensive. With the right methodology, and conducted with a knowledgeable and interdisciplinary group of experts, an effective analysis should be relatively easy to perform. It is preferable that the analysis be performed by a mix of local and international experts with knowledge of both the topic and the country, the latter contributing insights on context and culture and the former bringing a necessary objectivity and distance.

Impact evaluation through analysis of public opinion

The final goal of every project aimed at strengthening democracy is to improve the performance of an institution or organization. Therefore, public opinion of the organization ought to improve as programme objectives are achieved. For example, a justice reform programme should result in a better public image of the judiciary, and a programme for strengthening a parliament should increase public trust in the parliament. Although we acknowledge that public opinion reflects several components, and external influences play a relevant role, it important to monitor changes in public opinion and look for correlations between such opinion and programme results.

Monitoring public opinion is a complex process, and may be expensive. In some instances the cost may even be prohibitive. It is therefore advisable to conduct periodic

public opinion polls that monitor the legitimacy of democratic institutions. To this end, FORES, with the Libertad Foundation and the Torcuato Di Tella University School of Law (Buenos Aires), has developed the Justice Reliability Index (JRI), available since mid-2004, which focuses specifically on public perceptions of ROL issues. (For more information about the JRI, see the FORES website at <http://www.foresjusticia.org.ar>.)

This kind of index, along with more specific programme-related public opinion assessments, is an adequate tool for measuring the overall impact of ROL reform efforts. It is advisable to develop and measure this index independently of existing ROL programmes. We strongly encourage the donor community to support implementation of the JRI in other countries, and to use it consistently over time and realize its utility in the ROL project evaluation process.

Another useful strategy for analysing public opinion is to follow media discussions related to the programme, or which may affect it. For this purpose, it is important to analyse media cuttings and to have a media expert on the staff.

Evaluation case studies: FORES' experience in the evaluation field

FORES has used the above five-step methodology to evaluate the success of various ROL programmes. Although it is clearly preferable to use all the steps in combination, FORES' experience of implementation has demonstrated not only that resource and other on-the-ground constraints may limit evaluators' range of action but also that the methodology is sufficiently flexible to accommodate such constraints. Below, we describe three case studies related to the evaluation of recent ROL programmes in Argentina: first, the evaluation of PROJUM; second, the evaluation of the court reform programme in Rio Negro Province; and, third, the Justice Reliability Index.

The evaluation of PROJUM

Through competitive bidding, FORES and the National Center for State Courts (NCSC) were awarded the evaluation of the World Bank justice-sector reform programme—PROJUM—in 2005.[ii] FORES and the NCSC provided a multidisciplinary evaluation team composed of four lawyers with different backgrounds (judicial management, training, indicators, and judicial reform), two sociologists, two experts in quality norms, one political scientist and an information technology expert.

PROJUM was a pilot programme conducted in 12 Argentine federal courts. It implemented new court management methods and tools in order to improve the services delivered to court users. The programme's primary objective was to identify, establish and evaluate the existence of conditions that support judicial reform, and eventually to form part of an overall legal reform programme at the national level.

The first reform measure involved analysis of the existing court organization

and management mechanisms and the development of new administration policies, strategies, and a court management plan detailing operational standards and statistics for monitoring the progress of reform. The second element of the reform programme aimed to develop a permanent solution for reducing the number of pending cases within the selected courts and to improve the skill levels of court officials and personnel, through training in court administration and case management. The third component of the reforms comprised outreach activities and activities to evaluate the results of the model courts by creating judicial information centres, conducting user opinion surveys, disseminating information to the public, and evaluating the project.

The evaluation of PROJUM was primarily focused on the 'outcomes' of the programme. Its main goal was to review the level of implementation of each one of the components. Project design, implementation strategy and financial management of the project, as well as the new software installed in the pilot courts, were beyond the scope of the evaluation. All evaluation topics were determined by the World Bank and PROJUM teams, leaving FORES–NCSC unable to express their opinions on the terms and conditions of the evaluation.

FORES began with an analysis of court performance indicators, the core elements of the project evaluation. These indicators had been defined early in the project and suffered from two major difficulties. First, the selection of indicators was poor. They were defined and calculated in a way that made it difficult to attribute indicator changes over time to the reform programme, because they were susceptible to other variables that distorted the results.[3] Also, the judges who had participated in the identification of indicators did not understand how they worked, and were not ready to support them.

Second, there was a remarkable lack of reliable data and information that should have been produced and collected at the outset of the programme. Due to the delays in implementation, the software was not ready to store data during the programme's early phases, and proper information and statistics were not entered in time to facilitate comparison before and after programme implementation. This fact obliged the FORES team to make estimates, which ended up distorting the indicator results and reduced the reliability of those results. These difficulties demonstrate the need to observe strict technical standards when selecting and calculating indicators, and the importance of beneficiary involvement in the definition of indicators. Another lesson is that causal relationships between the proposed intervention and the expected results should be made explicit during the design phase.

Although these problems hampered FORES' evaluation efforts and probably distorted the results to some degree, other FORES information collection mechanisms helped to fill gaps. The analysis of documents, interviews, focus groups and *in situ* observation allowed the FORES team to analyse the new organizational chart and layout of the 12 target courts, the new quality control system, the availability of the information in the new software and of how up to date it was, the functioning of the

'administrative units' created to serve the 12 courts, as well as the training provided to judicial officials and employees.

The collection and analysis of key actor opinions took place mainly by way of interviews of the judges. The interviews were important in explaining the great delays in project implementation during 2001–2, and in identifying exogenous variables that had a negative impact on the programme, including power struggles between the Judicial Council and the Supreme Court.

FORES also assessed changes in court performance by surveying judicial officials, employees and court users. The analysis of survey results consisted of a comparison between an ex ante sample taken at the beginning of the programme, and the surveys performed during the evaluation process.

These surveys should be considered as only a part of the evaluation of key actors' opinion. Two main obstacles appeared during the process; the first was a methodological mistake in the design of the collection instrument, and the second was the difficulty of attributing a causal relationship between programme results and changes in opinion in the case of the court users. The latter was due (a) to the great number of external variables that were completely out of the control of the programme and (b) to the practical inability of court users to distinguish improvements related to the PROJUM reforms from others that were external to the project.

The implementation of PROJUM (1999–2005) was deeply affected by some important exogenous variables. Some of them were unexpected, such as the worst economic crisis in the history of Argentina (2001), while others were expected but uncontrollable, such as the creation of the Argentine Judicial Council (1998).

These exogenous variables had a negative impact on the programme, and yet the terms and conditions of the evaluation provided by PROJUM/the World Bank did not take them into account and did not ask for an institutional or an external influence evaluation of the sort proposed by the methodology recommended in this chapter. Nevertheless, the FORES' team used some of the techniques for analysing the institutional and external influences to better understand certain obstacles to implementation, particularly the enormous delays that made it almost impossible to evaluate results. This 'unsolicited' piece of the evaluation that FORES conducted was key to the production of informed conclusions and recommendations that were required as part of the evaluation outputs.

In summary, the evaluation of PROJUM had to overcome several unexpected challenges before it could reach conclusions and generate lessons useful for understanding the programme and facilitating further developments. The recommendations for future action and reform were a core part of the evaluation; if FORES had not applied its toolkit methodology, the outcome of the evaluation would have been very different.

The success of the toolkit methodology was validated by the consensus that FORES' recommendations elicited in a seminar at which they were shared and debated with major judicial reform actors in Argentina. This was possible only because

FORES had applied its expertise in programme analysis, particularly in the analysis of justice-sector indicators and in the introduction of institutional and environmental evaluation.

While the PROJUM evaluation experience demonstrates the toolkit's flexibility and the complementary nature of the five tools, the evaluation of the court reform programme in Rio Negro demonstrates the toolkit's financial efficacy. Although at first sight the toolkit's five components seem to require substantial investment, the participatory nature of the methodology helps offset the demand on resources and makes the methodology more accessible.

The evaluation of the court reform programme in Rio Negro Province

FORES with the support of the Management Development Institute of Argentina (Instituto para el Desarrollo Empresarial Argentino) implemented an innovative participatory reform methodology in a pilot project that included three courts in Bariloche city in 2004. The Superior Tribunal of Justice of Rio Negro Province sponsored this pilot project. Due to its overwhelming success, the project was replicated in every court in Bariloche in 2005; a third stage was expected to take place later.

FORES developed an innovative court management reform methodology based on the training of judges, court officers and judicial employees, and on technical assistance. The main components of the training programme were judicial process analysis, identification of best practices (understood as those that increase user satisfaction with court performance), 'benchmarking', change management and project management.

'Benchmarking' consists of an assessment of performance in comparison to the best performers in a particular area. The first step consists of defining the areas of practice (as there is no ideal organization, it is very likely that the target organization is the best in a particular practice but not necessarily in other practices). The identified best practice then acts as a standard that all other courts participating in the programme should emulate. Within the judiciary, the criterion for defining a best practice is client satisfaction. This satisfaction does not refer to winning a case but to having received adequate justice 'service'. Adequacy includes the notions of effectiveness (a fair an impartial solution) and efficiency (decision made in due time). The interest in a more effective and efficient court process involves not only the actual users of the courts but all the citizens, who will benefit from a better judicial branch.

The judges and judicial personnel in the Rio Negro courts selected the best judicial practices that they hoped to emulate. They also identified obstacles to effective judicial administration and to achieving the overall objectives of the reform programme selected. The intensive judicial reform training provided by FORES experts facilitated benchmarking as well as the identification of obstacles and objectives. The main judicial actors in Rio Negro not only participated in project design; they also selected and calculated the evaluation indicators. That is, judges and judicial personnel

measured, monitored and evaluated their own progress towards self-set objectives, with the assistance and training provided by professional FORES evaluators.

Although the lack of predetermined indicators prevented programme organizers and donors from foreseeing results in advance, it presented an opportunity to enhance programme sustainability and build a beneficiary coalition in support of the programme. This participatory methodology provides beneficiaries with a sense of ownership over the reform programme and ensures the main actors' commitment to the results, thereby generating deep, sustainable reforms.

FORES worked with key actors to evaluate each programme after six and 12 months of operation. However, the FORES evaluators did not come in to pass judgement on the programmes, but instead to collaborate with court officials and employees in identifying programme strengths and weaknesses and appropriate future work. A deep understanding of internal processes was necessary for the evaluation, and this could only be achieved with internal commitment and cooperation. At the end of the programme in Rio Negro, the judges themselves presented the results of the programme at a public event in the presence of the media.

Although the participatory strategy increased the original coalition of support for the project and facilitated project evaluation, it was not entirely without faults. After the collection of the information, the evaluation team identified mistaken data (including wrong numbers, incorrect interpretation and incorrect attribution of effects to causes). These mistakes were due to the fact that the people in charge of gathering and processing the data were judicial employees, not experts in the methodology. After the data were cleaned, the evaluation team interpreted them and proposed conclusions. During this stage, a multidisciplinary team of professionals helped the evaluation team identify possible external variables intervening and modifying the data. FORES professionals are therefore indispensable to the project design and evaluation process in that not only do they train judicial actors in reform techniques and project design; they are also available to 'fix' the mistakes that are unavoidable in programmes implemented by novice reformers.

Two external events helped the evaluation of the Rio Negro court reform programme. First, the Superior Tribunal of the province provided a financial incentive for those courts that met the performance standards, which promoted a culture of reform and demonstration of success through evaluation. Second, the evaluation results were disseminated at a public meeting in Bariloche city, with the participation of the local and national media.

Despite the small budget, the programme was a success and provided a new model through which reformers and donors can make a big impact and without spending huge amounts of money. (The budget for the entire reform programme, including its evaluation, was approximately 50,000 US dollars (USD) over two years.) Although the personnel training was taxing, and oversight by FORES professionals was essential, assisting project beneficiaries in the implementation of the toolkit methodology is an effective way of ensuring effective evaluation without extensive investment.

The Justice Reliability Index

Another technique for reducing the overall cost of the FORES evaluation methodology toolkit is to monitor public opinion consistently over time. Although such monitoring may be complemented by project-specific public-opinion evaluations, the general legitimacy level of a public institution is often sufficient to enable the application of the fifth tool in the toolkit, namely impact evaluation through public opinion.

As part of a continuous evaluation of public opinion on justice-related issues, FORES, the Freedom Foundation and the Torcuato Di Tella University School of Law, developed and periodically administer the JRI. The JRI is not an ROL programme itself, but a specific tool developed for periodically measuring public opinion of judicial administration and law enforcement in Argentina. It is designed to gather information about people's behaviour when facing concrete legal conflicts as well as citizens' opinions about the Argentine justice system in general.

This index works under the assumption that the reliability of an institution is reflected not only by what individuals say but also by what they do or are willing to do in connection with it. The JRI is therefore designed as a combination of two sub-indexes. The first relates to individuals' behaviour, in other words, what people do or would do when dealing with concrete legal conflicts in patrimonial, family or labour matters (the behavioural sub-index). The second sub-index measures the individuals' belief in the justice system's impartiality, efficiency and honesty (the perceptual sub-index).

The JRI has the following three characteristics. It has specificity: it is exclusively focused on the reliability of the justice system. It is two-dimensional in that it evaluates behavioural and perceptual elements. And it is systematic: it consists of three polls per year.

Since it was first measured in 2004, the behavioural sub-index shows higher scores than the perceptual sub-index (approximately double, in fact). This fact suggests that what individuals are willing to do in concrete situations in which they have the option to access judicial intervention does not correlate with the image they have of Argentine judicial system in terms of its impartiality, efficiency and honesty.

The JRI is an important tool that allows public perceptions of progress in ROL to be measured. It also contributes to the external influence evaluation of any programme in the ROL field in Argentina, by analysing the mood and opinions of the citizens with respect to the justice system. The JRI provides detailed and reliable data about public opinion related to justice issues and allows analysis of its evolution over time.

For example, in 2003, the Argentine Supreme Court began a process of renewal. Four of its nine members resigned, and two others were removed through impeachment. Argentina's president implemented a public consultation process evaluating each new judicial nominee prior to his or her appointment. This process enjoyed a high level of civil society and media participation. The JRI was the tool selected by one of the most important Argentine newspapers to monitor the impact 40 per cent of these changes

in the Supreme Court had on the public perception of the justice system (*La Nación* 2005 and 2006).

Clearly, the JRI does not attribute effects to particular ROL programmes. Nevertheless, it is an important tool that serves to measure the impact of the overall conglomeration of reform efforts. As a side effect, it helps show how difficult it is to attribute causal relationships in social behaviour: at the same time as people express distrust in the justice system, they are willing to go to court to resolve a conflict.

Conclusions and recommendations

The evaluation of ROL programmes is still in an early stage of development. There is not much consensus among experts, donors and practitioners about what and how to evaluate. There is also a lack of consensus about how particular interventions cause specific results, apart from a few exceptions. Multiple and complex obstacles to effective evaluation of ROL programmes hamper the field, although examples of successful evaluations do exist.

The ability to evaluate of the overall impacts of ROL programmes on all citizens, and not only on the specific beneficiaries of a particular programme, remains uncertain. It is also unclear how ROL programmes impact upon or affect the strengthening of democracy.

FORES' approach to the evaluation of ROL programmes consists of the implementation of five levels of evaluation: (a) institutional evaluation; (b) the analysis of hard data; (c) key actors' opinions; (d) external influence evaluation; and (e) impact evaluation through public opinion. This multi-step approach helps us to better understand the relationships between intervention and results, and to assess programme impact.

FORES' approach may increase the cost of the evaluation, by using multidisciplinary teams and training actors in evaluation skills. But the expense may be restrained by using local or regional evaluation experts and by incorporating programme beneficiaries into the evaluation process.

The participation of main actors is essential to the ROL programme success and evaluation. It ensures that assessment indicators and methodology will not be challenged at the evaluation stage, and builds commitment to the implementation of the evaluation. The use of local experts and the programme beneficiaries as evaluators has additional benefits: it helps build local capacity and ensure the sustainability of the results.

NGOs like FORES have a double role in evaluation—a technical role as implementers and evaluators of ROL programmes; and a social role as active members of civil society committed to long-term democratic development. These organizations are in a unique position to understand and analyse successes or failures in the current evaluation methodology, and to make recommendations on how to improve the current state of the art.

FORES' lessons for donors and the ROL community for improving ROL project evaluation can be summed up in three recommendations. First, request that every programme evaluation follow the FORES toolkit. Second, support the periodic assessment of public opinion on ROL issues through reliable tools, such as the Justice Reliability Index. And, third, support studies on evaluation in ROL programmes and the attribution of effects to causes.

Notes

[1] To illustrate, in 2005–6 the government of Argentina decided to repay all its debt to the International Monetary Fund with the sole purpose of avoiding IMF evaluation of its economic and financial policy.

[2] The actual findings of the project are not discussed here owing to a confidentiality agreement. To illustrate, the deposition into the archives of the files on cases that were judicially paralysed gave an impression of improved court management indicators, but in practice this did not signify any real improvement in case management.

[3] According to Spendolini (1992) benchmarking is a systematic and ongoing process to evaluate products, services and work processes of organizations recognized for having best practices, with the objective of organizational improvement.

Chapter 5

Exploring a human rights-based approach to the evaluation of democracy support

Chapter 5

Hanne Lund Madsen

Exploring a human rights-based approach to the evaluation of democracy support

This chapter highlights experiences and methods used in evaluating democracy support while also debating the new frontiers in evaluation thinking and practices within democratization and human rights. In particular it inquires into how the RBA may be applied in evaluations of democracy support. It suggests that the rights-based approach is useful as it provides a consistent framework for situation analysis, for programme design and for monitoring and evaluation. Moreover, it is applicable at all levels, from the global to the local community level. It thereby builds a bridge between meta and micro frameworks and between situation analysis and change analysis. Finally, the RBA provides a link between the development cooperation community and the human rights (treaty) monitoring bodies.

Introduction: general lessons from evaluations of democracy support

The increase in support to democratization and human rights as part of development assistance programme means that the challenge of evaluating democracy support must be met and explored. The International IDEA/Swedish International Development Cooperation Agency (Sida) workshop on this theme held in Stockholm in April 2006 confirmed the need to enhance innovation and cross-sectoral learning in the evaluation field but also the need to clarify what outcomes and impact democracy assistance is really aiming to achieve.

A perusal of existing evaluations of democracy support makes it clear that many evaluations identify the following issues and challenges:

- the vagueness of the objectives and of the definitions of democracy applied;
- the tendency to overkill and overload, both in the purposes and the scope of the evaluations and in the data compiled;
- the crucial importance of institutional capacity to learn; without a learning organization evaluations themselves will have little impact;
- the non-availability of data and baselines; and
- problems of aggregation and attribution.

These lessons and challenges are certainly not unique to human rights and democracy support evaluations. Many of them are typical stumbling blocks in all evaluations and thus not peculiar to democracy support evaluations. Any evaluation will have to deal with them; they cannot be shirked. Forss, too, clearly rejects the argument that democracy is too complicated to be subjected to evaluation, and refers to the many ways in which democracy and human rights have been studied and evaluated in Sweden (Forss 2002).

Much reflection on the evaluation of democracy support tends to treat the more general and inherent difficulties of evaluation as being specific and unique to democracy support evaluation, and the general difficulties tend to overshadow the question of whether support made a difference. In other words, what should form the real core of democracy support evaluation, namely measuring changes in the essential features of the substantive democracy that is being practised within a given country or context, receives less attention than it should. At the same time those characteristics and features of democracy evaluation that are really unique and very different from, say, the evaluation of health and of health interventions tend to be overlooked.

Finally, the evaluation of democracy support has been confused with the evaluation of democracy as such. The evaluation of democracy support is essentially intervention-oriented and seeks to measure changes brought about by a given intervention in a specific context. Democracy evaluation or assessment is a situation analysis based on a number of analytical dimensions or indicators, which is derived from a theory about what democracy is. Much debate has taken place as to whether it is necessary to have a definition of democracy in order to be able to evaluate changes brought about by democracy support. Many changes brought about by a democracy support intervention can in fact be measured without having an analytical model of democracy; but if the purpose is to measure the intervention's impact on the democratization processes or on the enjoyment of democratic features in a country, some idea about the essential characteristics of democracy is needed. Thus, democracy support evaluation needs to embrace both the dimensions used in democracy assessment and the dimensions used in measuring change brought about by a particular intervention's interplay with the existing situation.

In search of analytical frameworks

What is really striking is that, especially in the first phases of democracy support, little emphasis was given to reflection on analytical frameworks and methodologies in the design of the evaluations. Very few terms of reference actually requested the team of evaluators to develop analytical frameworks and methodologies and to reflect on their applicability and accuracy (Organisation for Economic Co-operation and Development 1997). An exception was the United States Agency for International Development (USAID)'s evaluation of its experience with democracy initiatives in the 1990s, which posed questions about what kinds of performance indicator are valid for measuring the results of democratic institution building, but without arriving at an answer. Interestingly, the study also used the impact on human rights as a parameter, but concluded that significant attitudinal or behavioural changes were not discernible, although this could be a function of the evaluation methods used rather than an actual indication of what really happened (United States Agency for International Development 1990: ix).

The development of such analytical frameworks is especially important when the interventions in themselves do not reflect or are not developed on the basis of a clear theory of change or conceptual framework, as was generally the case: 'This is compounded by lack of theory or conceptual framework for PD/GG assistance...' (Organisation for Economic Co-operation and Development 1997: 28).

The Danish NGO Impact Study published in 1999 (Danish International Development Agency 1999d) was a major undertaking that assessed the impact of support made through Danish non-governmental organizations (NGOs) over the period 1988–98. In 1998 Danish overseas development assistance amounted to 10,072 million Danish kroner (DKK, or c. 1,500 million USD dollars, USD), of which 9 per cent (920 million DKK) was channelled through Danish NGOs. The study included a desk study, three country studies and three in-depth studies that examined clusters of projects. It had two basic objectives:

- to document and assess the relevance and impact, including the main strengths and weaknesses, of development interventions supported by Danish NGOs in selected developing countries; and
- to compile, develop and test suitable methods to assess the long-term relevance and impact of NGO-supported development interventions.

The *Danish NGO Impact Study* assessed the impact of support through Danish NGOs on democratization in local communities in countries in the South, but without clearly defining democracy or identifying where changes would be located if they did take place. The glossary for the study tells us that democratization is 'the involvement of previously excluded groups in national political debate or activities, and the extent to which a development project has broadened the base of community

participation in development activities' (Danish International Development Agency 1999d). The impact study made a link between participation as a mode of project implementation and democratization, by concluding that 'In this overall scenario, the increased involvement of people in development projects—even if this involvement is limited to open discussions and consultation—could be seen as a useful first step for communities who previously had never been asked or consulted. It is a tenuous link but participation does appear to be slowly happening, albeit in a tentative form in some projects. And this increasing participation could be seen as an incipient form of democratization' (Danish International Development Agency 1999d: 47). In line with many other donors and scholars, the impact study believed that support to civil society was an important way of strengthening democratization. And, interestingly, it concluded that the impact on incipient forms of democratization may be larger than the impact on the development of civil society movements.

The picture is generally similar in more recent evaluations. The major Danish International Development Agency (Danida) evaluation *Danish Support to Promotion of Human Rights and Democratization* of 1999 mentioned the many obstacles and difficulties in democracy assistance evaluation, such as disagreements about what democracy is and 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' (Danish International Development Agency 1999a: 11). However, no alternative methodologies were proposed. The issue of how to measure impact on human rights and democratization was also given very little consideration.

The Danida evaluation consisted of four thematic studies and four country studies. In the thematic study on elections there is a relevant discussion about what constitutes a free and fair election and how a free and fair election process contributes to improving democracy. This in turn provides the starting point for designing some categories for measuring impact, combined with a pragmatic approach that relies on stakeholder identification of indicators: 'To assess the impact, the team has therefore chosen a very pragmatic approach, whereby a number of more general indicators of assumed relevance and qualitative assessments from interviewed stakeholders have been used (such as for example a more levelled playing field, improved NGO capacity concerning monitoring activities, improved Electoral Commission capacity etc.)' (Danish International Development Agency 1999b: 4). The evaluation did not, however, specify the notions of democracy that governed the decisions regarding the design of the aid interventions that were subject to evaluation.

The large technical cooperation programme of the United Nations Office of the High Commissioner for Human Rights (OHCHR) underwent a global review in 2003. Among several objectives of this review, the terms of reference clearly stipulated that 'The Review will focus on impact and achievement' and moreover 'assess how the assistance has contributed to the promotion and protection of human rights' (Netherlands Institute for Human Rights 2003). Very little space was given in the

review to discussing how the impact on the promotion and protection of human rights should be considered, or which analytical categories were used and which underlying theories of change were applied both in the programme formulation and within the perspective of the evaluators. In this case the review found the projects and interventions to be very scattered, with goals that were very vague so that there was no clear orientation on impact. This could be seen as a reason for the evaluators not moving to consider the impact and instead concentrating on other factors leading to inadequate project identification or development and management. However, the review did put forward many interesting observations regarding synergies between different strategies, which are considered further below.

Another striking feature of evaluations to date is the reliance on methodologies that reflect the traditional impact chain, starting from an assessment of the activities, the outputs, the outcomes and finally the impact that may have been produced by these activities and interventions. This use of a one-dimensional impact chain not only risks exaggerating the significance of the project activities, but also runs counter to the knowledge we have about the dynamics of change within democracy and human rights, which is very multidimensional, dynamic and unpredictable.

A third striking feature is how little attention has been given to the definition of impact itself. Some studies employ the traditional development impact definition of significant changes in people's lives. Others assume that increased NGO collaboration is an articulation of impact in itself. The proposition that democracy support evaluation can be conducted without starting out from a theory of democratization and some idea of what constitutes democratic change is indeed deeply problematic. Gaventa (2006) provides an illustration of the problem in which he distinguishes between four different approaches, all of which have been applied within the 'deepening democracy' school of thought—a school that exists outside the main schools within representative democracy and substantive democracy thinking. Depending on whether the intervention and the evaluation adopt a deliberative democracy approach or an empowered participatory governance approach, the choice of analytical fields of investigation will be very different.

A series of publications addressing 'the evaluability of democracy support' serve to reflect some of the difficulties touched on above. For example, in 2001 the government of Sweden even directly requested Sida to develop a method for assessing results in respect of the development cooperation objective of promoting 'democratic governance'. However, the study focused on evaluation methods and systems rather than addressing the hard issue of evaluating democratic change (Forss 2002). Elsewhere studies have confirmed the constraints of the logframe approach that has been used in some quarters, which takes too narrow a time perspective. But no alternatives have really gained ground over and above giving more recognition to such things as higher levels of uncertainty, greater risk preparedness, greater process orientation and so on.

In sum, notwithstanding that the field of evaluating democracy and human rights assistance is still fairly new, too little attention has been given to considering appropriate analytical frameworks, clarifying the theory of change, developing new methodologies and deconstructing the notion of 'impact' in new and possibly more relevant ways. To date the learning that has proceeded from the evaluation exercises has been quite limited and a *community of practice* with agreed standards, tools and approaches has as yet not developed.

The role of human rights in democracy support and the evaluation of democracy support

Democracy and human rights interventions should both be subjected to evaluation and impact measurement just like any other type of activity. However, the specificity of human rights and democratization work calls for distinctive methodologies and approaches. The difficulties of applying conventional evaluation approaches to human rights projects are considered elsewhere (Madsen 1998; Forss 2002).

But in what way is democracy and human rights support different from, say, support for basic needs or poverty alleviation? And how does impact assessment in the case of human rights differ from impact assessment for, say, agricultural development? While most international donors package human rights and democratization support together, there are major differences between the two. However, the common ground is that most actors and scholars—no matter what definition of democracy is applied—tend to agree that progress or regression in a number of fundamental human rights will help determine the democratic development of a country. In fact it is not easy to find a definition of democracy that does not contain some notion of human rights. In some instances donor democracy support evaluations have almost completely equated democracy with human rights, as in the evaluation of European Commission Positive Measures in Favour of Human Rights and Democracy: 'When talking about human rights, reference is made to universally accepted human rights standards, as codified by the United Nations... Democracy is the realisation of these rights... The following three groups of rights are considered to be essential for a functioning democracy' (German Development Institute 1995: 15). Most definitions of democracy actually specify certain political and civil rights[1]. And Gaventa (2006) argues that even without making explicit reference to rights, any view of democracy also implies a view of citizenship and the rights and duties associated with that.

Democracy—like human rights—is first and foremost an expression of a relationship and the qualitative characteristics of this relationship. Democracy expresses a relation between government and citizens that is different from a relationship governed by authoritarianism. Democratic development inherently involves some notion of power dynamics and struggles for human rights, as reflected in Gaventa's formulation: 'Democracy-building is an ongoing process of struggle and contestation rather than the adoption of a standard institutional design' (Gaventa 2006: 3).

The Democracy Assessment Tool developed by International IDEA explicitly recognizes human rights as a fundamental pillar of democracy assessment. It employs a number of traditional democracy parameters. At the same time the tool is very unconventional and progressive in that it encompasses economic and social rights as a vector for determining the development of democracy, in line with the following perspective: 'In some formulations, especially Latin American, this [democracy] view is also about extension of rights. Full democratic citizenship is not only obtained through the exercise of political and civil rights, but also through social rights, which in turn may be gained through participatory processes and struggles' (Gaventa 2006: 11). Furthermore the Democracy Assessment Tool seeks to cover the international dimensions of democracy, which in a human rights perspective means extraterritorial obligations. (Human rights treaties not only bind state parties to implement human rights within their borders: states undertake certain obligations towards persons outside their territory as well (Coomans and Kanninga 2004).) The Democracy Assessment Tool is designed to be usable for all, from politicians to students, for IDEA underlines the crucial importance of citizens engaging in the self-assessment of the conditions of their democracy. Each parameter is accompanied by a number of more specific and targeted questions.

International IDEA (2002: 16) outlines the parameters as:
- I.   Citizenship, law and rights
     - 1. Nationhood and citizenship
     - 2. Rule of law and access to justice
     - 3. Civil and political rights
     - 4. Economic and social rights
- II. Representative and accountable government
     - 5. Free and fair elections
     - 6. The democratic role of political parties
     - 7. Government effectiveness and accountability
     - 8. Civilian control of the military and police forces
     - 9. Minimizing corruption
- III. Civil society and popular participation
     - 10. The media in a democratic society
     - 11. Political participation
     - 12. Government responsiveness
     - 13. Decentralization
- IV. Democracy beyond the state
     - 14. International dimensions of democracy.

As we see, human rights are one out of four main pillars of the democracy assessment. Thus, it is fully justified and necessary to look at improvements in human rights as part of an evaluation of democracy support. A final justification is that we would

normally judge the quality of a political system by the way those in power treat the citizen. Also in this perspective the respect for the fundamental human rights of the citizen is a main determinant of the quality of the ruling system and the quality of the changes occurring.

This point of departure prompts the questions: What is an improvement in human rights all about, irrespective of the strategy or means employed to promote rights in any given situation? Can we extrapolate some constituent characteristics of human rights to guide assessments of the impact of interventions and projects? By their very nature the answers will then tell us something about strategy too. In other words, can we find a correspondence between analysis of the human rights situation and analysis of human rights change?

A further reason for looking at human rights in democracy support evaluation is that in recent years considerable attention has been paid to applying a rights-based approach (RBA) to the programming of support, both in mainstream development and within the human rights and democratization sector specifically. Applying it to the latter has been seen as tautologous by those actors who argue that human rights support is per se rights-based and that the rights-based approach has nothing additional to offer in this regard. In spite of that, the United Nations Development Programme (UNDP) and others have shown very convincingly that traditional access to justice programmes and legal reform programmes does undergo change when a rights-based approach is actively used (Golub 2003).

Considering that the respect for and the protection and fulfilment of a number of rights are crucial to democratic development, we will now explore the key element of a rights-based approach and later the significance of a rights-based approach to the evaluation of democracy support.

The rights-based approach

The human rights-based approach as promoted by the United Nations (UN) Common Understanding stipulates that:

1. All programmes of development co-operation, policies and technical assistance should further the realisation of human rights as laid down in the Universal Declaration of Human Rights and other international human rights instruments.

2. Human rights standards contained in, and principles derived from, the Universal Declaration of Human Rights and other international human rights instruments guide all development cooperation and programming in all sectors and in all phases of the programming process.

3. Development cooperation contributes to the development of the capacities of 'duty-bearers' to meet their obligations and/or of 'rights-holders' to claim their rights.

A rights-based approach is a conceptual framework for the process of human development that is normatively based on international human rights standards and operationally directed to promoting and protecting human rights. Essentially, a rights-based approach integrates the norms, standards and principles of the international human rights system into the plans, policies and processes of programme development (United Nations, Office of the High Commissioner for Human Rights 2003: 1).

Human rights, on the one hand, set standards for what individuals and communities—the rights-holders—are entitled to have, to do or to receive. On the other hand, human rights imply obligations and duties on others—the duty-bearers—and ultimately on the state.

Human rights are basically a regulation of the relationship between state and citizens, and human rights unfold at the interface between duty-bearers and rights-holders. Thus, the relational character of human rights determines the quality of the human rights situation, just as the relation between the elected and the electorate is a determinant of how democratic a political system is.

Both in the design of human rights and democracy support and in the evaluation, the focus on the relation—the interface—is crucial. It follows that there are four key dimensions for both programming and evaluation.

- The first dimension highlights the duty-bearers' obligations to respect, to protect and to fulfil human rights.
- A second dimension captures the process, which must be characterized by non-discrimination, accountability and the right to participation.
- The third is the duty-bearers' capability to comply with their obligations, that is, the extent to which the obligations are recognized as a duty and are backed up by adequate resources to act.
- The fourth concerns the rights-holders' capability to access and claim their rights, that is to say entailing the recognition of their rights, the legitimacy to claim, and the resources to access and claim the rights.

In figure 5.1 the RBA Navigator highlights the key dimensions and four corners of the compass. It can be used to make the initial problem or situation analysis on the basis of which strategies can be developed and implemented and subsequently evaluated.

Figure 5.1: The RBA Navigator

International and National Human Rights System

Capability to Comply
* Recognition of duty to act
* Authority/Legitimacy to act
* Resources to act

Guardians of Rights: Human Rights Commissions, etc

Duty-bearers
Government & Non-governmental entities

Process Rights:
* Non-discrimination
* Right to Participation
* Accountability

Obligations:
* Respect
* Protect
* Fulfil

Right-holders
All citizens

Capability to Access and Claim
* Recognition of right
* Authority/Legitimacy to claim
* Resources to access and claim

Human Rights Defenders Human Rights NGOs etc.

Source: Lund Masen, Hanne (2003)

The human rights system

Looking at each element of the RBA Navigator we see that the overall regulatory framework consists of both national law and binding international human rights conventions ratified by the country in question. It is relevant both to analyse the existing national human rights framework and to relate programme interventions to (a) the implementation of existing obligations, and/or (b) the reform of national frameworks with a view to enhancing compliance with international norms.

Actors and capabilities

Rights-holders and duty-bearers constitute the key actors. However, many other actors and institutions also play a role, which is illustrated by the reference to 'human rights guardians' and 'human rights defenders'. The part played by the media too is relevant. It is important to recognize that the roles of duty-bearer and rights-holder are not static or separate. Rights and duties go together. But, depending on the situation, the same citizen may carry the role of duty-bearer or rights-holder. For a policewoman, for example, the role of duty-bearer is to the fore. But in her role as a female employee in the district police, her role as a right-holder will be to the fore. NGOs that consider themselves defenders of human rights also carry certain obligations, as do political parties.

It is generally accepted that the ultimate responsibility for respect, protection and fulfilment rests with the state and its relevant judicial, legislative and executive institutions. However, it is also increasingly being recognized that all actors have a duty to respect and not directly interfere with the enjoyment of other people's rights. For example, private companies have a duty in particular to respect and where possible protect labour rights in the workplace.

The capability of the duty-bearers to comply may be assessed in terms of (a) the recognition of the duty/willingness; (b) the authority and legitimacy to act; and (c) the resources to act and comply, covering human, organizational, technical and financial means.

The capability of the rights-holders can also be assessed in terms of (a) the recognition (of their rights), (b) the authority/legitimacy for claiming their rights, and (c) the resources available for defending their rights in terms of skills, finances and organizational mobilization. The same applies to human rights guardians. The importance of 'recognition' is highlighted by lessons showing that large programmes of technical and financial support to national human rights commissions have tended to have weak results. This is due to lack of investment in developing a willingness to perform as watchdogs and call for accountability in particular cases either by judicial process or other means.

Obligations

The trinity of obligations concerns respect for, protection of and fulfilment of all human rights, be they economic, social and cultural or civil and political—the so-called substantive standards such as the right to health or freedom of association. (For an initial discussion of the trinity of 'respect, protect and fulfil' and the dialectics between duty-bearer and rights-holder (the RBA) in evaluation, see also Madsen 1998.)

Respect requires the state and all its organs and agents to abstain from carrying out, sponsoring or tolerating any practice, policy or legal measure, which violates the integrity of individuals or infringes on their freedom to access resources. It requires that legislative and administrative codes take account of guaranteed rights. It concerns appropriate legislation confirming the rights of all groups.

Protection obliges the state and its agents to prevent the violation of rights by other individuals or non-state actors. It requires that where violations do occur, appropriate remedies exist in the form of accessible and well-publicized complaints and inspection procedures. It concerns for instance the establishment of independent ombuds-institutions, but basically it is a question of the rule of law and access to justice, whether through formal or traditional justice systems. Determining state involvement in violations, asylum law works with the following typology. Violations are

- investigated and prosecuted;
- tolerated or ignored;
- sponsored; or
- directly commissioned.

Fulfilment obliges the state to provide opportunities for (facilitate) or to provide directly for the enjoyment of the right. It requires proactive enhancement of the opportunities of individuals or groups and the direct provision of benefits and services. It involves issues of public expenditure, the regulation of the economy, the provision of basic services and redistributive measures. The duty to fulfil covers those active measures that are necessary for giving opportunities to access the entitlements which citizens have rights to.

Failure to uphold human rights is generally considered in terms of acts of commission or acts of omission—that is, it is due either to direct action that violates rights or to inaction whereby the enjoyment of rights is infringed. Moreover, it is generally understood that the state has obligations not only in terms of how it acts—obligations of conduct—but also for the outcomes such actions have—the obligation of result. The level of enjoyment of rights among citizens corresponds with the obligation of result, which analysts who work in programme design would label as impact. The above categories are useful for both problem analysis and programme design—and for monitoring.

The potential strength of the RBA is not just its focus on key characteristics of human rights. Another strength is that if it is used for both programming and evaluation we will have an exceptional situation of the same analytical framework and methodology being applied both for programming and for evaluation. Very seldom do evaluators use the same paradigm as that used by the programmers—if, that is, the paradigm and theory of change are spelled out at all. Another potential strength, as is mentioned above, is that it provides a uniform framework for both situation assessment and change assessment. Even though there may be some variations in interpretation, there are a number of defined standards supported by case law and judicial precedent in addition to the authoritative interpretations provided by the treaty bodies and the special rapporteurs, on a long list of rights.

Normally, the first advantage of the rights-based approach to be mentioned is the common agreed framework (OHCHR 2005). Unfortunately, the international development cooperation community continues—even after human rights have entered the development agenda—to pay relatively little attention to the work undertaken by the human rights monitoring bodies. The national reporting on poverty reduction strategy papers (PRSPs) and the United Nations Millennium Development Goals often run in parallel with and isolated from the reporting obligations under the international treaties—the latter often receiving much less technical assistance and resources than the former. The RBA holds the potential to bring closer together the actual reporting processes as well as the analytical framework and the data sets used.

Programming and evaluation

The RBA Navigator illustrated in figure 5.1 provides a guide for analysis, programming, and monitoring and evaluation. It informs the analysis at local community level, at local government level and at the national level. It also applies equally at the international level. The RBA calls for programming in three steps—situation/rights analysis; role and responsibility analysis; and capability analysis.

In order for a programme to be effective at improving human rights it must capture—but not necessarily be limited to—all four dimensions based on a concrete assessment of the failures and potentials within each dimension. Programme components should, accordingly, be designed to complement each other within the overall framework. A template for analysis, programming and evaluation is exemplified in table 5.1.

Table 5.1: The RBA Navigator in analysis, programming and evaluation

RBA Navigator Situation analysis	RBA Navigator Programme design	RBA Navigator Evaluation
Human Rights Framework Analysis of the National Human Rights Framework at work in the given rights area Analysis of the human rights issue selected	• Specify which Human Rights the intervention will address. • Specify pertinent recommendations from UN Treaty Monitoring Bodies (to be) addressed. • Specify if any specific objective regarding improvement of the national human rights protection system.	Outcome in terms of improved national human rights protection system and compliance with Treaty Body requirements
Rights-holders Identification	Specify the rights-holders involved with relevant claims	Recognition achieved as claimants
Duty-bearers Identification	Specify which duty-bearers (chain) regarding the specific human rights issue	Recognition achieved as duty-bearers
Human Rights Situation • Respect • Protect • Fulfil	Specify improvements based on decreased violations and progressive enjoyment by rights-holders.	Impact in terms of • Respect • Protect • Fulfil
Human Rights Situation Analysis of • Non-discrimination • Participation • Accountability	• Specify improvements in terms of non-discriminatory practices (gender, HIV/Aids, ethnicity, language, etc.) • Specify which accountability mechanisms will be strengthened. • Specify how the right to participation will be improved.	Impact in terms of • Non-discrimination • Participation • Accountability
Capabilities to claim • Recognition of their human rights and of the nature of the violation • Authority and legitimacy to act (public litigation) • Resources (advocacy/skills/finances) to act and defend	Specify the specific objectives in terms of improvements in of the claim-holders' capabilities to claim	Outcome in terms of improved capability to claim
Capabilities to comply Recognition of duty and responsibility (willingness) to act Authority and legitimacy to act Resources (human, organisational and financial) to act	Specify the specific objectives regarding improvements of the duty-bearer's capabilities	Outcome in terms of improved capability to comply

Source: Madsen, Hanne Lund, 'Characteristics of Human Rights Indicators', Paper presented at the Danida Seminar on Human Rights, Democratization and Decentralisation', November 2003 (unpublished).

Support to a Human Rights Commission, for example, can be both relevant and justified. The results must be measured in terms of the commission's ability to fulfil its mandate as a guardian of rights, that is, to oversee the compliance of the duty-bearers. Similarly, support to human rights organizations will, if successful, strengthen the organizational capacity and in turn hopefully result in improved human rights lobbying and advocacy. Actual human rights improvements, however, will depend on the interface between rights-holders and duty-bearers. Thus, the strategic cocktail is often a multi-pronged approach, targeting guardians of rights, human rights defenders, rights-holders and duty-bearers at various levels around a particular human rights objective or problem.

Justice flow analysis and chain of justice analysis have been found to be very useful in systematically identifying the opportunities and barriers which a claimant faces in the search for justice at various levels, from the household level to the national judicial institutions, and the choices made at various junctures between formal and customary adjudication systems. In comparison, the duty-chain analysis proposed by the United Nations Children's Fund (UNICEF) looks at the duty-holders vis-à-vis the rights of the child at all levels in the chain, from the obligations of the parents and schoolteachers to municipal authorities and finally lawmakers (Jonsson, 2003: 50).

Such analysis helps to develop comprehensive programmes that address all weak links in the justice chain. The implications for the intervention model are clear:

> Traditionally, donors have often pursued a Rule of Law paradigm focussing on the procedural and institutional aspects of the justice system seeking to enhance the performance of the system within its own borders and with little inter-linkage or relationship to the users or the reality of the users. The relational intervention model or the human rights intervention model calls for investment both in the delivery of justice and the access to justice with the empowerment of users/citizens to access and holding the system accountable as a key ingredient. The intervention model would thus encompass both a rule of law perspective and a legal empowerment perspective. Finally, the intervention model would seek to establish or strengthen the mechanisms through which the various stakeholders may influence or negotiate reform of the system or reconstruct the delimitation lines of the justice system to include both formal and non-formal regulation methods (Madsen 2003: 4).

There are many examples of the rights-based approach being used in programming within human rights and democracy support, and recently many publications have begun to consider the first examples and the lessons, including the weak aspects (Gready 2005). There are fewer examples of the RBA being used in the field of evaluation.

Evaluating categories of aid, or the achievement of change

Evaluations of democracy and human rights support have often mirrored the existing intervention channels, levels or themes that have been the organizing principles or categories of support. In other words the strategic weaknesses of the mode of support tend to be replicated in the evaluation design. Evaluations have become compartmentalized into different strategies or thematic areas trying to assess the impact of single interventions areas within human rights education, legal aid, civil society advocacy, elections, and so on. But in regard to the enjoyment of human rights, impact seldom emerges from the application of just one strategy alone. Similarly we know that training by itself will seldom change behaviour. The impact of democracy and human rights support is first and foremost relational and therefore likely to occur as the result of a combination of strategies addressing both the electorate and the electoral institutions, or addressing both the rights-holders and the duty-bearers. Nevertheless, the evaluations considered here generally seek to establish a direct link between the support area and impact.

The rights-based approach calls for analysis of the *vertical* linkages and the *interface* between duty-holders and rights-holders and any improvements in this relationship. In contrast, democracy and human rights support evaluations have mainly been conducted *horizontally*, with either a focus on various levels of support or different thematic areas, which will be considered below.

The Danida evaluation *Danish Support to Promotion of Human Rights and Democratization* (Danish International Development Agency 1999a) is structured according to four thematic studies, each focusing on one intervention area:

- justice, the constitution and legislation;
- elections;
- the media; and
- participation and empowerment.

Each thematic evaluation attempts to establish a link between the intervention area and broader changes and impact—often with only limited success. Obviously, the support of free, professional and responsible media is critical for any improvement in electoral processes. And the disconnect that is evident in the evaluation study between strategies that are so clearly interlinked risks hampering the evaluation results. The Danida evaluation also included four country studies, where the interplay between the various intervention areas could be studied and where the resulting impact could be assessed. However, one main finding from these was that at the national level the various intervention areas often remained separate and with little interlinkage. This was due to weak programming efforts that were not comprehensive. Moreover, the obvious synergies with mainstream development sector programmes were found to be weak or non-existent.

The thematic evaluation of elections maintains an input focus, but broadens the scope of investigation beyond the study of selected specific projects to include all election support activities over a period of ten years:

> This approach provides the option for assessing the election support with the inclusion of both a contextual and a time and process perspective. This is considered rather important since in particular election support is strongly dependent on the establishment of good relations with collaboration partners in the recipient country. Also, the fact that many efforts regarding establishing possibilities for launching the election support—e.g. mutual confidence building, networking etc. are not reflected in the project documents justifies this approach (Danish International Development Agency 1999b: 4).

The horizontal focus of the evaluation was manifest in the choice to classify the various aid efforts into three horizontal levels—regime level, institutional level and citizen's level (Danish International Development Agency 1999a: 8). The horizontal structure is also used in the presentation of the evaluation findings, which implies a tendency to consider support to the citizens in isolation and disconnected from the efforts undertaken at regime level, and vice versa. The crucial interlinkage between the state and citizens, which is captured only weakly in the support design, is replicated in the evaluation design. The lessons learned are also mainly structured on the same three levels and thus miss out on the key question we are most concerned with in democracy assistance—how to improve the *relationship* between those holding power and those delegating it or who do not have any power.

The global review of the OHCHR technical cooperation programme, interestingly, uses the same combination of country studies and thematic studies, defined as:

•   the administration of justice;
•   human rights education;
•   national human rights action plans (NHRAPs); and
•   national human rights institutions.

As with the Danida evaluation, this too used the input structure as the organizing principle. In this case the weak goal and impact orientation of the OHCHR interventions may have necessitated this approach of the review, which has a number of advantages in terms of reviewing management and operational implementation issues. The global review, however, clearly moves beyond the intervention straightjacket and looks into the interlinkages and synergies between the different intervention areas. The conclusion is very clear, however: there were no synergies or interlinkages.

NHRAPs were originally foreseen as the umbrella and the other themes as its components. For instance, an NHRAP should ideally provide and lay the groundwork for the establishment and strengthening of national human rights institutions, among other structures for the promotion and protection of human rights in the country. Human rights education should be an objective stipulated within the plan, and the means by which it is to be delivered provided within its text. Human rights training for judges and police, the administration of justice theme, among many other measures, should be stipulated within the plan as well. This would form a coherent plan.

According to this logic, the NHRAP would be a management and planning tool to coordinate the expertise and potential guidance for national human rights institutions, administration of justice, human rights education and other issues, even the drawing up of reports and follow-up of the recommendations of the treaty bodies and special procedures, into one service package of OHCHR inputs in an NHRAP. At present, the NHRAP theme (or mandate) does not fulfil that role. It is run in parallel with and in fact completely separately to the other thematic mandates of the administration of justice, human rights education and national institutions. …The situation with the other themes appears to be the same. All themes run separately from the other thematic areas and there is little or no management or planning across the themes focused on the country-based programming (Netherlands Institute for Human Rights 2003: 51).

Even in situations where a national plan for improvement has been developed, the interlinkages and synergies remain a challenge. It goes without saying that the prospects for change and impact are weaker in such situations. It is beyond the scope of this chapter to list the many reasons and explanations given for this programming practice within the democracy and human rights assistance field.

The crucial interplay between different human rights strategies is well illustrated by the Human Rights Strategy Web, which is employed as a rule of thumb in the human rights community.

Figure 5.2: The Human Rights Strategy Web
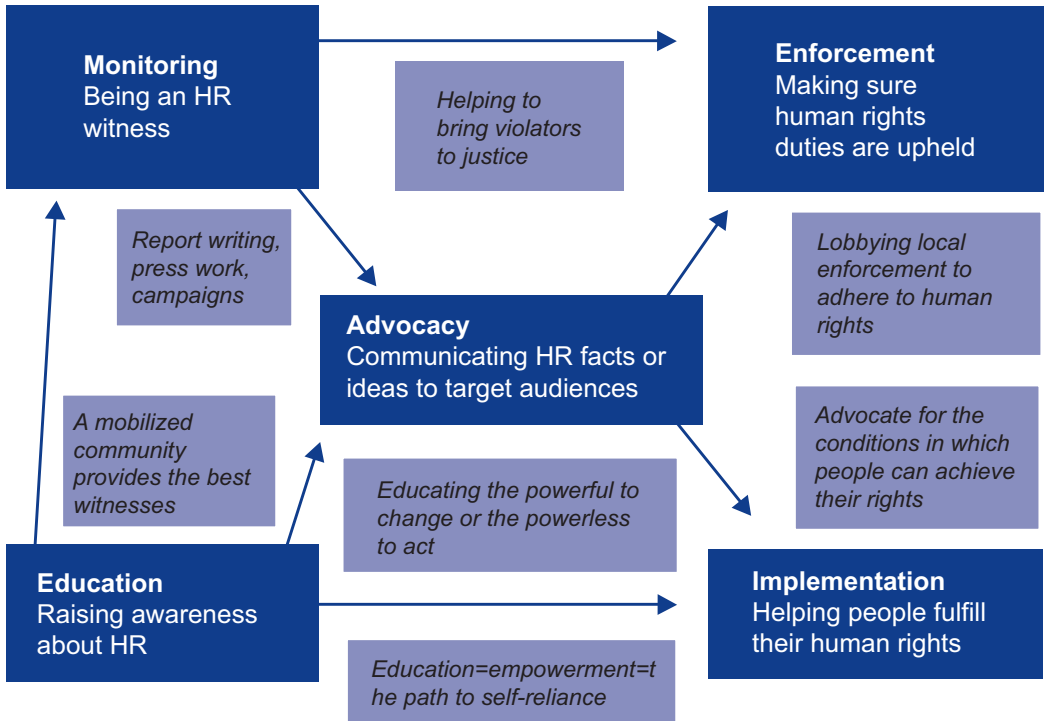
Source: O'Brien, Paul and Jones, Andrew, Human Rights and Rights Based Programming Training Manual (Nairobi: Care, 2002).

In terms of evaluation approach, the Human Rights Strategy Web reminds us that, even if a programme only employed the strategy of voter education or of civil society advocacy, we need to consider the interplay of this strategy with other strategies being pursued by actors who are 'outside' the programme if we are to grasp fully the synergies and the impact. If a monitoring programme provided the missing link in the given situation, and fed valuable testimonies and evidence to policy advocacy and litigation, then the impact of that alone could be significant. However, where the monitoring programme acts in isolation, with few linkages to or uptake by the other strategies, then immediate impact could well be low. Still, in a longer-term perspective the collection of evidence frequently in itself spurs advocacy and litigation initiatives. Evaluation of the impact of a civil society advocacy programme would need to go beyond considering the immediate improvements brought about for the constituency. It should look also at changes in the power relationship between citizens and the state, and see if (human rights) safeguards and legitimate channels of participation have been established. Thus, the RBA is also valuable when considering the third pillar in the IDEA Democracy Assessment Tool, namely civil society.[2]

Outcome and impact

With the rights-based approach we are guided to assessing impact in terms of changes in a relationship, namely the interface between rights-holders and duty-bearers. The definition of impact therefore moves beyond the traditional notion of impact that is employed in development cooperation, which focuses on improvement in people's lives. Human rights impact materializes in terms of enjoyment of rights—'respect, protect and fulfil'—and the process rights of non-discrimination, participation and accountability (see also the section on Process rights below). 'Respect, protect and fulfil' encompasses the notion of both obligation and right, and 'enjoyment of rights' is thus a relational term.

The need to move beyond the traditional definitions of development impact was strongly voiced at a workshop in 2001 (reported in Madsen 2001), which brought together 15 human rights organizations from around the world to consider how to assess the impact of human rights work. The workshop report concluded that further work should be undertaken to ensure that impact assessment methodologies reflect the characteristics of rights. The workshop supported the proposal that there was a need to:

> Revisit the very impact definition in a human rights perspective and consider that basically *a change in human rights* should be assessed with departure in
> – The *trinity* of respect, protect and fulfil, which together denotes the enjoyment of a right
> – The *relational* character of rights (duty-bearers–claim-holders) implying that an impact assessment should not only be undertaken of changes in the lives of people, but also and in particular of the *relationship* between duty-bearers and claim-holders vis-à-vis the respect, protect and fulfilment of various rights (Madsen 2001: 1).

We see that programme outcomes typically manifest themselves in terms of enhancing capabilities to comply or to claim. Moreover, positive changes in the human rights system are here considered to be an outcome, which may translate into enhanced actual compliance and enjoyment. The levels of outcome correspond with indicators as is illustrated in table 5.2.

Table 5.2: Human rights indicator levels

Level	Indicators measuring:
Impact Enjoyment Obligation of result	Actual enjoyment Improved respect for human rights Improved protection of rights Improved fulfilment of human rights Enhanced non-discrimination Right to participation institutionalized Enhanced accountability mechanisms
Outcome I Conduct	Changed conduct in terms of policy, programmes and practice that comply with rights obligations Changed conduct in terms of contestation and claims
Outcome II Capabilities	Duty-bearers' capability to comply Rights-holders' capability to claim and access Human rights guardians' capability to oversee compliance Human rights defenders' capability to promote compliance and support the empowerment of rights-holders
Output	Training conducted for duty-bearers Human Rights Commission established Legal awareness sessions conducted

With the above levels of impact, outcome and output in mind it becomes clear that most of the evaluations actually remain at the level of measuring improved capabilities—that is, the outcome level—and do not move on to the level of impact. The central question whether the improved capabilities actually translate into improved respect for, and protection and fulfilment of, rights remains unanswered. The global review of the OHCHR mentions increased capabilities within the administration of justice; the *Danish NGO Impact Study* mentions the strengthening of civil society; and the Danida democracy and human rights evaluation highlights empowerment at the citizen's level, changes at the regime level and performance changes among certain institutions.

However, we do not know how that plays out in a given context and situation in terms of greater enjoyment of rights. The collective experiences within the human rights movement show many examples of human rights organizations developing very good capabilities and strategies to promote and defend human rights, without there being an immediate resulting improvement in the human rights situation. In fact, the enhanced capability of the human rights movements may be met with countermeasures from several parties. The result then could be an increase in the number of violations of the rights of the human rights defenders, increased victimization of the groups they try to protect, and the passing of legislation that will allow greater surveillance and

interference with the right to privacy and so on. The negative spiral may be temporary and could be changed into a positive spiral in situations of growing international pressure or changing domestic constituencies. However, a true change often requires the capabilities to comply to be built up as well, that is to say, enhancing both sides of the rights equation simultaneously.

As the RBA carries a focus on power relations, a conflict perspective is inherent. For this reason power profiling tools, change agent analysis, risks assessment and safety and security mapping are extremely useful both in the design and in the evaluation phase.

The most challenging aspect of the rights-based perspective on impact is the notion 'obligation of result'. In international human rights law, states are responsible not only for their conduct but also for the results. Governments have the obligation, within all available resources, to ensure their citizens' enjoyment of their rights. If national resources are insufficient there is a duty to seek and extend international cooperation to that effect. This takes the notion of responsibility much further than the traditional development paradigm. The development logframe stipulates that the programme management is only directly responsible and accountable for the output produced by the programme. The outcomes lie beyond the direct control of the programme, and impact is a combination of factors that is beyond the ability of a project to determine. The rights-based approach challenges this narrow range of responsibilities. Governments are obliged to seek results that enhance the enjoyment of human rights, and if the outputs in the longer term produce a negative impact then mechanisms of complaint and redress should be established.

Selecting the data sets

The fact that the interface between rights-holders and duty-bearers or between electorate and representation is the key relationship to investigate in human rights and democracy support has implications for the choice of evaluation tools and the selection of the data sets. It will be important to have data that mirror or provide indications of the relationship between the rights-holders and the duty-bearers.

First, the event-based data should mirror reported acts of violations of human rights. This means violations committed by both state and non-state actors, and to an increasing degree the monitoring is not only of civil and political rights but also of economic, social and cultural rights, too. Although event-based data have been a major instrument in the human rights struggle for decades, in the quest for an end to violations and impunity, they are equally important when evaluating the impact of human rights and democracy assistance. The computerized Events Format developed by the Human Rights Information Documentation System (HURIDOCS) is particularly useful as it helps trace systematic patterns of violation as well as the types of violation that are most common. The casework being developed by the FoodFirst Information and Action Network (FIAN International) on violations of the right to food in several countries holds special promise in this respect. However, the

overarching weakness is that the various monitoring bodies and institutions do not currently employ a consistent and coherent format that will allow for the aggregation and cross-referencing of their information (see <http://www.huridocs.org>; and <http://www.fian.org>).

Second, the perceptions-based data that gauge people's perceptions of their relationship to the rulers or to the duty-bearers are also very important. To illustrate, fear of repression is often as effective as manifest repression in influencing human behaviour, and only perceptions-based tools will fully grasp its significance. Examples are self-censorship or 'shadow voting'. Low levels of public trust in the judiciary and in the rulings of the courts greatly influence which avenues people choose to gain redress in situations of conflict. Similarly, perceptions of the political systems have greatly influenced the growing voter apathy that we see developing in many European democracies. It is encouraging to see that perceptions-based data are gaining ground in several circles, ranging from the employment of scorecards among communities and the ranking of health services delivered by health clinics in Malawi to World Bank involvement in large-scale surveys portraying the voices of the poor. Yet, it is relatively rare for perceptions-based data sets to be employed in evaluations.

Process rights

Impact assessment is not only concerned with respect, protection and fulfilment within substantive human rights standards such as the freedom of expression, freedom from torture or the right to food. It also concerns the so-called process rights of non-discrimination, participation and accountability, which are among the most fundamental human rights principles.

Non-discrimination is one of the main basic principles of human rights enshrined in all conventions and is often called the rule of thumb number one of human rights promotion—seeking to promote equal opportunities and combat discrimination in all forms. Non-discrimination is also the basis for democratic development; it is enshrined in the standard principle of 'one voice one vote'. It is for good reason that the protection of minorities is found to be an important proxy of substantive democracy as distinct from procedural democracy. In terms of programme design, the process rights of non-discrimination require us to move beyond the target of providing voter education to, say, 80 per cent of the electorate or of districts and to give more attention instead to the non-discrimination imperative, which of course calls for attention to the remaining 20 per cent. Similarly, in an impact evaluation it will be crucial to assess the impact on patterns of discrimination. To allow for the tracing of discrimination on the grounds of gender, religion, language and so on, indicators and other proxies used in the evaluation need to be disaggregated to the greatest extent possible (Madsen 2003).

Assessing participation in a human rights perspective goes beyond measuring such things as the take-up rate of the various activities of a project, or the number of

people organized in the project committee. This perspective was used in the Danida NGO Impact Study. Rather, it means assessing the extent to which institutionalized processes have been put in place whereby people's participation is recognized and respected by all stakeholders, whether it is mandatory for the progress of the project, and whether redress measures are in place. Distinct notions of democracy are thereby ingrained in the human rights principle of the right to participation.

Accountability is the third main principle originating in the very notion of obligations. It lies at the heart of the relationship between the electorate and the elected. Accountability is the opposite of impunity, which appears to be widespread in many countries. Accountability of the government is of primary importance, but the demand for accountability also concerns all other actors, individuals as well as NGOs. It concerns both vertical accountability mechanisms—between the state and the citizens—and horizontal accountability mechanisms, that is, between groups of citizens or between members of an NGO and the governing board of the NGO. It implies the establishment of procedures for holding parties accountable, including avenues for presenting complaints and gaining redress.

Accountability has been defined (Humanitarian Accountability Project 2005) as involving two sets of principles and mechanisms:

• those by which individuals, organizations and government account for their actions and are held responsible for them; and
• those by which individuals, organizations and states may safely and legitimately report concerns, complaints and abuses, and get redress where appropriate.

Thus, in practice the focus must be trained on accountability mechanisms, whereby such mechanisms become much more than management tools: they are powerful tools in the quest for maintaining a balance or re-establishing a balance in the rights equation. This is also stressed by Mokhiber: 'Accountability means beginning with the identification of (1) an explicit standard against which to measure performance, (2) a specific person/institution owing performance (3) a particular right-holder (or claim-holder) to whom performance is owed; (4) a mechanism of redress, delivery and accountability' (Mokhiber 2001: 127).

The process rights have implications for both programming and evaluation. Any human rights and democracy intervention should ensure that the intervention itself is designed to give form to and realize the process rights and that the programme monitoring looks at adherence to the three process rights. Moreover, the changes brought about in strengthening accountability mechanisms beyond the accountability of the programme—that is, within society—must be investigated too.

The use of indicators

The main difficulty with indicators has in part resided in the difficulty of bringing together state-of-the-art frameworks for the situation analysis and frameworks for the change analysis. The same difficulty has recently been experienced within the development indicator debate, where it is hard to find congruence between situation, performance and impact indicators. Also within the human rights field there are great difficulties in reaching agreement on a common framework, both because of technical difficulties and because of the variations among the institutional perspectives of different actors. As democracy is about power, so is democracy support very much driven by different interests; the level of agreement on common frameworks is, as on many other matters, itself a product of politics.

The UNDP's Indicators for Human Rights Based Approaches to Development in UNDP Programming: A User's Guide (United Nations Development Programme 2006) is an example of one institution's attempt to employ a rights-based approach in situation analysis and programming and in monitoring and evaluation. While there is still room for improvement in the conceptual clarity and a need for the approach to be tested, it goes a long way towards binding together the situation-based indicators and the programme change-based indicators, which in the past have often been kept separate.

The United Nations Development Group (UNDG) also aims to develop indicators for rights-based development. This is expressed strongly in the following guidelines:

> Approaching development from the perspective of human rights creates particular demands for data that are not satisfied by traditional socio-economic indicators alone, and requires the selection and compilation of indicators on the basis of the following principles: (a) internationally agreed human rights norms and standards that determine what needs to be to measured; (b) a comprehensive human rights framework with sectors mirroring civil, cultural, economic, political and social rights; (c) integration of the 'rights element' into existing indicators by identifying (i) explicit standards and benchmarks against which to measure performance, (ii) specific actors or institutions responsible for performance, (iii) rights-holders to whom responsibility is owed, and (iv) mechanisms for delivery, accountability, and redress; (d) measuring subjective elements, such as levels of public confidence in institutions of governance, including among vulnerable or marginalized groups. All relevant indicators should be disaggregated, to the extent possible and where appropriate, by race, colour, sex, language, religion, nation, ethnic, or social origin, property and disability and other status such as woman or child head of household, etc. (United Nations Development Group 2003: 33).

While indicators should always be crafted to suit the particular purpose or programme, the RBA Navigator helps us pinpoint those dimensions of indicators that are indispensable if we want to say something about human rights and the state

of democracy or changes in these. The core indicators are as outlined:

- respect, protect, fulfil (both substantial and process rights);
- capabilities (recognition, authority, resources); and
- the human rights framework.

It is important not to rely too much on indicator-based monitoring and evaluation. Other forms of assessment may be used too, such as most significant change (MSC), context-in analysis, and appreciative inquiry.[3] This goes for all types of development intervention, but it may be valid in particular for human rights and democracy programming due to a number of factors.

- Human rights and democracy support do not necessarily lead to a situation of greater comfort or tranquillity. On the contrary, human rights activities, just like democratic rules, may lead to turmoil and discomfort as they have the effect of spotlighting conflicts and awkward power relations in society.
- Due to the relational character of human rights and democracy, an otherwise very effective human rights organization may not be able to impact on the forms of repression of governments. The increased ability of citizens to defend their rights may be countered by the government. On the other hand impact is produced by the mere presence of human rights organizations, which may prevent repressive action on the part of governments or on the part of non-governmental entities.
- Impact cannot always be anticipated. In human rights work processes and forces are set in motion that cannot be controlled easily. From a donor perspective this implies taking the risk of assisting a movement without knowing the ultimate results. It also implies taking the risk of supporting those who put their own safety at risk because they are willing to work for change (Madsen 1998).

Broadly speaking, then, human rights and democracy assistance is sometimes characterized by unexpected outcomes and by disputes among the actors as to the meaning and significance of certain outcomes. This limits the usability and relevance of predetermined and standardized indicators, as table 5.3 illustrates.

Table 5.3: The usability of indicators

Outcomes	Expected	Unexpected
Agreed meaning	Indicators are useful	Indicators are unlikely to be developed
Disputed meaning	Indicators, if developed, are of limited use	Indicators cannot be easily used

Source: Adapted from Rick Davies at a seminar on Most Significant Change, held in Copenhagen in 2004, with grateful acknowledgements.

While there are disputes over the meaning of both democracy and human rights, it is generally acknowledged, as is stated above, that human rights are internationally accepted norms that bind all member states of the United Nations. Human rights norms thus fall into the category of having a higher level of agreed meaning than democracy. Because of this, explicitly using human rights as one pillar in democracy evaluations will help move assessments in the direction of more generally agreed and more useful indicators. This of course presupposes a consistent use of the rights-based framework in the evaluation exercise.

However, it should be kept in mind that indicators are no more than tools to help diagnose a given situation (provide the baseline) or help identify the (intended) change brought about by democracy and human rights assistance—or any other assistance or policy measure for that matter. A quick perusal of human rights project documents from various international donors shows that the work on indicators is often confusing. This is because the design, programming and identification of intended change are not made sufficiently clear, and in some cases because of weak conceptual and strategic clarity in the given intervention field.

Recently, there has been a strong preoccupation with the search for indicators within both human rights, and governance and democracy. Some notable works include the Metagora Project,[4] the World Bank Governance Indicator Project, and the UNDP RBA indicators. In total this work presents a major leap forward, but it is striking that in general the debate and scholarly work around indicators have been quite disconnected from the consideration of evaluation methodologies and challenges within human rights and democratization support.

Many of the frameworks for indicators now being developed are not compatible because they use different definitions, terminologies and levels of aggregation. We see major differences not only between the human rights and development communities but even within the human rights community itself, and there is considerable variance between different but equally authoritative human rights mechanisms in the UN, too. What are the development agencies expected to make of this? A former UN special rapporteur on the right to education offered one answer when she suggested a monitoring format according to the dimensions of availability, accessibility, acceptability and adaptability (the 'Four A's'). Alternatively, the latest guidelines issued by the OHCHR on a human rights approach to PRSPs with regard to the right to education (United Nations, Office of the High Commissioner for Human Rights 2003) present eight key targets with 23 corresponding indicators.

While these two formats in no way contravene the letter and intention of the relevant articles of the conventions, it would be beneficial to arrive at a greater degree of consensus on which monitoring and evaluation formats key human rights institutions are using (United Nations, Office of the High Commissioner for Human Rights 2003).

The confusion has been exacerbated by recent suggestions from the UN special rapporteur on the right to health that consideration should be given to monitoring human rights by use of structural indicators, process indicators and outcome

indicators (United Nations, Economic and Social Council 2006). This is a basic typology of indicators that can be applied to any area of investigation, whether private-sector development, human rights or democratic development. It basically mirrors the three steps in any given intervention: structural indicators reflect the existing system, mechanisms and institutions; process indicators reflect the actions, policies or interventions being implemented; and outcome indicators reflect the impact (in this case, on health). But it will not help bring out the important and constituent dimensions of human rights and democratization, and it could risk reproducing the gulf between situation and change indicators. Many outcomes actually become manifest within the structural category inside institutions and mechanisms for human rights protection.

It would therefore be a major step forward if the RBA brought actors and institutions together in a common framework. This is happening already with the UNDG Common Understanding, which is developing impact indicators on the basis of the obligations of 'respect, protect and fulfil' of both substantial rights and procedural rights as well as outcome indicators relating to the three dimensions of capabilities to comply and claim. In this context the RBA Navigator is generic and covers civil and political rights, as well as economic, social and cultural rights. It covers compliance both at the national level and at the international level and is fully compatible with the compliance framework now being developed for non-state actors, in particular the corporate sector (see the Business and Human Rights Project, <http://www.humanrights.dk>). Once this platform is established many other indicators may be developed as they become relevant in concrete situations.

Applicability

When the RBA was first considered in development programming a great deal of concern was expressed quite naturally with regard to its applicability. Many years have passed, and agencies have had time to assess the feasibility of the RBA: some have taken the decision to make rights- and results-based programming mandatory. UNICEF is an example.

Now we are slowly starting to see initiatives to harvest the lessons learned and to review the processes and results of rights-based programming. However, due to the relative novelty of the RBA and the stepwise and gradual employment by most agencies of this approach, we are not as yet seeing many examples where a full RBA has been used full circle—that is, from situation analysis to programming design to implementation to evaluation. For obvious reasons there is more accumulated experience of the applicability of the RBA in the first stages than in the last phase of evaluation. There are, however, a few examples, some of which are mentioned below.

In 2001 the British Government's Department for International Development (DFID) initiated an appraisal-cum-evaluation with a strong rights-based focus on a

sustainable livelihood programme in Malawi, which was also very much focused on local-level democracy. The appraisal highlights the strengths and challenges of the proposed project from a rights-based perspective. Moreover, it recommends that the design of the evaluation and monitoring system should align with the rights-based approach:

> Presently, Oxfam seems to focus on measuring impact at household level, but the purposes of the project call for impact assessment at several levels. Moreover, the human rights approach specifically calls for assessment in the changes that occur in the relationship between duty-bearers and rights-claimants at all levels—be it between the District Assembly and the VDCs (Village Development Committees) or between the village headman and a member of the community. A human rights approach would not call for impact assessment with regard to people's ability to meet basic needs only, but also for the extent to which the right to food, water, education, etc. was respected, protected and fulfilled (British Department for International Development 2001: 12).

FIAN International has been central to the development of a rights-based approach to food and in promoting the right to food at both the international and the local level. FIAN International clearly employs a rights-based approach, and a Sida review of FIAN's programmes used the RBA Navigator to determine their strengths and weaknesses, in both strategies and results (Madsen 2004).

Save the Children UK has worked with a child rights programming approach for many years and has also been making efforts to develop a monitoring system that will reflect the child rights programming approach. The Global Impact Monitoring system operates with five dimensions of change. This has the potential to measure impact in terms of (a) children's enjoyment of rights ('respect, protect, fulfil'); (b) outcomes relating to changes in children's and communities' capability to claim and defend; and changes in duty-bearers' capability to comply, including changes in the national human rights legal and policy framework; and, finally, (c) two dimensions on outcomes relating to the process rights of children, namely non-discrimination and participation (Save the Children UK 2004).

With regard to the use of the RBA in the design of democracy support, there are no indications that progress has been any less marked. Several large multilateral agencies working in democratic governance support have committed themselves to human rights mainstreaming and to moving further into the employment of a rights-based approach. The UNDP is a prominent example. Several national governments too have used the rights-based framework in their development of national development programmes and poverty reduction strategies. Democracy and human rights strengthening plays an important part in these both as a specific plan and as a perspective that cuts across the main development sectors within areas such as health and education.

Caution is needed when considering the merits of the RBA in programming and evaluation. Before assessing its added value we need to ascertain that the programming actually carries the distinctive features and characteristics of a rights-based approach. A certain threshold is needed before we can say that a programme merits being called rights-based. The RBA Navigator provides this basic threshold. In many situations, owing to the factors described above, only some dimensions of the RBA are operationalized, which in turn implies that we cannot expect the same outcomes as from a more holistic approach. UNICEF has clearly demonstrated this. Among initiatives focusing on drawing lessons that can be applied to the operationalization of the rights-based approach, UNICEF has been particularly vigorous in conducting a series of case studies of projects and programmes and drawing lessons from them. Its experience clearly shows that some RBA dimensions are more easily operationalized than others. For example, the fundamental interface between rights-holders and duty-bearers is poorly operationalized: 'The relationship between duty bearer and rights holder lies at the heart of a rights-based approach. Duty bearers are responsible to respect, protect and fulfil the entitlements and freedoms of rights holders. Many of the case studies do not sufficiently understand this important relationship. This weakens their analysis, strategies and results' (United Nations Children's Fund 2004: 5).

The review of the UNDP's Global Programme on Human Rights Strengthening (Hurist) concluded in a similar manner. While some dimensions of the RBA were being actively pursued, in particular the process rights of non-discrimination and participation, others, such as the substantial standards of the various human rights and the corresponding obligations of 'respect, protect and fulfil', were much weaker. The relational and vertical intervention model that targets both duty-bearers and rights-holders around a particular human rights issue was rarely employed (United Nations Development Programme 2004).

The UK Interagency Group on Rights-Based Approaches also recently initiated a large multi-country study to identify the lessons from the RBA and to seek to document the achievements gained by employing this approach. The study is thought-provoking, because the methodology and analytical framework are informed by fields of investigation derived not from the RBA but instead from a combination of categories derived from various development paradigms, including the livelihoods paradigm (that is to say, asset accumulation).[5]

To summarize, until now more attention has been paid to the rights-based approach in programming than to what the RBA brings to a programme evaluation perspective. It is therefore too early to document the lessons for applying the rights-based approach in evaluations of democracy support. But, as is mentioned at the beginning of this chapter, we now see many examples of democracy assessments in the form of situation analysis or trend analysis, where rights constitute a fundamental pillar.

The rights-based approach and evaluation standards

Many practitioners have called for a greater focus on quality standards within evaluations and evaluation design, in the form of accuracy standards, feasibility standards, propriety standards and utility standards. Other practitioners, and some scholars, too, have made a contribution to best practices within the conduct of evaluations, covering issues such as participatory methods and process use of evaluations (Forss 2002). What does the RBA imply in terms of quality standards and evaluation design? As yet this question has not been subject to much debate or investigation. However, Theis makes a valuable point in saying that 'A rights-based evaluation is not just a technical exercise in data collection and analysis. It is a dialogue and a democratic process to learn from each other, to strengthen accountability and to change power relations between stakeholders' (Theis 2004: 104). By adapting the rights-based perspective to the evaluation exercise itself we can arrive at three main observations.

- The evaluation itself and the results operate within the *relationship between rights-holders and duty-bearers*, and may impact on the course of development. The inherently political nature of evaluation and the power associated with knowledge must be acknowledged and handled accordingly.
- The *interface between rights-holders and duty-bearers* is the epicentre in the evaluation and calls for the active engagement of both sets of actors, or sides of the equation, in the evaluation itself.
- The *focus on process rights* means that the evaluation design and conduct must be transparent and accountable, making the evaluation results public to all affected parties.

The RBA places emphasis on what is covered by the evaluation propriety standards—the fact that evaluations affect people. The propriety standards are intended to protect the rights of individuals. They promote sensitivity to and warn against unlawful, unscrupulous, unethical or inept action by those who design and conduct evaluations. Quality standards to protect the people engaging in or affected by evaluation are very rarely mentioned in terms of reference for evaluation assignments, by donors and NGOs. An exception can be found in the Minority Rights Group International (MRG) practice of issuing terms of reference that call explicitly for respect for the integrity of the partners and minority groups with whom the MRG works during the conduct of evaluations. The process rights of non-discrimination place emphasis on the most vulnerable groups and give them voice in the evaluation process. It is positive to note that guidelines on good practice in evaluations often consider issues of protection of information sources, equal treatment, and informed consent (UK Evaluation Society 2003). But there is no doubt that further useful reflection and development in the field of evaluation standards can be inspired by the RBA.

Conclusions

In this chapter the case for using the analytical framework within the rights-based approach in the evaluation of democracy support has been argued together with the case for a more systematic use of a rights-based perspective in the design of democracy and human rights interventions. The RBA is not offered as a 'grand strategy' or blueprint for democracy support and evaluation that could make all analytical approaches, frameworks and methodologies redundant. Rather, the point is to use the RBA framework as a common point of departure and to relate the findings, data and observations of other tools to the RBA, modelled as it is around the constituent characteristics of human rights. We may say that the RBA, like human rights, is universal, but the mode of implementation and operationalization is situation- and context-specific.

Human rights constitute a fundamental pillar in most notions of democracy and enjoy a higher level of international standard-setting and legal frameworks than is the case with democracy. The RBA provides the intersection between the democracy and human rights perspectives. In this regard, Gaventa is only one in a series of observers to note that 'there is a need to examine these debates and projects together to see how one strengthens the other' (Gaventa 2006: 24). The case for bringing together the frameworks for assessing the state of democracy with the frameworks for assessing democracy support has been highlighted, as well as the importance of paying more attention to the theory of change applied in democracy support evaluations.

The rights-based approach as agreed by the UNDG and as illustrated in the RBA Navigator provides a consistent framework for situation analysis, programming and evaluation. It can be applied at local, national and international level. Moreover it supplies clear guidance for understanding and categorizing the changes brought about by (programme) interventions and it brings consistency of indicators for multiple purposes. The RBA embodies the dynamic power relationship between duty-holders and rights-bearers, with the national and international human rights framework as the scaffolding. The actual strategies for promoting rights will have to be located in more context-specific analysis of how the capabilities and relative strengths play out in the given situation.

There is a major challenge in considering the new modalities of aid for democracy support and democracy support evaluation. At a recent Danida seminar reviewing experiences with general budget support it was mentioned that budget support could imply that the donors would have to retreat from such cross-cutting objectives as gender equality, human rights and democratization, and environmental protection.[6] In other words, the new aid modalities would give priority to national ownership and less scope for donors to impose their wishes through conditionality and so on. While this may be so, it will not necessarily mean that the case for promoting human rights and democratization 'from the outside' has become weaker. Two reasons can be given. First, the RBA is particularly well suited at the levels of overall sector-

wide approach and budget support to focus on mechanisms of regulation and the allocation of resources within a society. Second, the principle of national ownership that is cherished in connection with budget support carries with it a stronger focus on national responsibilities in terms of internationally binding human rights obligations. The emphasis may thus shift way from bilateral donor conditionality with questionable legitimacy, and towards accountability vis-à-vis the more authoritative human rights organs and mechanisms of the UN. Enhancement of domestic accountability is, however, the most crucial of all. Accountability lies at the heart of the RBA, and this is also a reason why the rights-based approach is being pursued so eagerly by civil society organizations, human rights defenders and national human rights institutions, and even by planning ministries around the world.

The academic preoccupation with the strengths and weaknesses of the rights-based approach, the questions as to whether it is merely a new fad in the practice of development, and doubts about how to operationalize it are to some extent founded on the perception that the RBA is a very new thing. In fact, the RBA just takes us back to the essential fact that the UN member states have voluntarily acceded to the goal of promoting peace, justice and development and have voluntarily signed up to a number of essential human rights conventions that regulate the exercise of power and the rights of citizens. Taking the argument to the extreme, we may say that rejecting the RBA means rejecting altogether the human rights framework and the obligations that go with it. Having taken on the human rights obligations, legitimate and accountable governments have for decades tried to design their national development programmes, annual budgets and performance measurements in accordance with them. Where there have been cases of regression or flaws, concerned citizens groups, political opposition parties, the media and independent human rights institutions have for decades subjected these governments to legitimate criticism.

The human rights obligations are so embedded in some countries that far too little notice is given to periodic reporting on performance and impact. However, when the general standard is not granted to certain groups, such as asylum seekers, refugees, migrant workers and so on, the issue then gains some publicity. It is interesting to note that the RBA framework is gaining renewed interest in domestic affairs in Europe, perhaps due to its significance in international relations. Three examples are the Norwegian White Paper on Agriculture and Food, which takes as its point of departure the right to food (Norwegian Parliament 1999); an independent rights-based audit of domestic developments in Ireland (Amnesty International); and a Swedish seminar on how to operationalize human rights within local municipal governance, held on 23–24 November 2006 in Malmö.

In countries where human rights compliance is not institutionalized, or where the rights equation is skewed and characterized by misuse of power and gross human rights violations, the rights-based approach provides an even more relevant platform for advocacy and action, programme design and evaluation.

Notes

[1] Gillies (1993) suggested a democracy assessment built upon key human rights.

[2] Windfuhr (2005) considers the use of the RBA in the fourth pillar, the international dimensions of democracy.

[3] Most significant change (MSC) was originally developed by Rick Davies in 1993 as a means of participatory impact monitoring. The MSC approach involves the collection and 'systematic participatory interpretation' of stories of change. Appreciative inquiry (AI) is a process for engaging people across the system in renewal, change and focused performance. The basic idea is to build organizations around what works, rather than trying to fix what doesn't. Appreciative inquiry was developed by David Cooperrider of Case Western Reserve University. It is a commonly accepted practice in the evaluation of organizational development strategy and the implementation of organizational effectiveness tactics. The context-in approach starts by identifying the changes in the context of a project and then seeks to draw the relationship (attribution) between these changes and the performance and outputs of the project. It turns the traditional impact chain bottom–up (Roche 1999).

[4] Metagora is a pilot project focusing on methods, tools and frameworks for measuring democracy, human rights and governance. Based on innovative initiatives, it aims to enhance proper assessment methods. Metagora is being implemented under the auspices of Paris 21, a consortium hosted by the Organisation for Economic Co-operation and Development (OECD) which aims to foster more effective dialogue between the producers and users of statistics on development issues (<http://www.metagora.org>).

[5] Presentation of the study awaiting publication by Sheena Crawford at the seminar entitled Rights Based Approach to Development, Copenhagen, November 2006.

[6] General Budget Support Seminar, Ministry of Foreign Affairs, Copenhagen, 16 May 2006.

Chapter 6

**Evaluating a democracy
support evaluation:
the Rights & Democracy
ten-year taking stock exercise**

Chapter 6

Michael Wodzicki *

Evaluating a democracy support evaluation: the Rights & Democracy ten-year taking stock exercise

This chapter traces the evolution of evaluation activities at a Canadian institution with 15 years of experience in democracy support. In particular, it evaluates a 'taking stock exercise' conducted between 2000 and 2002, an internally-led evaluation of the institution's first ten years of democracy support activities. Using a comparative, participative and qualitative methodology, the taking stock exercise aimed to propose a strategic vision and approach for support to democratic development. Drawing on interviews and available documentation, and examining the actual follow-up to the recommendations of the exercise, the methodology is found to be of mixed usefulness. While some links are established between the objectives of identified democracy support activities and their eventual impact, concrete guidance for future programming is less evident. The chapter offers three lessons learned, relating to participative and qualitative evaluations of democracy support activities.

The nature of Rights & Democracy's work made it *somewhat difficult to assess its contribution:* …where its capacity building activities may have lead to important changes in confidence, contacts and influence, which is difficult to measure. There are other actors that do similar work, some with greater human and financial resources than Rights & Democracy. *It is difficult to isolate the effects of Rights & Democracy's interventions* when these other actors are often present and sometimes

working in cooperation with Rights & Democracy (Canadian Office of the Inspector General 2003: 3, italics added).

Introduction

If human rights are protected and promoted, democracy will flourish. Democracy, therefore, as stated by former United Nations Secretary-General Kofi Annan in his report In Larger Freedom (United Nations 2005), must be pursued around the world. It is not a question of certain human rights, say, only civil and political rights often linked to elections, but rather of all rights as defined in the International Bill of Human Rights.

The indivisibility of human rights has been at the heart of the work in democracy support of the Canadian organization Rights & Democracy (R&D) for the past 15 years. The recurring challenge has been how to transform this well-established and now generally accepted link between human rights writ large and democracy into effective support for the democratization of countries. Indeed, it is difficult to say with authority, as the above quotation from R&D's latest five-year external review states, that the populations in many countries where R&D has been engaged are substantially better off today than they were 15 years ago, that their governments are more democratic or that their human rights records have improved thanks to R&D's work.

R&D is not alone in this critical self-assessment. There is widespread acceptance by many practitioners and observers of democratic development that the most recent generation of democracy promotion activities have not been as effective as originally hoped[1]. Expected transitions to democracy have more often than not ended up as 'transitions to nowhere' (Brumberg 2003). Democratic systems may well have been dangerously substituted by fecklessly pluralist societies (Carothers 2002). Consequently, many donors and democracy-supporting institutions scramble to understand better how and whether their work impacts on the long-term development of democracy.

As an institution that has attempted to support the development of democracy around the world, R&D has consistently evaluated the impacts of its work. This chapter reflects on those evaluations and aims to contribute to the larger discourse on measuring the effects of democracy support activities. It does not discuss specific indicators of democracy, nor does it contrast and compare different methodologies for evaluating democracy promotion. Rather its objective is to analyse R&D's experience evaluating its Democratic Development programme and to present what it has learned from that experience. Specifically, the chapter will present how R&D used a comparative, participative and qualitative methodology to undertake a ten-year review of its democracy support programming in 2000/2001.

The chapter is in three parts. The first reviews R&D's conceptual and practical approach to democracy promotion. Understanding how R&D approaches democracy will help elucidate how R&D has evaluated its work, in particular the ten-year taking

stock exercise described in the second part. The chapter concludes with a discussion of the usefulness of this methodology and the conclusions offer some final observations about measuring democracy promotion.

The Rights & Democracy approach to democracy promotion

Rights & Democracy (the International Centre for Human Rights and Democratic Development) is an independent Canadian institution created by an act of Parliament in 1988 (Canada, International Centre for Human Rights and Democratic Development Act 1988). It first opened its doors in 1990. It has an international mandate to promote, advocate and defend the democratic and human rights set out in the International Bill of Human Rights. In cooperation with civil society and governments in Canada and abroad, Rights & Democracy initiates and supports programmes to strengthen laws and democratic institutions through programmes, advocacy and research, principally in developing countries. It divides its work into four thematic programmes—democratic development; women's rights; globalization and human rights; and the rights of indigenous peoples. R&D also manages projects in developing countries funded by the Canadian International Development Agency (CIDA). Sizeable programmes are currently in place in Haiti and Afghanistan.

How does Rights & Democracy promote democracy?

Democracy was and still is understood by R&D as a long-term, dynamic social process and not just a series of institutions or periodic elections. Human rights are therefore a constitutive element of democracy. Democracy develops as rights are fought for through the 'institutions' of democracy—civil society, political parties, the media, elections, legislatures, and various state institutions. This struggle is rooted in socio-cultural contexts, occurring differently in every country and society. For the most part, R&D's efforts to promote democracy have focused on support to civil society in developing countries as the major long-term guarantee of democratic development (Thede et al. 1996). Civil society has been defined as the sum of all non-family social institutions and associations that are autonomous and capable of significantly influencing public policy.

On the basis of this understanding of democratic development, R&D developed a human rights framework by which it could qualitatively evaluate a state's democracy, producing what R&D now refer to as 'democratic development studies' (Gillies 1993). The approach is qualitative in the sense that seven criteria were used to assess state recognition and respect for the whole family of human rights—participation, security, well-being, the national political economy, non-discrimination, the rights of collectivities, and state institutions. The resultant analysis was to serve as a tool for identifying key issues and actors for democratic development, in particular offering a portrait of civil society as it exists in relation to each category in the

framework (Thede et al. 1996). From this portrait, democracy promotion activities were developed in conjunction with local civil society actors to support a society's democratic development.

Between 1991 and 2005, Rights & Democracy used the framework explicitly to publish nine democratic development studies and implicitly in the development of its programming support to civil society organizations.[ii] This has translated into approximately 12.5 million Canadian dollars (CAD; c. 10.5 million US dollars) of expenditure on democratic development programming over the same period, disbursed among 540 projects in almost 50 countries. Currently, the Democratic Development programme at R&D operates on an annual core budget of about 1.2 million CAD, an increase from previous years, in about 15 countries.

Lessons learned: Rights & Democracy's evaluation experiences

Evaluations play an important role in helping R&D understand how it conducts its democracy support activities. There are four broad observations to be made about how R&D has evaluated these activities. First, R&D's qualitative interpretation of democracy has led to qualitative evaluations of its work. Second, R&D has consistently faced challenges in balancing this qualitative approach with the quantitative measures that are often required by its donors. Third, a qualitative understanding of democracy combined with R&D's broad mandate has often made it difficult to measure effectively the impact of its programming. Finally, the fact that R&D is a small institution has also impacted on the way in which it carries out evaluations. Projects have tended to be small in size (valued at under 100,000 CAD). This not only affects the size and scope of evaluations; it also means that drawing links between grass-roots civil society efforts, where most R&D project work occurs, and the full-blown democratic development of countries is often very difficult.

From R&D's first days, evaluations have consistently been part of its organizational culture. Like many other similar institutions, R&D is required in its founding law to have an external evaluation conducted every five years (Canada, International Centre for Human Rights and Democratic Development Act 1988, article 31). The Board of Directors can also request specific evaluations of individual programmes or projects. Project evaluations have generally taken three forms: self-administered evaluations led by project officers; evaluations conducted during field visits to projects using terms of reference designed by an external consultant; and external evaluations. Over time, self-evaluations of projects were folded into the project cycle and external evaluations became the primary method of evaluating projects.

The five-year external evaluations of R&D's work have sought to provide a broad overview of its work with reference to its mandates and objectives. This has included interviews with partners, but focusing less on the impact of R&D's work on democracy support in a given country and more on how successful R&D has been in meeting the institution's broad mandate.

While always broadly positive, the regular five-year external evaluations of R&D have all mentioned the difficulties inherent in evaluating R&D's ability to promote democracy. The first external evaluation of the institution in 1993 noted that R&D 'scatter[s] its energy in funding a great number of projects which, taken separately, are all justifiable but probably won't have much impact as a whole' (Brodeur et al. 1993: iii). In 1998 the evaluators commented that 'we are convinced that the Centre would be more effective, efficient, and accountable if it were to focus on more intermediate or "outcome" level objectives' (Universalia 1998: 39). The quotation from the 2003 evaluation, cited at the beginning of this chapter, alludes to such problems as well.

The Democratic Development ten-year taking stock exercise

In 2000 and 2001, R&D conducted an internally led 'taking stock exercise,' evaluating the first ten years of its democracy support activities (Thede 2002). The general objective of the exercise stated in the terms of reference was 'To extract lessons to be learned from the democratic development work of R&D over…10 years in order to consolidate and enhance [R&D's] institutional experience and to propose a strategic vision and approach for the future consideration of [R&D's] management and Board of Directors'.

Several specific objectives were identified. The exercise would serve to evaluate impact in countries where R&D had been engaged for the long term (five or more years). Particular attention would be paid to evaluating attempts to support democratic development in countries where a very limited internal democratization process was under way, such as Burma (Myanmar). The process was also to serve as a team-building exercise for members of the Democratic Development programme, although it was not made clear why such team building was necessary or how it would take place. Ultimately, the ten-year review sought to raise R&D's level of institutional understanding of and commitment to the specific demands and methods of democracy promotion.

Between July and October 2000, an inventory of all the democracy support projects implemented by R&D was created (Spuches 2000). In November and December 2000, a field assessment took place. Interviews and workshops were conducted with partners, local participants, and observers of R&D's democracy support activities. The results of the in-field assessments and analysis were brought together over the course of 2001, culminating in October 2001 with an inter-regional workshop on the exercise and the future of R&D programming. The total cost of the taking stock exercise was approximately 100,000 CAD, excluding staff hours spent on the evaluation.

The methodology of the exercise was comparative, participative and qualitative. Some observers of such an approach have argued that it leads to better definitions, or indicators, of what democracy support projects attempt to achieve in a country, as well as more effectively attributing the presence of these indicators to the projects in

question.[3] Let us consider each of these three properties in turn.

Comparative evaluations of democracy promotion are those that attempt to identify trends in success or failure across different countries. This is seen as beneficial for evaluations of democracy promotion because it might help illuminate the causal relation between democracy support activities and the level of democratic development, and the potential role of outside factors. If a project in one country is deemed successful and in another it is not, why did it succeed or fail? Was it really due to the usefulness of the project? Comparing similar projects across very different countries can help shed light on such questions.

Participatory evaluations, in terms of their planning, design, analysis and interpretation, will be just as important in themselves as their findings. Research has shown that participation in these phases of an evaluation can have significant and lasting effects on the knowledge, attitudes and skills of the people involved (Horton et al. 2003: 19). This is doubly true when one is talking about democracy support and building the skills of consultation, negotiation, participation and flexibility. (Stiglitz makes a similar case for economic development, stressing 'the importance of the processes by which decisions are made—how consensus building, open dialogue, and the promotion of an active civil society are more likely to result in politically sustainable policies and to spur the development transformation' (Stiglitz 2002: 177).)

For this to be the case, a participatory evaluation must seek the active input of all stakeholders. Rebien (1996) identifies three threshold criteria to be met for an evaluation to be counted as participatory: stakeholders must be involved as active subjects rather than only as sources of data; stakeholders should be involved in at least the design and data analysis phases of the evaluation; and at a minimum the involvement of the beneficiaries, field staff, intervention management and donor representatives is needed. Drawing on Rebien, Crawford (2003b) proposes a participatory and qualitative approach for democracy support evaluations, which would include a political context study undertaken in a participatory manner by local experts on the democracy trends in a country, in particular at the sectoral or 'meso' level.[4] It is at this level, Crawford argues, that participatory evaluators will be able to make more plausible connections between external support and overall political change in a country. In other words, by reducing the scope of the impact of democracy promotion activities, from national democratic change to change in certain sectors of democracy promotion, we will be better able to argue whether more or less democracy is being built.

Finally, a *qualitative* approach to evaluation is one that attempts to capture people's socio-economic and political beliefs, opinions, perceptions and narratives (Kapoor 1996: 7). Indicators of success would be developed by all stakeholders of the project. This is seen as beneficial to measuring the impact of democracy promotion because it tries to capture how people feel about their democratic rights; because it allows for more local ownership of the evaluation process; and because it can increase

the usefulness of indicators, since more stakeholders choose the relevant methods and criteria (Kapoor 1996: 8).

With respect to the taking stock exercise, what did these concepts mean in practice? Who participated in the exercise and in what capacity? How were comparisons made, and with what results?

In regard to being comparative, the taking stock exercise analysed the trends in R&D's democratic development work by comparing six case studies of partners and countries where R&D had significant programmes—Kenya, Tanzania, Burma, Thailand, Guatemala and Peru. Five of the six had been the subject of democratic development studies. The sixth, Burma, has to this day received more attention and financing than any other country from R&D. In addition, Guatemala had been the location of an innovative R&D-managed, donor-funded project focusing on public policy research. Comparisons took place within R&D once the case studies were drafted and with 'insiders' and 'outsiders' at a workshop organized by R&D at the end of the exercise.

The broad cross section of projects, partners and countries made the taking stock exercise an interesting comparative evaluation for a number of reasons. First, it was a review not only of individual projects but of the entire Democratic Development programme at R&D. Second, it was backward-looking, evaluating what R&D had achieved, as well as forward-looking, inasmuch as it sought to identify those areas where R&D should focus its democracy support activities in the future. Finally, it had a long-term perspective. Some of the case studies examined had ceased receiving support from R&D several years previously. The presumption at the time was that comparing R&D's experience in several countries would necessarily deepen the level of analysis that could take place.

How was it participatory? Both 'insiders' and 'outsiders' of the projects contributed to the terms of reference for the taking stock exercise, the preparation of case studies, and the comparative analysis of the findings. 'Insiders' included people who had received R&D financial support, participants from workshops and conferences held with R&D support, and R&D staff members who had managed these democracy support activities. 'Outsiders' consisted of Canadian experts in the field of democracy promotion and national observers of democracy support in the countries in question, such as academics, members of parliament and foreign aid officials.

At the outset of the exercise, the three regional officers in charge of R&D's democracy promotion programmes were interviewed (see the questionnaire at annex 6.1). Based on these interviews, eight former R&D staff members were interviewed, including the former president, as well as nine outside experts, including university professors and government officials. The results of this second round of interviews fed into the questionnaire that would be used with stakeholders during the field assessments (see the questionnaire at annex 6.2). For instance, in Kenya, interviews were held with members of parliament, members of civil society organizations that had received R&D support, Canadian Government aid officials, and the Kenyan Human

Rights Commission. In Tanzania, current and former members of the Tanzanian Legal Commission, R&D's partner at the time, were contacted, as well as other local non-governmental organizations (NGOs), officials at the Canadian Embassy, other Canadian NGOs in Tanzania, and academics. Following the completion of all the field assessments, nine R&D partners, R&D staff, R&D Board members and Canadian academics collaborating with the Democratic Development programme came together for a two-day workshop hosted by R&D at its offices in Montreal.

As regards the qualitative dimension of the exercise, the interviews at the outset of the process fed into the identification of a list of issues that framed the field assessments that were to be undertaken later in the year. One issue was the broad scope of democratic development itself and the challenge for an institution like R&D, small but with a broad mandate, to identify those key areas of democracy support where its work could have the greatest impact. This included an emerging problematic link between R&D's conceptual approach to evaluating democracy (that is, its human rights framework) and its programming. The framework itself would be re-evaluated: how well had it contributed to advancing democratic development? What had been its unexpected results? The last issue was of course the need to identify the specific factors that contributed to the success or failure of programmes or projects. These were such factors as local capacity, the nature of partnerships, strategic vision, and the nature of R&D's specific contribution.

These issues were reflected in the interview questions prepared for the field assessments. Responses to the questions were considered at the time as the 'qualitative indicators' of the impact of R&D's work. Questions focused on the interviewees' specific knowledge of R&D's democratic development work; their personal understanding of democratic development; their priorities for democratic development in the country and region in question; and how well R&D had met those priorities in comparison to the efforts of other organizations; and they sought any advice as to how R&D could achieve greater focus and improve the impact of its democracy support efforts.

The case studies represented an attempt to understand the qualitative impact of R&D's work in these countries, by comparing responses to the questions. Initial findings from the field assessments were discussed within R&D and with selected Canadian experts, with the goal of writing up the six case studies.[5] Case studies were drafted by R&D staff and were generally five pages in length, presenting the context of the country's democratic development, an assessment of R&D's support, and perspectives for the future. These case studies fed into the writing of Democratic Development 1990–2000: An Overview (Thede 2002), which published the conceptual findings of the taking stock exercise. Prior to publication, a draft document was discussed at an inter-regional workshop, mentioned above, on democratic participation hosted by Rights & Democracy in Montreal. Discussions centred on the results of the taking stock study and its findings. The workshop served to fine-tune the results of the study and helped R&D determine its programming between 2002 and 2005, and beyond.

The usefulness of the ten-year taking stock exercise

Did the comparative, participatory and qualitative methodology of the taking stock exercise help R&D better understand the impact of its work? Given R&D's experience, could such an approach clear up some of the complexity surrounding democracy support activities in general?

Certainly, the methodology did have its benefits. Inside and outside stakeholders were involved from the outset in the design of the terms of reference and issues to be analysed. Interviews and workshops in the field focused on producing what amounted to a sectoral level analysis of the situation in the country. Stakeholders were further involved in the analysis phase of the evaluation, as exemplified by the inter-regional workshop hosted in Canada.

The outcomes met some of the expectations. Links were identified between relatively small-scale R&D democracy support activities and the sectoral level of the democratic development process in countries. For instance, in Thailand and Guatemala, universities were still using the democratic development studies produced by R&D and its partners. In Kenya and Peru, civil society organizations that were in large part established by R&D and its local partners continued to have an important impact in their relative fields of expertise, even after R&D had ended its financial support.

In going into detail with partners about their understanding of democratic development, in general and in their countries, R&D was able to identify sectoral and broad democratic trends globally. The main conclusion was that the greatest problem facing a growing number of countries where elements of democracy have progressed is the question of exclusion: understanding and devising strategies for sustained participation of citizens was consistently proposed as the course of action for R&D's programming in democracy support.

The reality of the follow-up to these conclusions, however, indicated that the methodology and indeed the evaluation process itself created certain difficulties. The findings of the taking stock exercise fed into the strategic plan of R&D, but little else; only one paper was published on the relationship of indicators to concepts in democracy promotion (Thede 2001).

Why was the uptake from the exercise slower than expected? What were the weak points of the taking stock exercise? Examining the exercise five years on, there are two main reasons that emerge.[6] First, there were internal factors specific to R&D that hampered the follow-up to the evaluation. Second, the methodology had its own built-in problems; the development and identification of indicators had not been as effective as originally hoped and thus attempts to make useful comparisons suffered.

R&D underwent a significant institutional transition in the year that followed the taking stock exercise. Soon after publication of the review, its author and the president of R&D left the institution. One year later, the director of programmes changed as well. These staff changes, often unavoidable in any organization, contributed to the

lack of immediate follow-up. For instance, a series of workshops planned for Asia, Africa and Latin America that were to build on the study's findings and follow up to the workshop hosted in Montreal were postponed and finally cancelled, as they were no longer considered a priority.

In terms of indicators, translating the qualitative findings of the field assessments into tangible lessons learned proved very difficult. Kapoor has called a participatory approach an 'attempt to assess results through dynamic, negotiated consensus' (Kapoor 1999). In R&D's experience, drawing on workshops, dozens of interviews and hundreds of background documents with the aim of learning lessons that could meet with consensus within R&D was no small task. In its attempt to identify trends and lessons learned across such a broad spectrum of programmes, the taking stock exercise produced an overwhelming amount of information.[7] The many priority issues in democracy support that emerged included women's rights, the culture of democracy, political parties, indigenous identity, accountability, constitutional development, and the internal organization of civil society organizations (Thede 2002). Consequently, the exercise left R&D where it started—painfully aware of the broad nature of democracy support activities but with little guidance as to what activities were most effective.

It is difficult to determine which of these reasons, institutional transition or imperfect methodology, is the more important in determining the lack of uptake of the taking stock exercise at R&D. On the one hand, the departure of the author and coordinator of the exercise coupled with the arrival of a new president with new ideas for the institution led to a natural disconnect between the exercise and its uptake. On the other hand, given that the exercise had generated so much information, one could argue that more precise findings might have made it more likely that these findings would be incorporated in R&D's programming, regardless of institutional changes.

This leads to a third potential weakness of the taking stock exercise: insufficient attention was paid at the outset to ensuring that the process included not only time for reflection on the findings, but also a commitment to examine those findings in terms of how they could be incorporated into programming and, if they could not, why not. The challenge for R&D remains to fine-tune its evaluations in such a way that they can potentially provide more focused results; results that can be more effectively compared and measured again at a later date; and results that can more readily be considered useful guidance for programming.

Three key lessons learned from the taking stock exercise are therefore the following.

First, participation should have been more systematic, in particular as regards the participation of partners in the development of indicators. In other words, the process should have been more participatory, not less—more participatory in the sense that more partners, more 'insiders' from the field should have been included from the outset in planning the process and agreeing on indicators. Increased participation would be conditional on developing specific indicators for the evaluation in question lest the

process also produce an overwhelming amount of information, as was the case on this occasion. Thus, during the exercise, interviews, workshops and conferences should have been facilitated in a way that could encourage concentration on identifying specific sectoral-level changes that could trace the democracy trends in a country more effectively.

Second, more attention should have been paid to the quality of the participation that took place—not only 'how many partners participated' but 'how' they participated. The process of participation, of how well 'insiders' participate, will be key in determining whether better focus can be obtained in the evaluation. For instance, available documents seem to show that all interviews were conducted by R&D staff. Workshops held in the field were also facilitated by R&D staff. One can reasonably assume that this affected the quality of participation and hence the nature of the results. There is arguably a need therefore not only to develop (jointly) indicators of progress, but to develop indicators of participation in the evaluation process. For an organization such as R&D, which aims to build democracy through improved participation, such indicators become all the more relevant.

Third, at the individual project level, evaluations need to be more consistent, and it is essential that more money, time and flexibility be devoted to them. Similarly, the monitoring of projects as they are being implemented also needs to occur in a consistent and systematic manner. This would require a commitment from project managers to put aside sufficient resources for monitoring and evaluations at the outset of projects, and for this to cover the entire lifespan of a project. In addition, it would require effective record keeping of how these evaluations took place, who participated, and how stakeholders participated, and what the results were. This could potentially make broader programme evaluations such as the taking stock exercise more effective.

These lessons learned are slowly being translated into practice at R&D. Following an increase in R&D's budget, a full-time evaluations officer position was created in 2006. For new projects, the post-holder will work with the officers and partners in designing the monitoring and evaluation processes. The objectives are to make the participation of partners in evaluations more systematic, to improve the quality of that participation, and to keep better records of the monitoring and evaluations of projects. For projects already in place, the evaluation officer works with programme officers to implement these principles. The evaluations officer has an independent annual budget, which allows her/him to conduct a series of monitoring visits with programme staff, depending on project cycles and other factors. In time, the evaluations officers should be in a good position to deduce from their experience of working on several projects useful guidance on programming in general at R&D.

Two substantial projects in particular are good examples. First, in Côte d'Ivoire, R&D has developed a partnership with a coalition of human rights civil society organizations. At the outset, this partnership served to help these organizations produce a monthly publication on human rights violations in the country. Building

on this collaboration, R&D's partners requested a 'training of trainers' project on the role of civil society organizations in Côte d'Ivoire, targeted at their smaller members from the country's rural areas. During the initial week-long training session, a day was set aside to discuss evaluation of the project. R&D officers and workshop participants mapped out what could be termed a success in the project, how to measure that success, and when to do so. The expectations of the project were not determined by R&D alone but rather in a collaborative, participatory effort between R&D and its partners. Monitoring in the field will be conducted by the local trainers themselves, and discussed as a group at predetermined stages of the project. The objective is to build collective ownership of monitoring and evaluations.

The second project is a three-year project R&D has recently begun in Haiti. R&D and its partners are trying to determine what the trends are regarding civil society's place in Haitian society, extrapolating from the experiences of two well-established organizations. This project essentially attempts to establish a sectoral or meso-level analysis of democratic trends. In this case, R&D is supporting the efforts of two Haitian civil society organizations to systematize their advocacy experiences. This process follows the growth of these organizations, from within and from without, and their success and failures in promoting their respective agendas. Based on these experiences, lessons learned materials will be produced to work with other, smaller Haitian organizations. At all stages of the project, R&D and its partners have jointly determined measures of impact and success; indicators will be agreed and impact will be determined collectively over time.

Conclusion

This chapter has traced the evolution of evaluation of democracy support activities at a Canadian institution, R&D, with 15 years of experience in democracy support. This evolution was located in a contextual understanding of democracy support activities in which evaluation of these activities, or more precisely the difficulty of conducting such evaluations, was identified as a factor that makes democracy support a complex endeavour. Overcoming this complexity will in part require better measurement of democracy support: institutions such as R&D have to prove to their partners, their stakeholders, and their partners' stakeholders that they indeed achieve what they set out to do. The R&D ten-year taking stock exercise, a comparative, participative and qualitative evaluation of R&D democracy support experience, was presented as one attempt to do so. In evaluating this evaluation the chapter has arrived at certain lessons about measuring the effects of democracy support activities as they impact on the democratic development of countries.

In analysing the taking stock exercise, its mixed results have been highlighted. Its methodology proved effective in making some links between democracy support activities and the democratization of the country in question. However, the methodology of the exercise also had a built-in fault of sorts, inasmuch as it

produced too many results, providing little in the way of concrete 'lessons learned'. The level and quality of participation of partners in the taking stock exercise have been questioned and it is proposed that participation should have been more systematic. Efforts to overcome these problems have been presented in terms of improving how participation takes place during the evaluation process, and using two examples of current projects supported by R&D that attempt to do so.

Several observations can be drawn from this evaluation of an evaluation. A first observation is a methodological difficulty that was encountered in conducting the research for this chapter—incomplete information. The taking stock exercise took place five years ago. The information presented here is largely based on archived documents and interviews with R&D staff from the time. In researching such material, certain difficulties are almost inevitable. Available documents and interviews are of mixed usefulness. Evidently, more complete information could have added to the thoroughness of the analysis presented here. This holds true for any attempt to evaluate democracy support projects, or, as in this case, an evaluation of an evaluation of those projects.

A second observation is that improving participatory evaluation methodologies is both a practical and a conceptual affair. Practically, these methodologies need to err on the side of being more participatory. Consequently, they will take more time and money if they are to be effective, and donor organizations will thus need to be more flexible. This is true in part because democratic development itself is a long-term process, but also because in the medium term progress will be very difficult to measure in challenging environments such as those in Côte d'Ivoire and Haiti today. Conceptual considerations are equally important: indicators of progress must be developed from the outset of a project or programme in close collaboration between partners, drawing on local understandings of democratic development. Evaluation cannot be an external requirement of a project, but must be an internal desire of the local stakeholders in the project. This may also help overcome local scepticism regarding democracy support activities supported by outside actors.

A larger dilemma emerges, however—one that is unfortunately well outside the scope of this account. Examining democracy support projects broadly, one can assume that there are four kinds: those that are deemed 'successful' and result in more democracy; those that are not deemed 'successful' but nevertheless result in more democracy; those that are deemed 'successful', yet do not result in more democracy; and those that are not deemed 'successful' and there is no growth in democracy. Perhaps the problem therefore lies not so much in tracing the causality of projects to more or less democracy, but rather in the fact that there is not enough of an understanding of what more or less democracy means.

Notes

[1] This chapter uses the terms 'democracy support activity, 'democracy promotion project' and 'democratic development activities' to refer to any development project(s) that have the stated objectives of making a society or a country more democratic.

[2] On Kenya see Gillies and Mutua 1993; on Thailand see Taylor and Muntarbhorn 1994; on El Salvador see Rivas and Gonzáles-Suárez 1994; on Tanzania see Halfani and Nzomo 1995; on Guatemala see Palencia Prado and Holiday 1996; on Peru see Ciurlizza and Acosta 1997; on Pakistan see Jilani 1998; on Mexico see Reygadas and Soto Martínez 2003); and on Morocco see Naciri et al. 2004.

[3] On the Canadian case, see Kapoor (1996) and the report from an IDRC workshop report (International Development Research Centre 1999: para. 24) on evaluating governance programmes.

[4] Crawford (2003b) also draws on Schmitter's and Brouwer's (1999) discussions on the micro, meso and macro levels of analyses.

[5] These case studies were only used for internal purposes and never published by R&D.

[6] Also see Horton et al. (2003: 20–1), which cites three barriers to using evaluations: the complex and dynamic nature of the development environment; shortcomings in the evaluation process; and internal factors specific to organizations conducting evaluations.

[7] Kapoor (1996: 8) referred to this tendency as an 'anarchy of particularistic viewpoints'.

Annex 6.1: Questionnaire for R&D regional officers in charge of democratic development

(From an internal draft document in the archives at R&D; translated by the author from French into English.)

1. What definition of democratic development do you apply in your regional programme? Is it different from the definition used by Rights & Democracy? If so, why?

2. How would you describe the evolution of the programme in your region in the last ten years? How has the programme changed and why?

3. What are the important achievements of R&D in your region? What were its failures?

4. What are R&D's strengths in democratic development? What are its weaknesses?

5. What would you need to improve your efforts to promote democratic development?

6. What are the major challenges for democratic development in your region for the next five years?

7. Regarding the case studies in your region, please provide:

- A brief history;
- Principle contacts;
- Your immediate perspectives;
- Logistical needs with regards to the in-field assessment.

Annex 6.2: Democratic Development assessment: interview questions (partners and regional experts)

(From an internal draft document in the archives at R&D.)

1. What is the specific context of your knowledge of the democratic development work carried out and supported by R&D (e.g. specific projects, countries, partnerships, etc.)?

2. What is your own definition of democratic development?

3. What are the principal priorities, in your view, for democratic development in the [country/region] that you are involved in?

4. In what way(s) and how well has the work of R&D addressed those priorities?

5. What are the principle shortcomings of the work of R&D in your view?

6. How would you characterize the work of R&D (for example, in comparison with other international democracy support institutions)?

7. What advice would you offer to R&D in its effort to give greater focus to its work within the broad field of democratic development?

8. Any other comments or advice?

Chapter 7

Gauging civil society advocacy: charting pluralist pathways

Chapter 7

Harry Blair

Gauging civil society advocacy: charting pluralist pathways

For several years, all international donors supporting democratization, whether directly (most bilateral donors) or indirectly (mainly the World Bank), have been engaged in backing civil society initiatives. In a democracy context, this means essentially supporting civil society advocacy efforts. But how can donors tell whether such efforts have been successful or not, especially when it comes to particular organizations and constituencies? This chapter concentrates on this question, attempting to develop further a civil society advocacy scale that can help evaluate achievement in civil society advocacy in terms of what benefits accrue to targeted constituencies and the long-term effects of such advocacy in promoting system pluralism.

Introduction

This chapter begins with a brief look at donors' objectives in promoting civil society and at problems in assessing the impact of programming. The second section distinguishes the scale the author has been working on from other, more long-standing, efforts in this area and then goes on to explain the approach. The third section then applies the scale to three well-documented efforts in civil society advocacy—one in India and two in the Philippines. Each has failed in at least some significant way to achieve its ostensible objective but at the same time can be viewed as an exemplary success in building pluralist politics. A fourth section draws lessons from these illustrations, and a concluding section explores the implications for future donor initiatives in promoting civil society.

Civil society, empowerment and advocacy

Although it has long been a contentious term among both scholars and practitioners, 'civil society' can be succinctly defined in the context of this chapter using the Swedish International Development Cooperation Agency (Sida)'s formulation: it is 'An arena, separate from the state, the market and the individual household, in which people organize themselves and act together to promote their common interests' (Swedish International Development Cooperation Agency 2004: 9). 'Empowerment' has been another much-argued term, but hopefully can be rendered into a useful concept for present purposes by accepting Deepa Narayan's definition as '[T]he expansion of assets and capabilities of poor [and I would add marginal] people to participate in, negotiate with, influence, control, and hold accountable institutions that affect their lives' (Narayan 2005: 5). The term 'advocacy' has been much less disputed and it should not be controversial to define it as the process by which individuals, and especially associations (that is, civil society organizations, or CSOs), attempt to influence public policy making and implementation. Thus the advocacy analysed here is state-centred, that is to say, directed at institutions of the state, but *mutatis mutandis* it could be employed to assess advocacy in other contexts as well, for example, efforts within a religious community to promote (or oppose) female clergy, or CSO initiatives to pressure business corporations to change their policies with respect to the environment.

International donors generally put these three concepts together by incorporating civil society in democratization strategies primarily as a means to improve the lives of poor and marginal people by empowering them to advocate for their own interests and by holding state institutions accountable. A smaller group of donors (see e.g. Swedish International Development Cooperation Agency 2004: 14) see building civil society and empowerment also as ends in themselves—promoting participation, enhancing accountability and advancing democratic pluralism. When civil society is used as a means, it is the organizations (and their individual members) that are thought to benefit, whereas when it becomes an end, it is the political system as a whole that supposedly benefits (although presumably individuals will also be better off when the polity has more pluralism). This chapter takes the larger perspective, looking at civil society as the actors, empowerment as their goal, advocacy as their method, and a two-tiered set of beneficiaries—groups (and their members) in the immediate sense (as being rewarded when advocacy succeeds), and the polity in the larger or ultimate sense (as becoming more pluralistic, which means responsive to more citizens, when advocacy succeeds).

When we look at the impact of civil society programmes, then, we must ask what is happening and who is benefiting at two levels—the constituencies (along with the individuals comprising them) and the larger system. Must both levels gain or lose together, or could one profit and the other suffer at the same time? Could a CSO successfully influence the state to adopt policies that benefit its constituency while

also harming (or at best having no effect on) the larger political system? Perhaps an ethnic minority could make a deal with the ruling elite to get included in the system's largesse while freezing out all other aspirants. One thinks of the Habsburg Empire after it admitted Hungarian elites to join the Austrians in the 'dual monarchy' in 1867, leaving all other groups out in the cold. Could a CSO fail to deliver much of substance to its constituency but at the same time enlarge systemic pluralism by enlarging the political space within the polity? Could it in other words lose the battle but contribute materially to winning the war? We will see several examples along these lines below. Civil society advocacy, in short, might work out in different ways.

Gauging CSO impact on the well-being of both constituencies and systems should be of signal interest to donors who are concerned about whether their democracy support programmes are actually doing anything to support democratization and poverty alleviation. It is to this topic that the analysis now turns.

A civil society advocacy scale

In earlier work (Blair 2004), the present writer endeavoured to develop a civil society advocacy scale that could indicate how far a CSO (or coalition of CSOs) is advancing in promoting significant benefits for its constituents and for the polity within which it functions. While it is hoped that this succeeded to some extent in the first objective, the writer is less sanguine that he made much progress towards the second. This present study, then, will try to develop the scale further and test its utility against the experience of three major civil society initiatives.

Several points would be in order here regarding the principal focuses of this chapter before proceeding to explore the advocacy scale. First, the scale is intended to gauge particular advocacy initiatives, whether single CSOs or coalitions (or even—as especially in the Philippines—'coalitions of coalitions' that approach becoming social movements), as opposed to the overall progress of civil society. Accordingly, it is not put forward as an alternative to such instruments as those developed by Civicus (see 'The Civicus Civil Society Index') or the Johns Hopkins project on civil society (<http://www.jhu.edu/~cnp/>), which attempt to measure the overall status of a political system with respect to civil society. Second, the focus is on group advocacy and empowerment rather than individuals (although the latter may well be the beneficiaries of advocacy efforts, as with affirmative action programmes that give preference to members of marginal groups in hiring or education). Finally, the account will concentrate chiefly on how empowerment is (or is not) achieved rather than on how far or to what extent it has been achieved. The focus, in other words, will be mainly (although not exclusively) on the dynamics of advocacy rather than its results.

The scale shown in figure 7.1 has three major components, corresponding to the three core elements that the democracy literature has embraced as its essence—participation (Dahl 1998); accountability (Schmitter and Karl 1991); and contestation

(Schumpeter 1942). Citizens participate as individuals and (for our purposes more importantly) in groups (or CSOs), providing inputs or demands to the political system. CSOs seek accountability from the political system, asking that it respond by modifying its activities (outputs) to comport with their demands. And because there are many CSOs seeking accountability, the level of democratic contestation within the overall political system improves, thus making it more responsive to citizen needs and wants than periodic elections with their blunt and crude policy agendas could ever do. Prior to participation, however, people must become aware of their situation within the political system in a process that can be labelled 'social capital accumulation' (see figure 7.1). The figure, then, in effect has three-and-a-half components.

Figure 7.1: The civil society advocacy scale: a logical chain
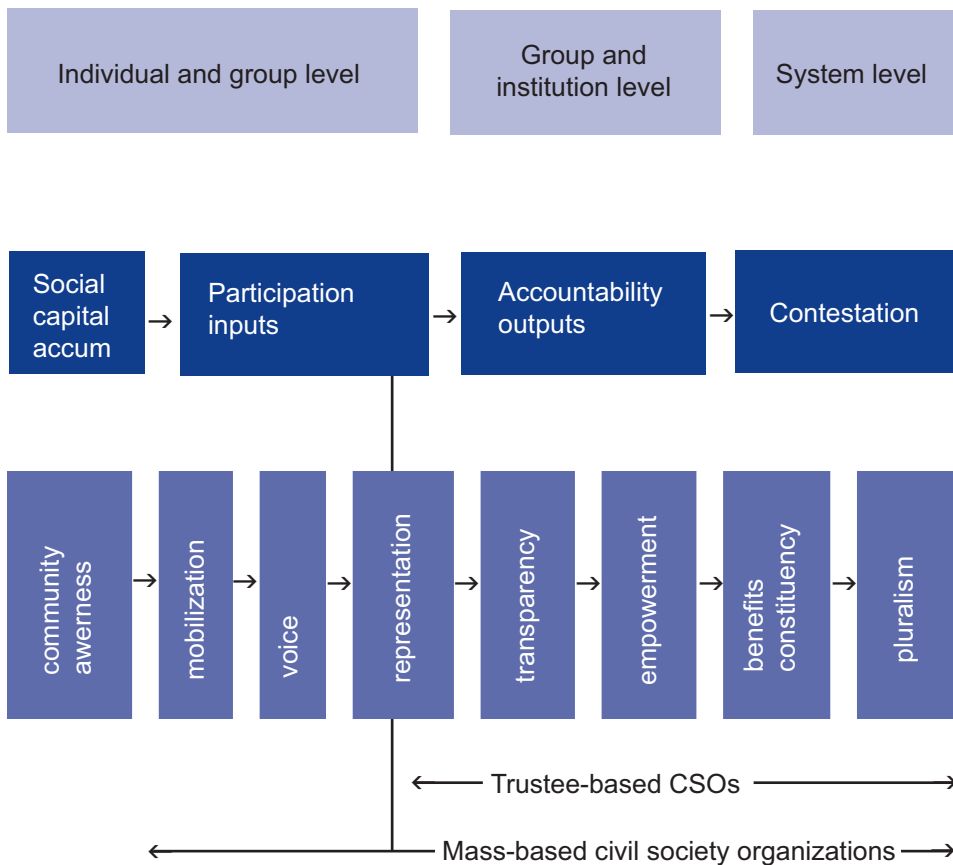
Source: Based on a figure published in Blair, Harry, 'Assessing Civil Society Impact for Democracy Programmes: Using an Advocacy Scale in Indonesia and the Philippines', *Democratization,* 11/1 (2004), pp. 77–103.

The civil society organizations that operate along the advocacy scale may be divided into two basic types. Mass-based CSOs traverse the entire scale shown in figure 7.1, beginning by promoting (or more likely harvesting) community awareness, and then moving into the participation stage, by organizing people for political participation (mobilization), developing agendas to bring into the public policy discourse (voice) and bringing their constituencies' demands to the attention of public authorities (representation). The accountability stage opens when a CSO has enough credible representation to compel the state to justify its actions (the beginning of transparency). Empowerment comes when the state finds it must meet at least some of a constituency's demands by modifying public policy decisions, but only when those decisions are put into action do constituency benefits occur. If enough CSOs representing enough constituencies get into the game in a serious way, then finally we can say that the level of pluralism has increased.

Mass-based CSOs can be further subdivided into membership organizations such as labour unions, professional associations, and constituency-based groups such as neighbourhood slum dwellers, petty traders, ethnic minority groups and so on, where active 'membership' is much more flexible and fluid. (For more on these distinctions, see Ottaway 2000 and the essays in Eade 2002.)

The second type of CSO can be called trustee-based, in that organizations operate on behalf of constituencies that cannot act for themselves (see Ottaway 2000). Human rights CSOs provide an excellent example, generally consisting of small cadres of (often foreign) elites who investigate abuses, publicize findings and pester governments on behalf of people who are unable to act on their own behalf, such as political prisoners or lower-status women. Environmental activist CSOs try to advance the cause of a constituency that is for the most part inherently inarticulate (indigenous inhabitants can be mobilized as a mass-based constituency to defend their environment, but then the CSOs representing them would fall into our first type). This second type of CSO in effect bypasses the social capital accumulation and participation stages of advocacy to concentrate on the accountability stage. The main focus here will be on the first or mass-based type, but there will also be evidence of the trustee type of CSO in two of the cases presented below.

The scale illustrated

An imaginary (admittedly ideal) example will show how the process illustrated in figure 7.2 might work. Village mothers talk about the school their children are attending, deploring the collapsing buildings, the lack of basic supplies such as textbooks and the common absence of teachers themselves (who supplement their meagre incomes by tutoring pupils for a fee rather than attending their classes); community awareness is building. A group of mothers, perhaps inspired by a story one has seen dramatized on the community television set, get together more frequently to vent their grievances (mobilization). Some start making a list of things that ought to be done (voice).

A group of several dozen mothers organizes itself to demand an audience with the elected village council, which, after initially brushing them off, begins to think of the election coming up in six months' time and decides it really should meet them (representation). A new constituency has begun to participate in the local political arena.

Figure 7.2: The civil society advocacy scale: an imaginary case
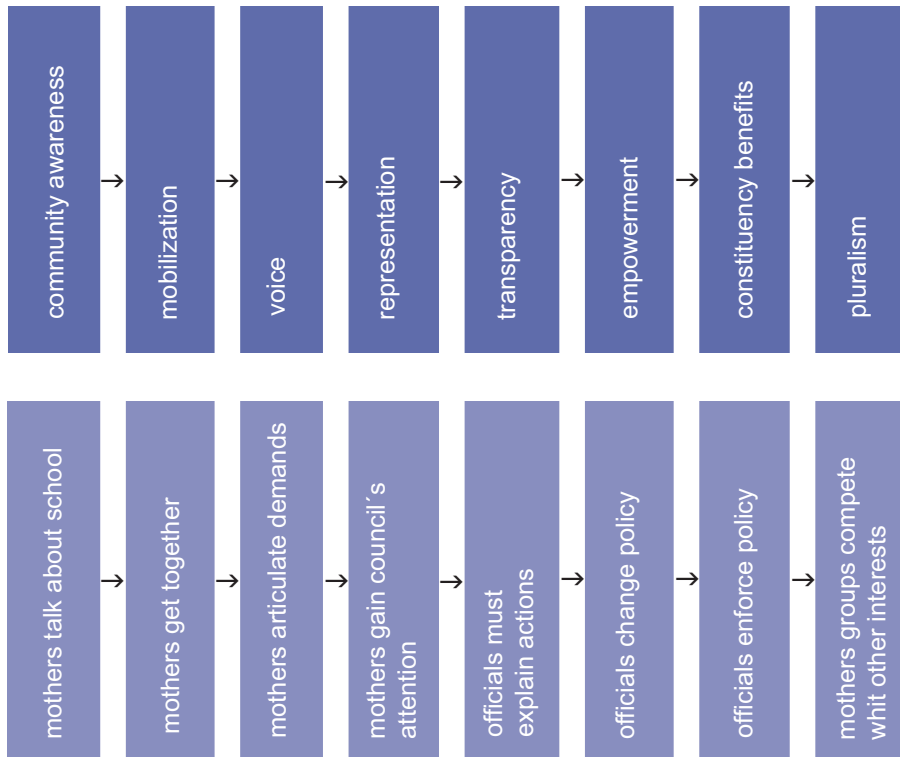
In the course of several meetings with the mothers' group, the council finds itself pressed to explain why it has done nothing to insist that the district education office repair the school roof or demand that the teachers show up for duty (the start of transparency). Exploiting kinship networks, the mothers' group links up with dissatisfied parents in neighbouring villages and the group becomes larger. Several mothers find some satisfaction in their advocacy work and make representations on behalf of their now much larger constituency to the district (that is, higher-level) council. These council members, now contemplating their own re-election chances, formulate a directive demanding that teachers attend their classes (empowerment for the mothers' organization), although nothing is done to enforce the new order. A couple of the mothers have husbands who work for the district newspaper, and who interest its manager in doing an investigative piece exposing the fecklessness of

the teachers and the indolence of the council (more transparency). With the election looming, an embarrassed district council follows up on its directive to the teachers, sacking several egregious absentees and inducing the remainder to begin taking their jobs seriously. At the same time, it decides to divert some of the Education Ministry's funding that it had devoted completely to patronage efforts back into repairing the school roof and buying textbooks. Teachers start actually teaching, the roofs are repaired, books are distributed, and pupils begin learning (constituency benefits). The political system has become accountable to a significant constituency among its citizenry.

The newly empowered mothers are not the only constituency to get involved in politics, however. The schoolteachers' union, heretofore largely somnolent as its members enjoyed the perks they enjoyed with their no-show jobs, stirs itself into action, demanding pay rises for its newly hard-working constituency. Local contractors sense that there are business opportunities in school repair and reconstruction and begin lobbying for increased funding to upgrade the educational infrastructure. The district council, now being pressed on different sides by a growing chorus of demands, finds that it must balance resources against them, seeking the best possible calculus to respond to the public. The level of pluralism, in short, has increased. This final stage will not necessarily work to the benefit of those who launched the process, it must be noted. The aroused teachers' union may roll back the district council's new demands on their services, while the contractors may contrive some way to siphon off virtually all the construction money into a combination of graft for themselves and pay-offs to the politicians letting the contracts.

What has been achieved in this imaginary example? Civil society advocacy has served as a mechanism to produce concrete benefits for children and their parents, and at the same time the advocacy experience has increased the capacity of local people to manage their own affairs. Advocacy has been both means and end. And even if the mothers don't get to the actual benefits level, or if teachers and contractors erode any improvements, they will have learned valuable lessons about political activism, which they can use to fight another day, perhaps on another political battlefront. Next year they might launch a campaign for improved drinking water or an electricity supply. All these things could be counted as achievements.

Three case studies

To illustrate the civil society advocacy scale with some real examples, three cases are presented below. All have been extensively documented and analysed elsewhere. Two of them have unfolded over more than a score of years, while the other, although it did not go on so long, attracted immense attention worldwide while it was in progress. All three cases include at least some elements of all parts of the advocacy ladder. The first is set in India, while the other two took place in the Philippines. Both countries can fairly be described as democracies, although they underwent a period

of authoritarian rule in the 1970s, and even today democratization has not been fully attained. In India, democracy remains under threat from the fundamentalist Hindu right and from destabilizing violence in Kashmir, the north-east and Maoist rebels in the Gangetic Plain. Several significant elements (including military factions and the Mindanao rebels) in the Philippines remain unconvinced that democracy is 'the only game in town' and periodically seek to overthrow the democracy. Both countries therefore continue in transition towards democracy and can thus provide good examples and insights for the democracy promotion community.

It should be noted that all these cases were essentially home-grown; none stemmed primarily from donor-supported efforts. This qualifier could be seen as limiting their suitability for an exercise like the present one, aimed as it is at providing guidance for future donor strategies. So why choose them? Principally because there are no instances of donor-sponsored civil society efforts that are anywhere near as rich in terms of analysis available from such a wide variety of sources and that cover so well the entire range of the advocacy scale presented in the section above. Some idea of the effectiveness of advocacy can be gleaned from donors' experience with specific CSOs (see e.g. Blair 2004), but generally the focus among donors has been on measuring outputs or indicators rather than overall impact. In addition, donor support for CSOs has tended to be for relatively short periods of a year or two—scarcely long enough in most cases to generate the effects to be examined here. If we are to attempt to gauge the impact of advocacy, we have to go to where the evidence lies. Hopefully the lessons to be found will be useful in informing donors' thinking about their own programmes.

The Narmada Dam

The origins of the Narmada Dam controversy lie back in the 1940s[1], when irrigation engineers began to study the potential of the Narmada River in western India to supply water to the arid regions of what later became the state of Gujarat. The centrepiece of this multi-dam project was to be the final dam at Sardar Sarovar, rising to some 138 metres (m) when finished. It was projected to irrigate 1.8 million hectares of agricultural land, supply potable water to 30 million people, and furnish 2,700 megawatts (MW) of hydropower.

The Narmada flows through three Indian states (Madhya Pradesh and Maharashtra as well as Gujarat), and prolonged arbitration was necessary before work could begin, finally, at the end of the 1970s. A unique feature of the settlement regarded persons displaced by the project, or 'oustees', most of whom were adivasis (tribal people, ethnically and culturally distinct from the majority population). For the first time in an Indian dam project, oustees were to be given land elsewhere, at least equivalent to what they would lose through submergence in the Narmada project. This was the 'resettlement and rehabilitation' (R&R) programme, a significant improvement on earlier practice of providing at most a small cash payment for land seized by dam

projects. Most of the funding was to be provided by the Indian Government, but the World Bank agreed in 1985 to step in with a 450 million US dollar (USD) 'start-up' loan (the Bank anticipated further loans later on) that would cover about one-seventh of the Sardar Sarovar's then-estimated total cost. Altogether the cost of the complete project was then estimated at some 15 billion USD.

Not surprisingly, as construction geared up, many potential oustees objected, and some formed organizations to oppose or modify the project, setting in motion the process depicted in figure 7.3. By the early 1980s, in addition to various CSOs in the immediate region of the dam project, others in New Delhi had become interested in it, and even some international CSOs such as Oxfam in the United Kingdom and the Environmental Defense Fund in the USA became engaged, lobbying the World Bank, which responded by commissioning the first of several studies on the project. By the mid-1980s, marches and large-scale demonstrations had been mounted, and court cases were filed. The first charismatic leader emerged in the person of Medha Patkar, a social worker turned activist in the cause of the adivasis, and a well-known Gandhian leader, Baba Amte, joined in as well. CSO demands at this point focused on R&R for the displacees, although some had begun to question the wisdom of the Narmada project altogether. By the end of 1987, the Gujarat state government announced better R&R terms.

Figure 7.3: The civil society advocacy scale: the Narmada Dam

community awareness → mobilization → voice → representation → transparency → empowerment → benefits constituency → pluralism

Initial oustees (late 1970s) → Early CSOs formed (1979) → R&R demands articulated (1980s) Demonstrations, marches Worldwide support split between R&R and oppositionists (1990s) → GOI negotiates with NBA (1993) → World Bank Mores Report (1992) NBA legal actions (1990s) Five Member Group (1993) → MEF as ally (1980s onward) Gujarat improves R&R terms (1987) GOI drops World Bank loan (1993) Supreme Ct hears cases (1995, 1999) → Supreme Court stay (1995) Supr Ct rescinds stay (2001) Supr Ct renews construction halt (2005) → Contending state govts, pro-Narmada CSOs

At this point, what was becoming a movement began to split between one element which focused on getting the best R&R terms for the oustees and a second faction which took up a position of total opposition to the Narmada project. Those in the

first camp worked mainly in Gujarat, cooperating with the state on a tactical basis, assisted by an ally at national level in the form of the Ministry of Environment and Forests in demanding compliance with government R&R regulations. On the other side, opposers coalesced around a new umbrella coalition formed in 1989 and calling itself the Narmada Bachao Andolan (NBA, meaning Save Narmada Movement).

Rather than devote attention to improving the lot of the oustees, whose numbers it estimated to range upwards of 1 million people (although others, e.g. Gandhi 2003: 484 and Gupta 2001: 75, put the numbers much lower), the NBA focused on the anticipated ecological degradation, the concentration of benefits expected to flow mainly to the rich, and the high cost overruns deemed inevitable. It intensified the anti-dam campaign both within India and on the world stage. Large rallies, demonstrations, road blockages, and a 6,000-person march covering 200 km on foot were among the tactics employed at home. Abroad, CSO pressure led to several US congressional hearings and a new review commission funded by the World Bank. Released in June 1992, the Bank-sponsored report found serious flaws with the Narmada project's R&R provisions, calling for the Bank to 'step back' from the project and take a fresh look at it. The next month the European Parliament passed a resolution calling on member countries to tell their representatives at the World Bank to cancel the project.

Stung by the severe criticism, the World Bank did indeed step back, insisting that the Indian states improve their R&R offers and laying down a six-month deadline. Embarrassed by the bad publicity, as well as annoyed by what it deemed foreign interference (but at the same time knowing that World Bank support only amounted to a small part of the total Narmada project cost), the Indian Government announced in March 1993 that it would terminate the Bank contract.

Despite the setback, the Indian Government and the state governments pressed on with Narmada, while the NBA continued its campaign at both national and international levels. Within the state of Gujarat, the Narmada canal system became a mantra fervently grasped by all political parties—'the lifeline' for this largely semi-arid state (on the 'lifeline' theme as a powerful force in Gujarat politics, see D'Souza 2002: passim), and politicians of all persuasions pushed it relentlessly at both state and national levels. For its part, the NBA intensified its efforts with a hunger strike in Bombay, followed by a dramatic 'death by drowning' campaign in 1993, featuring 'drowning squads' pledging to stay with their homes even as rising dam waters submerged them. Internationally, the NBA worked in concert with an international Narmada Action Campaign involving CSOs from 15 countries, inter alia sponsoring a full-page advertisement in *the New York Times* in September 1993. An advertisement headlined 'Why Thousands of People Will Drown Before Accepting the Sardar Sarovar Dam' appeared in The Times on 21 September 1992 (Wood 1993: n. 18). The Indian Government felt obliged to set up a Five-Member Group of prominent citizens to forge a compromise, while for its part the NBA filed a writ petition before India's Supreme Court seeking to halt construction as technically, environmentally,

economically and socially 'not in the national interest'. In January 1995, the Supreme Court halted construction of the Sardar Sarovar dam at 80 m, or just less than 60 per cent of its planned height of 138 m.

While the Supreme Court was deliberating the case, the NBA kept on with its campaign regionally, nationally and internationally. A major coup for the NBA came with the recruitment of Arundhati Roy, the prize-winning author of *The God of Small Things,* who, beginning with her essay 'The Greater Common Good' in the spring of 1999 (Roy 1999)[2], became a major activist and publicist for the cause, marching with the demonstrators and proclaiming that the anti-dam movement had replaced writing fiction as her life's main focus.

Despite her presence and support for the NBA (or, as some would later argue, in part because of it), the Supreme Court found against the opponents of the dam in October 2000, deciding that the Sardar Sarovar dam should be completed according to the design laid out in 1979 (Routledge 2003: 253). The Court's decision provoked a firestorm of outrage from the NBA and its allies, fuelled by intense media coverage. But construction resumed and, despite periodic flare-ups—in December 2002, for instance, a protest demonstration resulted in Arundhati's arrest and eventual sentencing by the Supreme Court to one day in jail for contempt (see Yates 2002; and Harding 2002)—the movement gradually lost steam, as is clear from the data shown in figure 7.4, which indicate press clippings about the Narmada declining from a high of over 200 the month after the court's October 2000 decision to near zero a couple of years later. This remarkable collection of clippings, mainly from Indian newspapers but also from the international media, was maintained by the Friends of River Narmada, and most of the items are still available online on their website (<http://www.narmada.org>).

Figure 7.4: The Narmada Dam: monthly clippings, 1999–2004
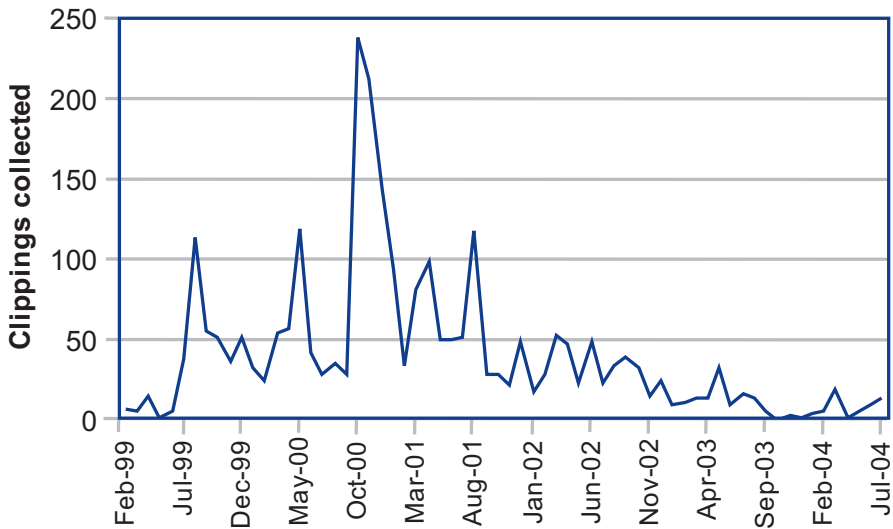
Source: Compiled by the author from the listed of clippings on the Friends of River Narmada, <http://www.narmada.org>.

The struggle is not yet over, however. In March 2005, the Supreme Court again intervened, stating that the R&R orders included in its October 2000 decision to resume construction had been seriously violated. The Sardar Sarovar dam was halted at 100.6 m, still far short of the intended height. Presumably, however, the dam authorities will meet the R&R requirements and construction will in time resume. In early 2006, Wood's prediction of more than ten years still looks good—that 'the [state and national] governments and developers will eventually have their way… the unfolding story of the Narmada controversy indicates that too much is at stake—economically, legally and politically—for current development plans to be reversed' (Wood 1993: 969). The Narmada controversy has been (and still is) a convoluted one, with a good deal of movement back and forth along the advocacy scale shown in figure 7.3. The same path was seemingly traversed more than once, but there does seem to have been a logical progression from awareness up through community benefits (and their apparent withdrawal) and finally an increased degree of pluralism in the political system.

Ousting a president in the Philippines

When the Supreme Court of the Philippines approved the removal of President Joseph Estrada from office in January 2001, a long saga capped by mammoth public demonstrations finally ended. The immediate reason behind his departure from the

Malacañan (presidential) Palace lay in the armed forces' abrupt withdrawal of support (a move quickly approved by the Supreme Court), but that was only the last act. The campaign against him had been building for months, built largely on investigative reporting carried out by the Philippine Center for Investigative Journalism (PCIJ), which had first launched its research effort more than a year previously.[3] Taken as a whole, the case provides an excellent illustration of how public policy, if it is to represent the popular will, requires accountability, transparency and the involvement of long-term constituencies. It also shows the messiness, inconclusiveness and mixed consequences of what might seem at first glance to be a clean and clear-cut denouement of a lengthy drama, as illustrated in figure 7.5.

Figure 7.5: The civil society advocacy scale: the ousting of President Estrada
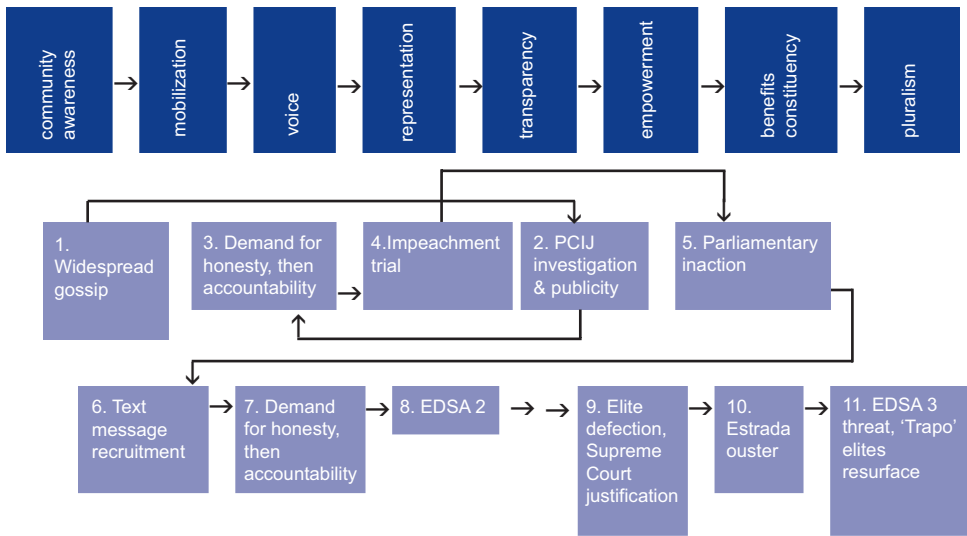
Our case began with wide community awareness, as rumours of presidential corruption swirled around President Estrada, beginning almost immediately after his election in 1998. A former film star, Estrada had long been known for his lavish spending on his wives and mistresses, as well as keeping shady company in his business dealings, and there was much apparently well-founded talk of his continuing to do so after becoming president. Gossip and political jokes did not lead directly to any mobilization of outraged citizens, however, to say nothing of voice and representation on the advocacy scale. Instead, it was the PCIJ that decided to investigate the rumours, beginning in January 2000 by researching corporate registration and financial records at the national Securities and Exchange Commission, which allowed individual citizens to retrieve three records each business day. In terms of the advocacy scale shown in figure

7.5, then, the movement was from community awareness directly to transparency. The PCIJ amounted to a 'trustee-based CSO' operating on a public-interest basis on behalf of a constituency (in this case a potential one) that as yet had not coalesced into an active group.

In July the PCIJ was ready to release its first set of stories linking Estrada's wealth to a set of some 66 corporations that he had failed to list in the annual declaration of assets required of all public office-holders in the Philippines. The mainstream press, largely cowed by presidential bluster, intimidation and reportedly some bribery, showed little interest in the stories, but they were picked up by several small papers and created a modest stir. The following month, the PCIJ released a second set of stories, also compiled through painstaking research this time into public real-estate records, detailing a string of luxurious mansions built for Estrada's wives and mistresses since he took office which had been carefully hidden from public view. This time larger papers picked up the story, and the PCIJ succeeded in getting it aired on a television network (despite strong objections from the network's owners).

By now, the investigation had developed into a public scandal. Interest and indignation built up. No formally structured CSOs of any size emerged to engage and develop a constituency of indignant citizens, but public demands for some response to the PCIJ's revelations began to build, and then another front against the president opened with claims by a provincial governor that that he had been paying off Estrada in connection with a massive illegal gambling scheme. By November the national House of Representatives had filed formal impeachment articles against the president, with three of the four principal charges based on the PCIJ newspaper stories. Further incriminating articles from the PCIJ strengthened the case as it unfolded. Huge anti-Estrada crowds demonstrated against the president, mobilized into action through what was probably the world's first people's movement activated through mobile phones and text messaging.[4]

The impeachment proceedings stalled in the Senate on an ostensible technicality in mid-January, as Estrada loyalists engineered a resolution by a one-vote majority not to accept evidence linking the president to bogus bank accounts. By this time, however, the affair had gained such momentum that millions of citizens were watching the Senate proceedings, and after the vote popular indignation erupted in mass protests. Mammoth rallies—quickly dubbed 'EDSA 2' after the EDSA demonstrations that were instrumental in ousting President Ferdinand Marcos in 1986—resumed, again facilitated through mobile phone text messaging. (EDSA refers to Epifanio de los Santos Avenue, Manila's ring road where the demonstrations were held, that led to the ousting of presidents Marcos and Estrada.) In short order, the vice-president, major Cabinet members and important members of the president's party resigned, and prominent leaders along with the Roman Catholic prelate Cardinal Jaime Sin demanded Estrada's resignation. The final step came when military leaders withdrew their support, on 19 January 2001. The Supreme Court then ruled that the president should be stripped of office and that Gloria Macapagal Arroyo (the newly resigned vice-president) should be sworn in.

The coco levy case in the Philippines

Our last case has been in play even longer than the Narmada controversy and is illustrated in figure 7.6. Coconuts form the basis of a major sector of the Philippine economy, accounting for some 30 per cent of the country's export earnings in the early 1990s and providing income for as much as one-third of the country's population. The coconut levy saga[5] goes back more than 30 years, beginning in the early years of the Ferdinand Marcos' dictatorship, when he established successive 'levies' on the sale of coconuts to the millers who produce coconut oil. Ostensibly intended to support funds for price stabilization, the levies collected soon went into the hands of Marcos cronies charged with managing them. Despite the dictatorship, the levies were onerous enough for a great deal of farmer opposition to develop, and finally they were ended in 1982.

Figure 7.6: The civil society advocacy scale: the coco levy
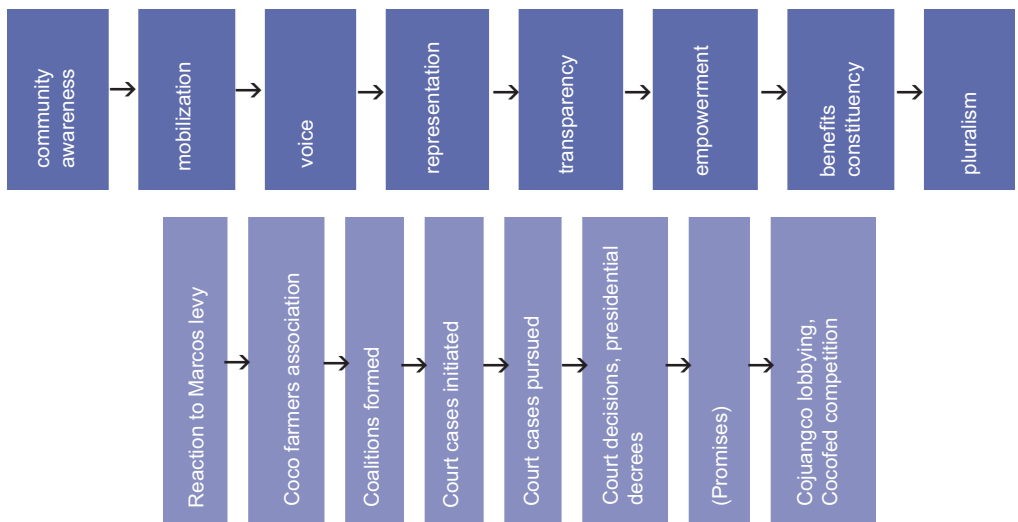
By this time, however, the larger of the two levies had collected almost 10 billion Philippine pesos (PHP), and much of the money had been invested through various mechanisms controlled by cronies into a number of industries, including the San Miguel Corporation, by far the largest brewery in the country. The lead Marcos crony involved in the coco fund at the time was Eduardo 'Danding' Cojuangco, who along with the Marcos family fled the country after the EDSA revolution in 1986. The following year, the Presidential Commission on Good Governance, tasked with recovering the illegal gains siphoned off during the Marcos regime, filed suit with the newly established Sandiganbayan (anti-corruption court). But Cojuangco eventually resurfaced, returning in 1991, and resumed his role in San Miguel,

claiming control over some 47 per cent of the corporation's stock. At the end of the decade, the prize had become a huge one indeed, worth about 1 billion USD; San Miguel was ranked third among all Philippine corporate enterprises in the year 2000 by *Asiaweek,* and one of only seven operations in the country that ranked among the top 1,000 in Asia overall (see <http://www.cnn.com/ASIANOW/asiaweek/features/asiaweek1000.2000/c_philippines.html>).

Many of the farmers who had been subjected to the levy continued to seek its recovery, and a number of CSOs were pursuing this objective. By the mid-1990s, two coalitions representing small coconut farmers' associations were active on this front, the Coconut Industry Reform Movement (COIR) and the Pambansang Koalisyon ng Magsasaka at Manggagawa sa Niyugan (PKSMMN, a coalition of NGOs representing small coconut farmers).[6] Over the course of the 1990s, the two umbrella groups received funding from various donors, including the United States Agency for International Development (USAID), which sponsored an initiative it called Building Unity for Continuing Coconut Industry Reform (BUCO), intended to assist the two coalitions.

BUCO's main efforts were devoted to advocacy through the government, the political arena and the media to obtain a release of the San Miguel shares to a trust fund that would benefit the small farmers who had involuntarily created the stockholding through the levies. BUCO's argument was that a majority of Cojuangco's 47 per cent holding rightfully belonged to these farmers. The alliance pulled together a Multisectoral Task Force (MTF) including businessmen, religious leaders, academics, legislators and even former Cabinet secretaries. Their main objective was to induce then President Joseph Estrada to issue an executive order setting up a trust fund for these shares that would benefit the coconut farmers.

In its campaign, the task force generated immense publicity, with video documentaries and heavy press coverage, aided greatly by the convenient target presented by Cojuangco with his past as a leading Marcos crony and widespread current allegations that he had a similar role in Estrada's inner circle. Press stories appeared frequently in Manila's leading daily newspapers, often on the front page. (BUCO assembled a collection of 64 clippings appearing in the national press between January and July 1998. See Building Unity for Continuing Coconut Industry Reform 1998. Between May and August 2000, more than a dozen stories appeared in just two Manila dailies, the *Philippine Daily Inquirer* and *Philippine Star.*) BUCO also pushed a legal case in the Sandiganbayan.

The BUCO coalition was not the only group attempting to pressure the president. Cojuangco himself insisted that all the shares were rightfully his personal property, and not purchased with coco levy money. And BUCO faced a rival claim from another coconut group, the Coconut Producers Federation of the Philippines (COCOFED), which claimed that they, representing larger farmers and mill owners (who actually paid the levy collections to the government), were entitled to the contested shares (BUCO responded that the mill owners simply deducted the levies from what they

paid the growers and thus acted merely as conduits in the process, not as the actual payers). The fact that COCOFED elected a representative to Congress under the 'party list' electoral system in 1998 gave it added clout in pressing its claims.

President Estrada repeatedly promised to issue an executive order resolving the issue, but, pressed on different sides by BUCO, COCOFED and Cojuangco, he kept postponing a decision. Then in November 2000, as the impeachment movement against him appeared to be gathering steam (see the above section), he issued a Solomonic decision, awarding part of the prize to Cojuangco and part to be auctioned, with benefits going to BUCO's two component coalitions and also to COCOFED. The move created great confusion and anxiety, but the events surrounding Estrada's removal from office the following January overtook everything else, and the situation remained unresolved.

After Estrada's departure, all sides took up their separate causes again, with Cojuangco and the various CSOs lobbying the new president, pursuing the case in the Sandinganbayan, and trying to marshal public opinion in their favour. The case continued to excite much public interest. The Internet archive of the *Philippine Daily Inquirer,* for instance, indicated almost 200 articles on the coco levy during 2001 (see <http://www.inq7.net>). In July 2003, the Sandinganbayan ruled that the disputed San Miguel shares belonged to the government, which was helpful to the farmers, but scarcely disposed in their favour (Pazzibugan et al. 2003). In early 2006, the dispute was still in full play. The PKSMMN coalition, which had earlier denounced its partners for trying to sell out the small coconut farmers, asserted that its members approved a negotiated settlement with Eduardo Cojuangco, while the MTF, formerly an umbrella alliance that included the PKSMMN, now said it would gather at least half a million signatures to prove that its members disapproved of the settlement. The MTF charged that President Arroyo and Cojuangco had made a sleazy deal to deprive the farmers of their rightful due (Calumpita 2006). In short, more than 30 years after the coco levy had begun and more than 20 years after it had ended, the principals involved were still wrangling over its ownership, with no end to the dispute in sight.

Lessons to be drawn

At first glance, our three case studies might appear to have few lessons to offer donors. As noted above, neither the Narmada campaign nor the anti-Estrada movement was initiated—or even supported in any serious way—by official international donor agencies. Nor was the coco levy movement, except for some modest USAID assistance when BUCO was formed in the late 1990s. And yet such cases do provide excellent examples for donor learning.

After 'management for results' became the mantra for evaluating programme assistance in the 1990s, donors became fixated on quick measures and indicators of success, rather than in-depth analyses of what worked and how. Moreover, donor assistance to civil society has largely been delivered in short-term grants, which has

obviated any real interest in looking at longer-term impact. Not surprisingly, there has consequently been little donor interest in analysing in depth cases like those presented here. But it is precisely by studying such cases that we can build a picture of how advocacy efforts function to influence public policy over both the shorter and the longer term. By building an understanding independent of donor programming of how civil society groups succeed or fail in attaining empowerment, and of the impact of their experiences on systemic pluralism, we can find many useful lessons to inform future donor strategies.

Success

What is 'success'? The concepts of 'success' and 'failure' are especially elusive in the Narmada case, in part because the identity of the principal constituency itself has been in dispute since early on in the controversy: was it the oustees, whose homes and livelihoods were to be wiped out by the project? Or was it the environment of the Namada River Basin, which would be forever changed by the dams? Yet again, perhaps it was the wider cause of environmentalism in India. If it was the oustees, then perhaps the campaign has succeeded in that they will eventually get a better R&R settlement than could ever have been dreamed possible a couple of decades ago. Even illegal squatters were to receive some compensation for being moved by the dam project.

But if the river system's ecology was to be the principal beneficiary, then perhaps the NBA has failed, because the dam has been in good part built and, once the current turbulence occasioned by the Supreme Court's 2005 decision has been smoothed out, construction will doubtless continue. And, finally, if it is the Indian environment in an overall sense that was to be the main beneficiary, the Narmada controversy may well turn out to have been a tide-turning event which, even if the immediate battle was lost, energized the cause of environmentalism in such a way that the long-term effort to preserve the environment has been greatly enhanced. Among environmentalists and long-time champions of the poor there has been much argument and dissent. Many supported Arundhati Roy and the NBA, while others accused them of romanticism at the expense of the adivasi oustees who could have achieved an even better deal than they were in line to receive if the NBA had concerned itself more with real people and less with opposing the idea of development at all costs—see, for instance, the intense irritation at Arunadhati Roy expressed by prominent environmentalist Ramachandra Guha (2000a, 2000b) and the pro-poor activist and intellectual Gail Omvedt (1999a, 1999b).

In the Estrada case the constituency ostensibly succeeded in its goal of ousting a corrupt president, but was this truly a victory for Philippine democracy? Some would argue this, including a great many Filipinos at the time, and the country's leading opinion polling organization found a majority in support of the president's departure (Mangahas 2001; Reid 2001). However, many of the same cronies soon resurfaced

around the new president, and the same old crowd of 'trapos' (traditional politicians) quickly resumed their leading roles in the system. The country's dominant oligarchy (of which the new president's family were virtually charter members) had little trouble in coming back to where the power was.

Perhaps worse, the EDSA approach to politics showed fair promise of becoming a habit, as Estrada supporters mounted an attempted—although unsuccessful—'EDSA 3' to restore their man after his removal, mustering their own gigantic demonstrations. Five years later, President Gloria Macapagal Arroyo faced a new impeachment trial on grounds of corruption, along with Cabinet resignations and an attempted (albeit failed) coup within the military. A repetition of the same destabilizing movement scenario—including an 'EDSA 4'—cannot be dismissed out of hand. After an impeachment drive failed in late 2005, an attempted 'people power' effort combined with a military coup was launched in February 2006. The movement fizzled out fairly quickly but the extra-constitutional impulse seems nevertheless to be running strong in the Philippines (see Mydans 2006a, 2006b; and Gomez 2006). As for EDSA 2 and the ousting of Estrada, it is arguable whether the PCIJ and the movement it initiated really benefited its constituency or the political system.

With the coco levy saga, it would be hard to argue that the past three decades of advocacy—first against the levy itself and then to recover the funds that were levied—have brought any concrete success at all to the farmers involved. Indeed, the prospect of obtaining any benefit from the coco fund must seem more like an ever-elusive chimera to the hundreds of thousands of coconut growers than a cause with some hope of succeeding. Presidential promises and decrees, as well as court rulings, have come and gone, but in 2006 Eduardo 'Danding' Cojuangco continued to control the same 47 per cent of San Miguel shares that he had at the end of the Marcos dictatorship in 1986.

Achievement

Perhaps it makes more sense to think about democratization 'achievement' rather than immediate campaign 'success' as the gauge of civil society advocacy impact. In the end, perhaps it is experience at 'doing democracy' that is really important. Thus NBA marchers and 'suicide squad' members became more effective participants in the practice of democracy, as did PCIJ journalists (and their readers) and EDSA demonstrators, and small coconut farmer-members of the COIR or the PKSMMN. The skills these participants developed, whether as advocacy leaders or as foot soldiers, will prepare them to take on other causes and eventually become more successful players in the democracy arena.

Are some levels of achievement (that is, steps on the advocacy scale) more valuable than others? Certainly, people will become discouraged if they get no further than, say, the Voice stage after repeated efforts. And the longer they stay at that level, the more dedicated and charismatic the advocacy leadership will have to become if they

are to induce group members to carry on the struggle. One can think of the US civil rights movement and its organizations, which took roughly a century after the American Civil War of the 1860s to achieve much in the way of concrete benefits for their constituency. Women's suffrage in the Western countries represents another long-term saga, where it took decades to achieve any real progress. And, within a shorter time frame, the gay rights movement traced a similar path. (At least in the USA and the United Kingdom, women's suffrage organizations collectively took from the 1880s to the 1920s to achieve the vote for their constituency, while formal CSOs campaigning for gay rights took perhaps a couple of decades, from the 1970s to the 1990s, to begin attaining legal rights in the form of anti-discrimination laws.) But without those lesser stages having been achieved it is surely fair to say that the later and more concrete attainments never could have been realized. So, even if the accountability stages are what count in the end, the participatory stages are critical building blocks. This perhaps will not be pleasing to donors who want to think in terms of three- and five-year democratization programmes, but it accords well with experience in the developed countries.

The impermanence of success

Whatever successes CSOs attain can always be reversed. The NBA obtained a stay from the Supreme Court in 1995, only to have it rescinded in 2000. The anti-Estrada movement gathered enough support to get the Congress to take up an impeachment case in the autumn of 2000, only to have it rejected early in 2001. The coconut farmers have several times appeared to be on the verge of winning a settlement, but each time they have been denied it in the end. The movement against Estrada did win the battle to remove him from office, of course, but if its ambition was to eliminate cronyism and corruption, it lost the overall campaign, as both elements resurfaced almost immediately in the succeeding Arroyo government. Even seemingly permanent victories can come undone, as the administration of President George W. Bush in Washington showed on the environmental front by rolling back such landmarks as the Kyoto Protocol and much of the Clean Air Act and Clean Water Act that previous administrations had welcomed. Few things in politics are ever immutable (although some things, like a 120-m-high dam, would be difficult to undo). But, just as achievements can be overturned, so can losses: the environmentalists, anti-corruption campaigners or coconut farmers may well win in future rounds.

A logical/ordinal scale, not a chronological one

It should be clear from our case studies that progress along the advocacy scale was not at all necessarily chronological. While the coco levy example did proceed along the scale, the other two definitely did not, backing and filling, at times moving a step or more back before going forward again. The Narmada example tracked back and forth

several times over its course, while the Estrada case began essentially in the middle of the scale as a 'trustee-based' effort. For the PCIJ took up its investigation originally on its own internal initiative; there was no real constituency insisting on action or even mobilized. Only when the PCIJ had made a case against the president was it possible to develop a constituency, and it was other groups that mobilized interested citizens to demand accountability, not the PCIJ. Still, it seems evident that a campaign does not begin to move seriously into the empowerment or constituency benefits stage unless all the logically prior steps have been taken. The Gujarat state government would not have improved the R&R package, nor would the final steps of Estrada's removal from office have taken place, nor would presidents and courts have decided (at least temporarily) in favour of the coconut farmers, if all those logically prior elements had not been in place.

Assessing advocacy

How can achievement along the advocacy scale be assessed? It would be wonderful to develop a set of metrics for gauging such progress. In our imaginary case, for example, mobilization might be measured by what proportion of pupils' mothers got involved in the group's initial days. Transparency could be assessed by asking how far school officials found themselves having to go to explain themselves (a letter from the principal? a meeting with the school board? a court case requiring disclosure of official records?). But, even if such metrics could be crafted, they would be useful only for school cases, and probably only for certain schools in a particular country. Donors spent much time and effort in developing measurement systems during what might be called the 'evaluation decade' of the 1990s, including the democracy sector, but with somewhat dubious results (see Blair 2000, 2002). It is most doubtful that they could construct better schemes for dealing with the advocacy scale.

A far better approach would be to undertake a thorough analysis of specific advocacy efforts, using intensive interviews with participants and officials involved, perhaps along with some surveys—or, given the expense of good surveys, 'good enough' focus groups could suffice—with 'thick description' being the principal technique. The three case studies presented here provide a good deal of useful material. But they come from a variety of sources who were concerned with asking questions different from those posed here, so that the writer's own analysis has had to be done very much at third hand. The kinds of query a solid assessment of civil society advocacy would ask would run along the following lines, paralleling the advocacy scale:

• When did any appreciable number of people become concerned?
• What exactly did the initial mobilizers do?
• How did they formulate and articulate their agendas?
• Why did the state find itself having to pay attention in some fashion?
• What did it do to explain itself?

- How, why and in what ways did the state actually respond?
- How if at all did constituency members actually benefit?
- What effect did all this advocacy activity have on the political arena?

These would be the basic questions to ask, with the main academic approach being political anthropology. Evaluation budgets would not need to be large (two- or three-person teams working with local experts would be quite adequate). Three or four countries could be selected, and in each one four or five carefully chosen donor-sponsored civil society advocacy initiatives could be assessed. Would such an effort yield a complete understanding of how best to support civil society advocacy and promote democratic pluralism over time? Of course not, but the results of a well-conducted assessment would go far in giving the international donor community a clear picture of what works and how in furthering both these objectives.

Notes

[1] The literature on this topic is immense. Book-length treatments can be found in Fisher 1995 and D'Souza 2002 among others. This chapter relies primarily on the relatively short but comprehensive and insightful accounts given by Dwivedi 1998 and Wood 1993, up to 1998. Except where otherwise stated, substantive data in this chapter come from these two sources.

[2] The essay, which first appeared in the Indian journals Frontline and Outlook in May 1999, has been reprinted by Friends of River Narmada, an international CSO supporting the NBA. It is available on the Friends' website, along with much other material (including some critical of the NBA and Arundhati Roy), at <http://www.narmada.org>.

[3] For a thorough account of the PCIJ's role in the removal of Estrada, see Møller and Jackson 2002. As with the Narmada case, except as noted, facts are taken from the well-researched account cited here, while I have taken considerable (although I hope reasonable) liberty in interpretation.

[4] Mobile phone use facilitated the gathering of an almost instant assembly of thousands at first, and in Estrada's final days crowds estimated at over 1 million. By the end of 2000, mobile phones had become plentiful; the two major companies had perhaps 4 million subscribers between them, a very large proportion of whom lived in metropolitan Manila. See Alcantara 2000 and Chandrasekaran 2001. Text messages could be sent to multiple recipients ('phone trees') for nominal cost—0.02 USD or less, in contrast to voice calls, which were much more expensive. Internet use became critical as well; the *Philippine Daily Inquirer* later reported over 1 million hits a day on its website at the height of the crisis (Magno 2001).

[5] There is a great deal written on the coconut levy, but it tends to be scattered pieces, mainly newspaper stories. A comprehensive analysis of this fascinating story remains to be written. The present account is based largely on what the author learned during an assessment for USAID in the Philippines in 2000 (Blair 2001). Other available reports are Gregorio-Mendel 1998, Parreño and Gaborni 1999, and Matute 2001.

[6] There were two parts to the 47 per cent holding: 27 per cent was claimed by BUCO as coco levy money, while Cojuangco had evidently acquired the remaining 20 per cent by other means.
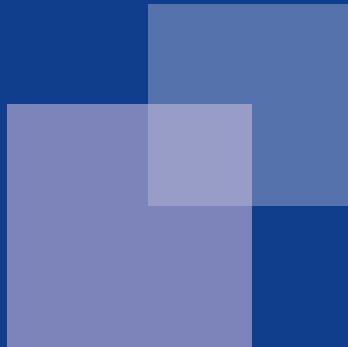
Chapter 8

Evaluation of the utility of community-level democracy support for conflict resolution: the Community Action Investment Programme in Tajikistan

Chapter 8

Natalia Mirimanova*

Evaluation of the utility of community-level democracy support for conflict resolution: the Community Action Investment Programme in Tajikistan

This chapter presents a framework and methodology for evaluating the utility of democracy, and thereby international support for democracy, for conflict resolution at the community level. It is based on an evaluation of the Community Action Investment Programme (CAIP) which was implemented by the Mountain Societies Development Support Programme (MSDSP) of the Aga Khan Foundation (AKF) in Tajikistan in 2003–5. This programme coupled democracy support with the provision of infrastructure projects as a strategy for conflict resolution. The democracy support, therefore, incorporated both its own objective and a strategy for achieving other objectives. The usefulness of democratic institutions and practices for conflict resolution is found to depend both on the quality of the interventions and on sound theory linking democracy and capacity for conflict resolution at the community level. An interdisciplinary framework of evaluation is essential. Attributing change on the ground to the democracy support component specifically is especially challenging. The chapter offers some analytical and methodological responses to the challenge.

Introduction

This chapter seeks to establish a framework and methodology for evaluating a multi-objective programme in which democracy support was simultaneously a means to promote democracy and a strategy for achieving other objectives. Attribution of an intended change on the ground to the democracy support input in multifaceted programmes of which democracy support is one component along with others is particularly challenging. In this chapter analytical and methodological solutions to this problem are proposed. The study draws heavily on an examination of the CAIP.[1]

The CAIP was a community-level democracy support initiative aimed at enhancing the community's capacity to manage and resolve conflicts. The democracy support was evaluated from the perspective of its utility for conflict resolution at the community level. The democracy support component in the CAIP–MSDSP took the form of the establishment and capacity building of a new democratic decision-making, governance and accountability institution at the level of the village, or Village Organization (VO). Democracy support was a strategy in a multi-strategy intervention in the complex environment of a Tajik village located at the geographic, political and economic periphery of the country. A parallel strategy was the provision of a vital item of infrastructure to the target communities. The creation and proper operation of VOs as participatory community-level bodies characterized by egalitarian decision making were the conditions for the provision of vital infrastructure to the target communities. In other words, the provision of much-needed resources was not only preceded by the institutionalization of democracy at a village level, but was also conditional upon the success of the latter. This was intended to enhance the target communities' capacity to manage and resolve conflicts.

The usefulness of democratic institutions and processes for conflict resolution was evaluated in two steps. First, the *quality of democracy* had to be assessed in its particular aspect as the institutions and processes designed to facilitate conflict resolution. In programmatic terms this meant assessment of how far the standards set corresponded to the performance characteristics of the democratic institutions and processes being supported. Second, the *theory of practice* that linked enhanced democracy at the community level with improved conflict resolution capacity of the community must also be tested. For this purpose, then, an interdisciplinary evaluation framework was applied that combined assessment of the quality of the new community-level democratic institutions and the relationship to a conflict intervention evaluation framework.

Evaluation of the utility of democracy support for conflict resolution: analytical framework

The belief at the heart of the CAIP was that democratic rule at the community level is manifested in everyone having equal opportunity to participate in decision making

on issues of collective concern. This would strengthen the communities' capability to alleviate poverty, effectively and transparently manage the collective resources at their disposal and raise additional resources, and plan and implement development projects. A successful intervention would facilitate the internalization and institutionalization of democratic values and behaviour to improve the means to deal with communal conflicts.

In the CAIP democracy support was coupled with the provision of infrastructure support because the CAIP was 'based on the hypothesis that encouraging communities to find solutions to the problems that most immediately affect their lives will ease tensions at the local level and reduce the potential for violence' (Ehmann, Morriss and Alimkulova 2003: 3). In other words, the CAIP was a social experiment where good performance by a newly established democratic institution at the community level would be rewarded with a vital infrastructure project, which was bound to be greatly appreciated by the target communities. From the perspective of an evaluator, however, the overlap of the democracy support programme and the infrastructure provision programme complicates the evaluation of the utility of the democracy support programme per se.

The evaluation took place at the end of the third year of the project when all the infrastructure projects had been completed. The infrastructure projects were designed as finite and one-time resource provisions. In contrast, the VOs were conceived as lasting and sustainable democratic community-level institutions that would amplify the communities' capacity to manage conflicts in the future, even when no more external resource support would be made available.

The evaluation of the link between the level of democracy at the community level and of the community's capacity to manage and resolve conflicts was divided into two tasks (see figure 8.1).

Figure 8.1: An interdisciplinary approach to evaluating the utility of democracy support
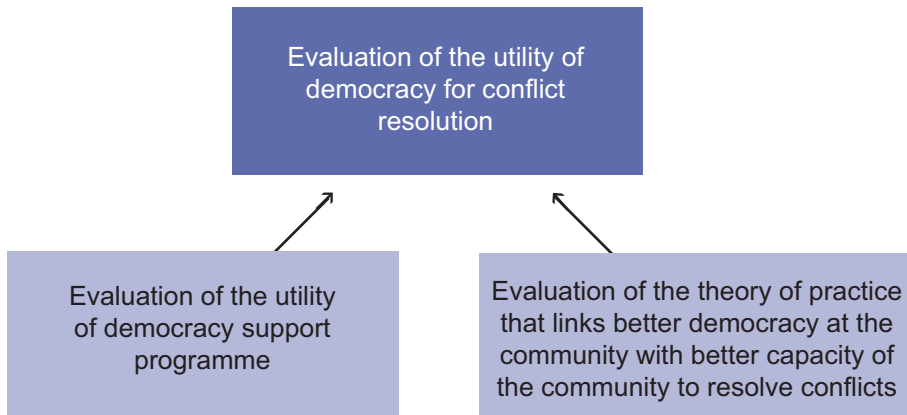
Given that the circumstances of Tajikistan are not well known in the world outside, before proceeding further it will be useful to provide some context to the democracy support programme in the form of background information.

Background information about the site of the democracy support programme in Tajikistan

Tajikistan declared independence in September 1991. Civil war broke out shortly after the first multiparty elections results had been contested by the opposition. The civil war unfolded along the lines of regional identity, which had always corresponded to the fault lines between the advantaged and disadvantaged. People did not exercise a choice, but were 'ascribed' to one camp or the other on the basis of their origin. The old communist nomenklatura, most of them from the northern Leninabad region, joined ranks with the Kulobis from the region south-east of the capital, Dushanbe. In the past the latter had always been under-represented in the political establishment, and the powerful Leninabadis hoped to exploit the Kulobis' anger and frustration in the fight against the opposition. The opposition consisted of three groups: the Qarategins in the Rasht Valley, who had been relocated to the cotton regions in the south during Soviet times (they formed the core of the Islamic Renaissance Party); Dushanbe-based intellectuals (they formed the Rastokhez popular movement with a nationalist agenda); and La'li Badakhshan, a Pamiri-centred party that promoted democracy and greater autonomy for the Gorno-Badakhshan Autonomous Oblast (region). Later some of these groups formed the United Tajik Opposition (UTO), while the Leninabadis and Kulobis were referred to as the 'government'. Both the opposition and the government had paramilitary groups on their side. The peak of the fighting was in 1992–3, when about 50,000 people were killed, many of them unarmed civilians.

The fiercest fighting and atrocities against civilians took place in the south and in the Rasht Valley. According to the United Nations High Commissioner for Refugees (UNHCR), 600,000 people were internally displaced and about 80,000 people became refugees, mostly in Afghanistan. The Tajik civil war continued as a guerrilla war in the mountainous regions until it ended with a peace agreement in 1997. By the end of the war the Kulobis prevailed in the 'government' block, and their leader, Emomali Rakhmonov, signed a peace deal with Abdullo Nouri, the UTO's leader. Post-agreement arrangements were aimed at power sharing between the government and the former opposition, which many commentators today view as a co-optation of the opposition. It paved the way to single-party rule.

The social capital of the Tajik nation was severely damaged by the civil war. The conflict itself was an outgrowth of decaying social capital made manifest in a division of the nation into the advantaged and disadvantaged, structured along lines of regional identity (Colletta and Cullen 2003). Today, the cohesiveness of Tajik society seems on the surface to be improving, and is clearly orchestrated from the president's palace. It

would be unfair, however, to say that the people resist this. Tajikistan cannot possibly survive another civil war or any violence on a mass scale, people say. It seems that all Tajiks cherish the new-found peace, and for most of them the peace agreement and hopes for a better future are personalized in President Rakhmonov—or at least they say so.

The fact that Tajiks so want the peace to continue is a serious argument against the use of violence as a means to resolve conflicts. According to Barnes and Abdullaev, 'Their history of war and violence has led many to prefer a government capable of sustaining a "negative peace" based on life without war at the price of not enjoying their full range of personal rights and liberties' (Barnes and Abdullaev 2001: para. 6). However, the ghosts of the civil war continue to paralyse a creative analysis of the current political and economic inequalities, the clanism and the lawlessness in a country where democracy is mere window dressing.

It is important to stress that at present Tajik politics and the economy are developing within the same system of regional disparity that led to the civil war. A few Kolobis are at the top of the political establishment today, which means that powerful Kulobi clans are prospering but Gharmis and Pamiris are marginalized. Inter-regional antagonism has not disappeared; moreover, the economic dimension has become more salient as a result of the economic marketization that has taken place. Only rudimentary state payments are made to the less well-off. Gharmis (in the Rasht Valley) and Pamiris are thus marginalized in both the political and economic sense. That is not to say that even ordinary people among the 'privileged' Kulob enjoy dividends from having promoted their men into the top positions. For some women and children conditions in the cotton-producing region of Kulob still resemble slave labour. The overwhelmingly popular response to the challenging economic situation throughout Tajikistan is labour migration of able young men to Russia. To avoid conflict is both an individual and collective response to an oppressive and discriminatory political climate. A word of wisdom one is often offered in Tajikistan these days is that peace can be sustained only at the expense of democracy and individual and collective freedoms.

In sum, the society at large lacks post-conflict trauma healing, restorative justice and genuine reconciliation. What it does have are poverty and significant regional and social economic inequalities; fear of a return of conflict and violence; and a lack of institutionally, procedurally and culturally guarded individual rights and freedoms such as the free expression of opinion and political participation. Democracy at the national level in Tajikistan is seriously defective regardless of the formal presence of democratic institutions and democratic procedures. That is why international humanitarian and development agencies try to incorporate local-level democracy support into their programmes in Tajikistan.

Against this background, the Rasht Valley was selected as a target region for the CAIP. It was one of the regions where civil war unfolded as full-scale violent conflict; the human and economic costs of the war were very high there. The Rasht Valley was

the base of the mujahedin fighters. After they were forced to retreat to the mountains in 1993, the government's army entered the valley, and agriculture, its major asset, was devastated. Every family in the valley was affected.

The programme unit was a village-based participatory structure. The aim was that it should reach collective decisions on the priority infrastructure projects for a particular village, calculate the costs of the projects selected, assess the monetary equivalent of the community's in-kind labour contribution and assume responsibility for the ongoing maintenance of the projects. The MSDSP-initiated village-level democratic decision-making structures were the VOs. The VO model of village-level governance, decision making and representation had already been piloted by the AKF in Pakistan and proved to be an effective self-governance institution there. Its operation in Pakistan was associated with a substantial increase in the standards of living in the villages (World Bank 1995). Beginning in 1998, the MSDSP began fostering the development of VOs in the Gorno-Badakhshan Autonomous Oblast of Tajikistan. CAIP village organizations were also set up in the Rasht Valley.

Each VO has four main leaders—a president, a women's group leader (who usually sits as the vice-president), a manager and an accountant. Participatory democracy at the village level is enacted in the Village Development Planning Process (VDPP). The VDPP takes place once every three years, or more often depending on the emergence of new community needs that were not factored into the previous VDPP or on the availability of new donor projects that could benefit the village. The VO also holds monthly meetings that are open to all villagers. An important feature of a VO is the Village Development Fund (VDF) that is formed from grants, and individual donations and fees. This is a revolving fund; villagers can take low-interest credits. Their interest payments help to replenish the fund.

An important feature of the democracy support component of the CAIP was that the formal local government structure (the khukumat) was required to contribute to the infrastructure project, along with the village itself. Khukumat-level officials in Tajikistan are appointed, not elected. A khukumat usually does not have any resources, nor does it have a transparent mechanism for funding communities if resources are put in place. It often solicits funds for urgent social projects such as sewage disposal or electricity or water supply from international aid agencies. Overall, people do not see the khukumats as being responsive to their needs. By making VOs a primary unit of the infrastructure provision programme and involving the local government institutions as a partner, the CAIP aimed to develop a representative function on the part of the VO and make official governmental structures more responsive and responsible to people on the ground. It was thought that this would narrow the gap between people on the ground and local government structures, and enable the concept of partnership between formal and grass-roots governance structures to take root.

Evaluation of the democracy support

The democracy support was evaluated throughout the CAIP by the Programme Evaluation Units (PEUs) of the central and regional offices of the provider organization, the MSDSP. The evaluations were carried out mainly on a quantitative basis. The indicators were the number of VO training events completed, the number of VOs established, and the accounting books of the VOs in their role as managers of the revolving village funds. In the present evaluation study the quality of democracy at the community level was assessed in two aspects that were relevant for further evaluations of the utility of the democratic institutions established for conflict management and resolution at the village level:

• the performance of VOs as facilitators of democratic decision-making in the implementation of infrastructure projects; and
• the sustainability of VOs as a democratic decision-making and governance community-level structure after the infrastructure projects were completed.

Qualitative research methods were chosen for the evaluation we consider here. Objective, quantitative indicator-based assessment of the difference between the pre-CAIP and post-CAIP situation as regards the state of community-level democracy and capacity to mitigate conflicts was not feasible. This was because the key 'state of democracy' and 'conflict resolution capacity' variables to be measured had not been identified or measured before the CAIP started. Moreover, even if a set of indicators had been designed and measured both prior to the CAIP and after its completion, valid conclusions could not have been drawn because CAIP projects varied substantially across communities, as did the communities themselves. The time frame of the evaluation field research did not allow for a large sample to be taken, which would be necessary in order to approximate a quasi-experimental comparative study.

In addition to the impossibility of quantitative indicator-based evaluation, another factor made qualitative research methods preferable. Quantitative study of the impact of the democracy support component of the CAIP on the communities' capacity to resolve conflicts and mitigate the sources of conflicts would not have revealed the meaning of the democracy-based provision of infrastructure projects in the target communities to the communities themselves, to the provider (the MSDSP) or to the local authorities.

Subjective perceptions of the performance and sustainability of the new democratic institution and decision-making procedures at the village level were collected in *semi-structured* interviews with ordinary village residents, VO leaders and traditional authority figures, namely religious leaders and elders. These were supplemented with *direct observations* of village life and people's interaction with VO leaders after an infrastructure project had been completed. Examples of the questions asked by the evaluators are listed below.

- How were decisions made on the priority infrastructure project (description)? How did you personally feel about the process? What was the procedure? Did people vote? Was there a discussion? Were people trying to convince each other using rational arguments or cases? Did anyone try to apply pressure?
- What were alternative proposals and how many people supported them?
- Were all villagers happy about the project selected? Did those whose project idea had not been selected feel resentful? What was done about that?
- Do you think someone from your village—a rich or a well-connected man—can take individual advantage of (privatize, for example) the communal water system (other infrastructure project)? Why? Does the VO have a role here?
- What is your attitude towards the Village Development Fund? Do you think it is transparent? Do you agree with how the money is spent? Do you think all village residents have equal access and responsibility as regards the VDF? Should they?
- What if someone does not pay the borrowed money back or does not pay the interest? What role does the VO play in these situations?

The purpose of these interviews was to discover whether the beneficiaries of the programme could see the value of VOs and democratic procedures irrespective of the offer and provision of new external resources.

Conflict evaluation framework: reconstruction of the theory of practice of the programme

It is important to emphasize that the provider organization, the MSDSP, did not have experience or expertise in conflict resolution programmes prior to the CAIP. The MSDSP incorporated the CAIP into the scope of its activities and carried out the project as just another development project. This posed a challenge when it came to evaluating the impact of the democracy support (CAIP) component on the community's and the MSDSP's capacity to manage conflicts and mitigate the sources of conflict. In the CAIP documentation the relationship between democracy support and conflict resolution at the community level had not been grounded in theory, nor had it been operationalized in the course of the programme.

The first step in the evaluation of the utility of democracy support for conflict resolution was to make more explicit what was at the most an implicit CAIP theory concerning the relationship between democracy support and conflict resolution. A conflict intervention evaluation framework implies that there is a theory of conflict (why there are conflicts), a theory of conflict resolution (what needs to be achieved to resolve conflicts) and a theory of practice, or a conflict resolution strategy (how to get to the stage of conflict resolution) (Church and Shouldice 2003).

There was no conflict theory in the CAIP, and 'conflict' there appeared to be a generic term for any level and type of social tension. Some informants mentioned that most often conflicts at the community level emerge over scarce resources. The

infrastructure provision component was incorporated into the programme to tackle the root cause of these resource conflicts. A democratic process of decision making was necessary if the community was to generate consensus on the priorities among several possible infrastructure improvements. It should be pointed out here that the Aga Khan Foundation in contrast has always operated with its own theory of conflict, which states that 'all conflicts have roots in poverty, therefore poverty alleviation will reduce conflict' (Ehmann, Morriss and Alimkulova 2003: 8). The designers of the CAIP, in contrast, were not specific about their theory of the causes of conflict in the targeted regions, apart from being made aware of the general background of civil war, poverty and unemployment.

No distinction had been made in the approaches to conflict that unfolds at different levels: within one community, between communities, and between a community and a local government. The conflict typology outlined in figure 8.2 would have been useful for VOs in planning their conflict interventions at the village community level. In practice the conflict management training that was delivered to the VOs throughout the CAIP focused only on conflicts over scarce resources within and between communities. The parties to the conflicts were assumed to have equal power. However, when unequal power distribution is acknowledged as a possible cause of conflict and when the parties significantly differ in their power, then democratic decision making within a VO is unlikely to be sufficient to manage or resolve the conflict. To elaborate this point, figure 8.2 incorporates resources and power as sources of conflict.

Figure 8.2: Typology of village community conflicts

Level of conflict / Source of conflict	Between groups within one community	Between communities	Between a community (-ies) and a local government	Type of conflict
Resources (shortage of water, arable land, pastures, finances, seeds)	Groups within one community compete for a scarce resource	Communities compete for a scarce resource	People expect local government to provide vital resources; local government is incapable of providing the resources	SYMMETRIC
Power (access to the available resources, access to decision making, especially on the matters of immediate relevance for the community)	Powerful figures in the village usurp access to a resource (water, land), and the powerless experience shortage of the resource; confrontation	One community has access to a resource, while another does not; confrontation	Local government abuses power and denies a part or all of a community(ies) access to vital resources	ASYMMETRIC

At any one time one and the same village can be engaged in several different conflicts that require different conflict mitigation strategies, because their causes differ. Careful conflict analysis would help design a tailored and cost-saving conflict resolution strategy. The CAIP, however, was designed to address symmetric conflicts only—conflicts that appeared to be symmetric only because no analysis of the power imbalance had been carried out.

The conflict resolution theory of the CAIP was that improved life conditions in the form of short-term and long-term jobs and vital infrastructure will rebuild communities and bring hope, thus diminishing the likelihood of another violent confrontation. The strategy for improving living conditions was to develop a community's ability to decide democratically on matters of common concern, coupled with the provision of vital infrastructure aid. This was the CAIP's theory of practice. In the communities characterized by high levels of poverty and few resources, democracy support was coupled with the provision of resources. The new participatory democratic structures were considered unlikely to succeed unless villagers associated them with tangible improvements in their material conditions. The provision of external resources—chiefly money for projects—was supposed to increase public acceptance of democracy at the community level as a strategy to collectively manage village resources, raise additional new resources and resolve conflicts. This theory of practice is summarized in figure 8.3.

Figure 8.3: Conflict intervention evaluation framework: theory of practice

Provision of resources and technical training to the community	Democracy support at the community level

Village organization (VOs)

Provision of the missing vital infrastructure defuses tensions within and across village communities	Participatory decision-making on the priority infrastructure projects and required community contribution to the implementation of the projects ensure transparency in the communities, on the one hand, and encourage them to assume collective responsibility for the communal infrastructure, on the other	Involvement of local authorities as partners and contributors democratizes relationships between formal authorities and informal democratic village representative institutions (VOs)

Capacity of local communities to mitigate conflicts over resourse and power is strengthened

Potential for violence reduced	More effective management of resources, better capacity for the community to raise new resourses for development

The strategy of the evaluation research was to elicit the perceptions of the value of community-level participatory democracy as a mechanism for managing the community conflicts of both the beneficiaries of this democracy support and the provider organization staff. Evaluating the utility of participatory decision making at the community level (in the VO) for communal conflict resolution in operational terms

meant eliciting the extent and the nature of the changes as experienced by community members.

The methodology for evaluating the utility of democracy support was the same as that used for the general evaluation of the democracy support—semi-structured interviews and direct observation. In addition to these methods a scenario-building method was applied to encourage people to think about the dynamic of hypothetical or actual conflicts in the communities where VOs had been established, and also in communities that did not participate in the democracy support programme.

The following research questions focused on the past, current and future performance of the community-level democratic decision making, governance and representative structure.

- Did VOs through the CAIP contribute to the mitigation of sources of community conflicts and develop the village's capacity to manage community conflicts?
- Did VOs through the CAIP offer the communities new, satisfactory and peaceful methods for dealing with community conflicts?
- What elements of the design and implementation of the democracy support were likely to strengthen the VO as a sustainable democratic institution at the community level—one that can be an enduring facilitator of collective resource management and conflict resolution long after the CAIP? What elements of the design and implementation of the democracy support may have weakened the VO for this purpose?

In order to retrospectively reconstruct a conflict baseline in the target communities, village residents were asked the following questions.

- Were there conflicts between the villagers [between villages] because of the lack of infrastructure (shortage of water or restricted access to water, or overcrowded schools and children being 'territorial' about a school in their village and hostile towards children from other villages)? What were the manifestations of the conflicts? What would people do to prevail in the conflict? Were there attempts to find mutually satisfactory solutions?

Challenges facing the application of the conflict intervention evaluation framework and some methodological solutions

The conflict intervention evaluation framework was developed after the programme had been completed. The conflict intervention impact of the democracy support had not been incorporated into the original criteria for selecting the communities. The baseline of conflict-proneness of the communities selected had not been recorded. The communities selected differed in their proneness to conflict, and yet the same steps towards establishing VOs, facilitating village development plans and training

VO managers were taken in all these communities. This made it difficult later to attribute anything to do with conflict-proneness to the implementation of democracy support. This is because, whereas in some villages the strategy might have resolved a pre-existing conflict over, say, water or an overcrowded school, in some other villages where no such conflict had existed the intervention could merely have helped people improve their daily lives. Clearly, all these weaknesses could have been remedied very easily by thinking through the situation more carefully in advance.

The indicators that were used throughout the CAIP were the numbers of training events, infrastructure projects, communities served and others that were descriptive of the scope of work, staff performance and the allocation of the funds. These indicators could not offer information about whether democratic ways of approaching communal conflict resolution had been taken to heart. Communities had been selected for participation in the CAIP on the basis of their infrastructure needs only. Even when the communities that were chosen were appropriate in terms of conflict, there was a 'disconnect between the conflict issue identified in the initial profile and the infrastructure project that follows, even when a fairly clear relationship exists between the problem and the project' (Ehmann, Morriss and Alimkulova 2003: 5).

This is a common situation in evaluation research when values for the variables that constitute the evaluation framework have not been reflected on prior to the intervention or programme. Subjective assessment of change by the beneficiaries was employed as a methodological solution. In order to reach valid conclusions, triangulation of the assessments of change as perceived by beneficiaries (different stakeholder groups, if applicable) and by the provider organization, and of the assessments of the actual state of affairs at the time of evaluation as observed by an external evaluator, should be carried out.

The combination of democracy support and economic assistance to prevent potential violence and to facilitate the institutionalization of democratic means and the internalization of democratic values and behaviour was a sound approach, given the circumstances in the targeted region. However, it posed a particular challenge to the attribution of changes on the ground to the democracy support component of the intervention. To address this challenge, conceptual and methodological strategies were employed. At the conceptual level, the infrastructure provision could be incorporated into the democracy support component. This is because of the required sequence of the democracy support activities as such (training and consultations for VOs in the facilitation of VDPP, conflict management, accounting and so on) and the decision to make the support conditional on a democratic approach to decision making. This addresses the attribution issue. However, other methodological strategies were employed to try to identify the impact of the democracy support component.

One strategy was 'scenario writing' with the beneficiaries and the provider organization staff. They were all asked to carry out 'mental' experiments and contemplate situations where either the democratic support component or the economic assistance component had not been in place. In addition, they were asked to

paint scenarios of the future performance of VOs as facilitators of conflict resolution and or mitigators of sources of conflicts, in the event of outside resource assistance being reduced or withdrawn.

A second strategy was to compare the conflict potential and the VOs' preparedness to facilitate conflict prevention and resolution in villages where infrastructure projects had been completed some time ago, or where no infrastructure projects had yet been provided (another round of economic assistance programmes was anticipated).

The conflict intervention evaluation framework: findings and recommendations

Applying the conflict intervention evaluation framework yielded several findings and these can be supplemented with additional recommendations to enhance the nexus between democracy support and the conflict resolution capacity of the communities.

First, the participatory decision making on infrastructure projects and the required contribution by the communities and local government, married with the external provision of financing to implement the projects, was recognized by the communities to have defused tensions and helped the development of a sense of personal and collective ownership of the projects (Fuller and Mirimanova 2005). All the villages benefited from the establishment and operation of VOs and the subsequent provision of vital infrastructure. One of the cases made an educational film to teach new VOs and village residents conflict resolution through democratic mechanisms (see annex 8.1). An example of how the VO process reconciled the rich and the poor in a village is described in annex 8.2.

Second, however, the evaluators' forecast of the sustainability of VOs after the provision of external resources ceased was not unreservedly positive. The internal village hierarchy was not carefully factored into the design of the VO operation. In time this may weaken the VO's conflict resolution capacity, as the principles of its operation (egalitarianism, participation, transparency and accountability) could well run counter to the established ways in which the other community institutions operate. Competition between VOs and the traditional hierarchical structure of the community institutions may well unfold.

In fact mixed results were obtained regarding the unique role of the VO vis-à-vis other community-based informal institutions. Village residents often confused traditional informal authority and decision-making community-level structures, such as the choikhona (mosque), mahalla (community), the hierarchies of clan, wealth and political connections, and the aksakals (elders') council, on the one hand, and the VO, on the other. This confusion may have stemmed from there not being enough awareness training in the new, more participatory structures. In many cases traditional village leaders had been elected as VO leaders. In the Gorno-Badakhshan Autonomous Region, for example, in many cases VOs fused with the choikhona structure. Women

participate in social life far more freely there than in the Rasht Valley, for instance, and this fusion did not affect the performance of VOs significantly. However, this kind of overlap could well impair the operation of VOs in other more religiously conservative regions.

On the plus side, it should be stressed that the introduction of a VO allowed new leaders to emerge among women and the young people, the categories that would not have been able to take leading positions in the traditional village hierarchy. This was an important change. Further efforts to nurture this kind of development will be necessary: the fact that the villages' internal hierarchy was not factored in to the operation of a VO may lead to democratic decision making being hijacked by the most powerful figures in the village. It was discovered that in some cases open voting in the VO meetings made it difficult for some residents to make their voice heard, or they were pressured to vote in ways that were contrary to their own interests. In such circumstances even a one-person-one-vote system does not always ensure democratic decision making. The people's trust in democracy could be badly affected. Annex 8.3 illustrates how an internal village hierarchy can come to disrupt the objectives of the VOs.

Third, in the CAIP democracy support strategy for conflict resolution the provision of external resources for infrastructure projects was a substantial part of the solution. However, VOs' capacity to deal with conflicts must be developed beyond using this tactic to get community-level democratic structures accepted. It is quite possible that if it had not been for the CAIP's offer of money, the tensions found in the communities in conflict in most cases would not have been significantly eased, even with the VOs in place. The VO-based democratic decision-making process could have remained just another training exercise.

Notwithstanding the generally excellent implementation of projects by the MSDSP, the fact that CAIP conflict resolution theory was not matched with a theory of the causes of conflict weakened the systemic impact of the programme on the way conflicts are dealt with in these rural communities. Problems of a different order in the communities were often confused with conflict. The assessment of conflict potential in the Rasht Valley that had been carried out prior to the CAIP identified land distribution and lack of infrastructure as potential sources of conflict. Improvements to the infrastructure were selected as a focal point for the CAIP initiative. However, while lack of infrastructure is a problem, it is not necessarily a cause of the conflict, or may not be the sole cause. It can become a divisive issue if one section of the community enjoys, say, access to a school or a water supply system and another section does not, especially if this inequality is mapped onto other differences such as ethnicity or class. To prevent the recurrence of similar conflicts in the future, conflict resolution must offer not only a material solution but also the restoration of damaged relationships and structural change, once the causes of conflict have been identified as lying in the political, economic or social structures. The restoration of relationships was assisted by the introduction of participatory decision making and most importantly by the

elimination of the most divisive issue. However, the structural causes of conflicts were not identified or addressed through the CAIP.

Bringing in external resources to equalize power is not always an option. Besides, some conflicts that are in fact over power and rights cannot be resolved in this way. Moreover, the resources may actually stir up conflict where the conflict potential is of an asymmetric nature and the existing power imbalance is dramatic. CAIP-style development projects can thus actually cause difficulties over setting priorities regarding the use of much-needed extra resources for the village—as well as affecting the decision-making processes in the village. Those village residents whose preferred choice of project is not favoured could feel resentful, especially if the projects that are selected are then mismanaged. Another potential source of tension may be the unwillingness or inability of some residents to assume full responsibility for the maintenance and other costs of the new infrastructure improvements. Although people kept saying that this could not happen, because of the sense of personal shame that would result, the poverty in the region is such that some people may simply be unable to contribute their share. This illustrates how democracy and democracy building can become entangled with both conflict and development issues at the community level.

In situations where the conflict between communities is asymmetric, the external provision of resources may be necessary but is not sufficient (see annex 8.4). A more penetrating analysis would indicate that the nub of the issue has more to do with how resources are used and shared among and between communities, and that some attention must be paid to the rules that govern these arrangements. Annex 8.5 provides a graphic illustration.

Fourth, it is fair to say that the democracy support component of the CAIP helped advance the VO's role as a facilitator of participatory decision making and a manager of collectively owned resources in symmetric conflicts over access to scarce resources or interpersonal conflicts. However, where the root of the problem lay in an imbalance of power between the conflict parties or the cause of conflict was more structural, such as discriminatory laws or corrupt law enforcement practices, the VOs had neither the expertise nor the mandate to intervene. In these situations more sophisticated democratic procedures are required. It seems that asymmetric conflicts must be addressed by more popular grass-roots structures than the current VOs, whose inability to resolve disputes by means of democratic procedures over such issues as land conflicts is a major weakness.

To conclude, the CAIP created a unique situation where local governments were given a chance to demonstrate their concern for ordinary people and a desire to accommodate the community's needs, in contrast to the existing local authorities in Tajikistan, which are not elected. The CAIP helped to strengthen the legitimacy of the VOs in the eyes of the communities and the local government, both as a community representative and as a community service institution. In time, ordinary village residents through the VO or some other community-level and inter-communal representative

body might develop into a partner in negotiations with the local authorities. The MSDSP could assist communities and their grass-roots representatives in working towards that end. It is true that some conflict prevention and resolution mechanisms already existed within village communities. Examples are the rule of the elders and ways of reducing material inequalities among villagers through collective help (hashar) or even the provision of charity by the rich villagers, and the self-regulation of inter-family violence and revenge. Nevertheless, these mechanisms work best in cases of symmetric conflicts between families or individuals who are of comparable wealth and influence. Such mechanisms do not work well when conflicts are asymmetric, that is to say they involve parties one or more of which dominate the situation by virtue of their greater wealth, political power or physical capacity to intimidate the rest. In these situations the traditional mechanisms for conflict resolution might actually work to the advantage of the most powerful party. Thus imaginative new citizen-based institutions for conflict resolution will be required. Possible examples are public hearings, consultative councils and conciliation commissions. Given that the MSDSP is in the vanguard of the creation of participatory community institutions and procedures, that programme is well placed to nurture such new participatory conflict transformation mechanisms.

The utility of democracy support evaluation frameworks at the community level

Democracy is believed by many to be the most effective conflict resolution mechanism both in the international arena and at the community level. Democratic institutions allow for timely and comprehensive conflict analysis and early warning. They can usher in structural and procedural conflict management and non-violent conflict transformation and resolution through political and social change.

Democracy was approached in this evaluation case as an empirical, not an ideological, concept. Its utility for conflict resolution at the community level was tested within a conflict intervention evaluation framework. Although the evaluation was modest in scale, face-to-face meetings with the beneficiaries, trust-building and good knowledge of the context and the local language enabled the evaluators to elicit the meaning of the democracy support intervention for conflict resolution as viewed by the people on the ground.

The overall research strategy for evaluating multi-strategy interventions in actual or potential conflict situations should take the form of most similar system design (MSSD) and most different system design (MDSD) when selecting cases for comparison. The target communities should be combined into sub-cases according to the value of the variables of concern, for example, according to the type of conflict, its stage, the level of background poverty, any previous record of successful or unsuccessful management and resolution of community conflicts, and so on. In multi-strategy interventions like the one described in this chapter, analysis of the conflict potential

and the established community-level institutions' capacity to facilitate conflict prevention and resolution in those villages where democracy support has been paired with infrastructure support should be compared with the situation in villages where democracy support has not been paired in this way. This could constitute a next step in the research. In the real world, however, that kind of preparatory reflection seems not to happen often enough. In this case, therefore, a retrospective reconstruction of the communities' pre-intervention baseline had to be carried out at the evaluation stage. A participatory scenario-building exercise in which the evaluators solicit the views and insights of the project beneficiaries and the provider organization after the programme has been completed offers a practical alternative to researching baseline indicators in advance of the project itself.

Democracy support programmes are often implemented in situations where a complex assortment of political, social, economic and cultural problems has already emerged. This may mean situations characterized by protracted violent conflicts, economic decline, poverty, decades of a totalitarian regime and massive human rights violations. Brainwashing, xenophobia, corruption and extreme nationalism may also be present. This chapter has focused on one such situation, where a combination of democracy support and economic assistance was employed. An interdisciplinary framework for assessing the meaning and impact of democracy support intervention at the community level was applied. The present democracy support programme is an example of what has been called the 'promotion of democratic structural stability', which Bigdon and Korf (2004) recommend as a basis for development work in conflict societies. The current generation of development projects even in non-conflict situations tends to include participatory and empowerment approaches as an integral component anyway.

The evaluation framework presented here could be a useful tool for comprehensive assessment of development–democracy interventions. However, to realize its full potential it must be approached as a theory-driven evaluation and ideally should be designed at the time when the intervention is being proposed, and not afterwards. The analysis that is undertaken of the causes of the conflict on the ground must shape the intervention—and that will have implications for the evaluation methodology too. Some of the conflicts described here were simple resource conflicts and they could be resolved once the community arrived at joint decisions about the use of external resources. The democracy support component aimed to teach the communities a process for addressing conflicts. However, this mode of intervention will not be adequate for every case of conflict, where the underlying causes may be both deeply rooted and diverse.

Attempts to build democracy from the grass roots could be the most promising entry point for the mitigation of sources of community conflicts, and for fostering democratic means for conflict resolution. This applies especially where the local traditions that used to maintain the cohesion of communities and mutual help have degraded under the impact of economic marketization, outward migration

and widespread poverty. Democracy support at the community level offers the only sound, non-confrontational approach to improving conditions for the many, in the presence of authoritarian rule at the national and local level or where formal democratic institutions have long been emptied of their democratic content. In the final analysis a more carefully worked-out integration of democracy support expertise and conflict analysis expertise is needed to amplify the value of the supported democratic institutions and procedures for the management and resolution of communal conflicts.

Note

[1] The final evaluation mission was carried out in October 2005 by two international consultants, Natalia Mirimanova (Russia) and Stephan Fuller (Canada). The central and regional MSDSP staff wholeheartedly participated in the evaluation and devoted their time and resources to supporting the evaluation research and to learning from the findings.

Annex 8.1. Conflict resolution: the movie

The film presents a real-life story of the conflict between two villages in the Tojikobod district over drinking water. One village is fortunate as it is upstream, while Novobod village is downstream and is destined to rely on what is left for its drinking water. There has always been a shortage of drinking water for all the villages along the stream and its use has had to be regulated, otherwise the downstream village would always have ended up without water. As a measure for the regulation of the water use and for the prevention of violence, an agreement was reached between the two communities, according to which the downstream village (Novobod) had access to drinking water before lunch, while after lunch time the upstream village blocked the water flow and had all the water for itself. However, on occasions the agreement was being violated by the upstream villagers: the regulation mechanism complicated their life. The agreement was not binding, after all, and the upstream villagers' good faith was the cornerstone of it—otherwise there was no incentive for them to share water with the downstream village.

In the film, the conflict is presented through the prism of the needs and frustrations of the Novobod community, the downstream village that is vulnerable and does not have the leverage to prevail over the upstream village. The Novobod community decided to address the issue of the supply system for themselves. The VDPP meeting was filmed in a great detail, depicting debates, proposals and arguments, priority-setting procedures and voting processes. As the VO did not have the financial means to buy the materials needed to construct a water system, it applied for CAIP assistance. The CAIP supplied materials, and villagers donated their labour in the traditional form of hashar. The cause of the conflict between the communities was eliminated. The evaluators visited this village and learned that not only was the issue resolved, but relationships between the two formerly rival villages had been restored. A party was held to celebrate the completion of the water system construction.

Annex 8.2. Overcoming established political inequalities

A mountain spring has always been the sole source of drinking water for village X in the Tavildara district. One day three rich and hence influential and well-connected families usurped the spring and diverted the water pipe into their households. This stirred anger among other village residents: occasionally fights would break out; but the majority of people were so intimidated by the possible consequences of a clash with the three powerful families that they preferred to keep silent. When a VO was established, and everyone joined in, including the three powerful families, a collective decision was reached to install an additional water pipe system and water taps in the village. Peace in the village was restored.

Annex 8.3. The limitations of the Village Organization

In the village of Ezgand, in Tavildara district, a new drinking-water system was installed, and all the taps ended up being within the households of just five families, who happened to belong to the same powerful clan. This was counter to the general MSDSP/VO policy as regards the installation of water taps in the villages that receive CAIP assistance. The MSDSP rule was that the taps should be installed in the streets so that every villager had access to the water at all times. The clan leader argued that after a cow had damaged one tap the village made the decision to move the taps inside the households and that the VO agreed to this. However, the dominant position of the powerful clan leader could lead to serious conflict in this village in the future. After all, many families from Ezgand had moved to the cotton-producing areas of the country some time ago, and one of the reasons for this was the absence of a proper drinking-water system. Now they are returning to their village of origin, in part because of the improved drinking-water situation.

As the number of returnees grows the present arrangement with the water taps may spark conflict between those who never left the village and the returnees. The VO as a democratic conflict resolution mechanism is in this case defective, because it is led by a representative of one of the potential conflict parties.

Annex 8.4. Infrastructure support as a temporary solution

The village of Qualakum in the Gharm district (which does not have a VO) was involved in a conflict over drinking water with the neighbouring village. The latter had its own water pipes and taps, but Qualakum had none. Their options included taking water from the ditch or walking to the neighbouring village to fetch water from their taps. Conflicts would regularly break out among people competing for the same resource.

The solution came with CAIP–Mercy Corps assistance. The khukumat granted the village the right to construct an additional feeder line from the main water pipe but with a limited number of taps. Mercy Corps provided the materials and the village donated the labour. For the time being the issue is resolved, but population growth could cause additional demand for taps, and local disputes may re-emerge if the communities cannot fund further improvements out of their own resources.

Annex 8.5. Conflict resolution beyond simply providing resources

The same village of Qualakum was involved in two conflicts over water. One was over irrigation water and developed between the households on two streets. Another was over access to drinking water and developed between two villages.

The first conflict erupted when people on one street blocked the mountain water stream going to the perpendicular street and channelled it to their own gardens—something that tended to happen at night. These attempts to divert water ended after the villagers themselves decided to establish a system whereby the streets would take turns to obtain irrigation water. A feast in the choikhona crowned the successful conflict mitigation. No external resources or expertise were solicited.

The Sholonak VO in Gharm district acted intuitively and proactively to reduce conflicts that might have arisen following the installation of a new water supply system that villagers were obliged to pay a fee for. They decided at the village meeting to subsidize poor families and let them use the water free.

Chapter 9

The evaluation of democracy support programmes: an agenda for future debate

Chapter 9

Patrick D. Molutsi

The evaluation of democracy support programmes: an agenda for future debate

This chapter presents arguments in favour of and experiences relevant to developing a global indicator for the evaluation of democracy assistance programmes. It is no easy task to evaluate the overall impact of international donor-financed projects involving civil society groups, central and local government entities, election management bodies, a parliament, political parties and the like on democratic development in individual countries. Different methodologies are used by development agencies, recipients of assistance and scholars to assess the impact of democracy assistance. Different actors appear to be measuring different things as they seek to address the specific concerns of different audiences. The time has come to develop a common tool and indicator for impact evaluation—one that all can use. The examples of Freedom House, Transparency International, and the Human Development reports of the United Nations Development Programme in developing what have become globally influential indicators show that, while meeting the challenge may not be a smooth process, it is nevertheless possible. With regard to the goal of an indicator for assessing and evaluating democracy assistance, International IDEA among others has made a good start through its State of Democracy project, and should now proceed to develop this further by incorporating appropriate modifications.

Introduction

The papers presented at the workshop on Methods and Experiences of Evaluating Democracy Support, jointly organized by International IDEA and the Swedish International Development Cooperation Agency (Sida) in April 2006, specially

revised for this collection, carry one common but important message concerning the evaluation of democracy support. The message is that there is a great deal of good, empirically rich work taking place on democracy support around the world. However, these papers, coming from different agencies and analysts operating in different countries and regions, also show clearly that when it comes to a common method or methods for assessing the impact of the billions of dollars currently being invested in democracy assistance programmes, 'the jury is still out'. There is no common approach to programme design or to monitoring and consequently evaluation. The absence of one evaluation methodology and set of indicators that are commonly accepted by democracy supporters, the executors of democracy support programmes and their targets and beneficiaries also appears to cloud project goals and the objectives of democracy assistance programmes. Moreover, the question whether evaluation should be based on project goals and objectives or on the impact of projects on the wider political system seems to remain a subject of discussion as well.

However, reading through the chapters in this volume and the general literature on this subject, it comes across very clearly that the researchers and those involved in democracy support usually agree on three key points. First, they agree that the method or methods for assessing/evaluating democracy support programmes must always be participatory in nature. Second, they agree that whatever tools or methods are used for evaluating democracy support, programmes must use both qualitative and quantitative techniques to arrive at indicators of what has or has not been achieved on the ground. The third point on which the contributors to this collection have invariably agreed is that the pressure is building on both the donor agencies and the recipients of democracy support to show results for what is now some decades of democracy assistance in the countries of the 'third wave' and beyond. This pressure is indeed widespread, coming from donor governments and their taxpayers, the populations of democracy support recipient countries, and academic researchers. All these key players in the democracy support 'chain' are concerned about what appears to be slow progress or even regression taking place in the democratization process. Unfortunately, recent developments in countries as far apart as Haiti, Fiji and Thailand, where democratically elected leaders, corrupt as some may have been, were overthrown by the military, have not helped the cause of democracy support. These countries have been among the leading recipients of democracy assistance over the past two decades. Their democratic reversals, when added to the lack of democratic breakthroughs in other countries such as Burma (Myanmar), Vietnam, North Korea and Cuba, and in the Middle East, show the size of the challenge that the democracy support agencies still face in justifying their demand for more funding for democracy support.

The apparent disillusionment with the lack of progress on democracy assistance cannot be taken lightly. The challenge to both funding agencies and the recipients of democracy assistance is to show that the funds they disburse and spend are indeed making a difference not just to the limited number of programme beneficiaries but

to the system-wide democracy processes in each country. The search for a common evaluation methodology is clearly urgent. What should be the next actions to develop such a methodology? What form and level of focus should such a methodology take? Should the evaluation focus as it currently does on achievement (or lack of achievement) of the project goals and objectives? Or should it focus much more broadly on the impact that individual projects have on democracy at the country level? This contribution to the debate on the search for an evaluation methodology attempts to make proposals on these key questions. This chapter, unlike most of the others in this volume, is based not on case studies but on the author's own experience of designing and promoting a particular democracy methodology and working with other donors in the search for common methods for supporting and evaluating democracy.

Democracy support in context

It is now close to two decades since governments, intergovernmental organizations and major non-governmental organizations (NGOs) in the Western world began massive investments in programmes to promote democracy in Africa, Asia, Latin America and the terrritory of the former Soviet Union. It is estimated that between 3 and 4 billion US dollars (USD) in total are being disbursed annually by the USA and the European Union for the purpose of assisting democracy development abroad.

Democracy assistance money has gone to numerous state and non-state agencies in the recipient countries (see the case of Sida examined by Fredrik Uggla in this volume). Among the main recipients have been electoral management bodies (EMBs) and election-related processes; rule-of-law and judicial activities; work to strengthen parliaments and local government; civil society groups in human rights, media advocacy, and so on; political parties; and academic research on democratization process. Carothers (1999) provides a broad overview. Some recipients of democracy aid have received support for the third and fourth cycle of their activities from the same donor(s).

The rise of democracy assistance has also witnessed the emergence of democracy assistance 'middlemen' based mainly but not exclusively in the donor countries. The middlemen or 'players' comprise new and old institutions which have emerged solely to promote democracy or, like the United Nations Development Programme (UNDP), have recast part of their mandates to enable them to facilitate the processing, disbursement and monitoring of democracy assistance in countries especially but not only in the global South. Institutions such as International IDEA and NGOs including the political foundations, Transparency International, Rights & Democracy (Canada), and political party institutes/centres in such countries as the United Kingdom (UK), Norway and the Netherlands are examples of this category of middle players in the democracy assistance chain.

Traditional development assistance agencies in developed countries, such as Sida,

the Canadian International Development Agency (CIDA), the Norwegian Agency for Development Cooperation (NORAD), the UK's Department for International Development (DFID), the United States Agency for International Development (USAID) and others, have become the natural and main vehicle for dispensing democracy assistance. These agencies have, however, approached democracy/ governance assistance programmes as 'normal' assistance programmes where funds are given to the provider of a service/activity and the results/outputs have to be produced following a year or two of programme implementation. It has largely been in the politically highly sensitive area of political party funding abroad that development agencies have opted out and preferred to pass the responsibility to newly established political agencies or institutes (as has been the case in the Netherlands and Norway), or to political foundations or political institutes (as in Germany, the United States and the UK), which can reach out to their partners and counterparts abroad.

The approach to democracy assistance has, however, raised a number of new questions and challenges for the development assistance community. Among the key questions that have arisen and are increasingly being raised with the passage of time are the following. Is democracy being achieved? How can we tell, and what results should and could be shown to provide an answer this question? Who are the beneficiaries of democracy assistance and are they the worthy targets?

Those who take stock of democracy assistance seem to be divided as to its impact. On the one hand, development agencies believe that the assistance is making some impact. Hence the continued and increased funding to democracy projects and programmes. Every three years or so major evaluations are undertaken, which show that projects/programmes were conducted as planned and that the target group(s) was/were reached. On the other hand, the more general analysts, such as Carothers (2006a), say that there is no or at best very little progress being made, and that any such progress is limited to just a few countries. Elsewhere, democracy analysts believe that the democratization process appears to be regressing. They are even doubtful as to whether the minimal progress being made in a few countries can justifiably be attributed to the impact of democracy assistance programmes.

Development of a common methodology: the experience of the past and lessons for the democracy assistance community

In the past two decades or so, the search for quantitative measures of progress has increased. International agencies, some governments, mainly in the developed world, and citizens have over the years shown a desire for performance in various areas to be measured in terms of simple but quantifiable indicators. Such indicators can be useful to determine policy interventions and in the revision of strategies in the case of development agencies and the recipients of aid. Even the business community has increasingly come to rely on economic, political and corruption indicators to make decisions about whether a country is suitable for their investment. It is in this

context that organizations such as the London-based Economist Intelligence Unit, among others, have come to prosper. Elsewhere, the world has seen the emergence of institutions and organizations, primarily based in the developed countries, which are dedicated to developing global indicators on various global trends. Indicators have clearly gained currency as measures of performance.

One organization which took the lead in the development of indicators early on is Freedom House. Based in the United States, Freedom House has gained prominence as an authority on measuring countries' performance in the area of political freedoms and civil and political rights. It developed its methodology through the expert use of specialists to evaluate political and civil rights in different countries around the world. Some of the experts, according to Freedom House, are based in the individual countries, and every year they file their returns on the country's performance. In short, Freedom House's indexes of political freedoms and civil and political rights are based on theoretical and empirical methods of performance evaluation.

The Freedom House classification of countries has sparked a great deal of criticism and rebuttal both by some countries which consider themselves to have been given too low a rating and by independent scholars who question the rigour of the methodology used. Many have claimed that the Freedom House figures were biased and detached from the reality on the ground, while others have felt that by concentrating on political freedoms and civil liberties Freedom House was missing the fact that not every country recognized these particular freedoms as the central tenets of its political culture. Needless to say, through insistence and regular publications, the Freedom House indexes and classification of countries on the political freedoms scale have gradually gained wide acceptance throughout the wider body of the academic and development assistance literature. Other institutions, such as the World Bank, International IDEA, Polity IV in the UK, the Barometer group, which started in Eastern Europe and has now spread to Africa, Asia and Latin America, Eurostat (the Statistical Office of the European Communities) and so on, are also developing some quantitative measures for assessing democracy and political development in the world. These initiatives have not, however, made as much impact as the Freedom House classification of countries.

Yet another initiative which might offer useful experience in the effort to develop a common project evaluation methodology for democracy support programmes is that of Transparency International (TI). Established as an international NGO in the 1980s, TI found a niche in focusing its work on monitoring corruption around the world. It began with descriptive reports of the state of corruption in individual countries. Like Freedom House, however, it soon realized that a quantitative indicator was much more desirable and attractive to the development assistance community. It was then that TI developed its reporting into the Corruption Index, which has become its flagship project, accounting for both Transparency International's popularity and its notoriety. Since the introduction of the Corruption Index, TI has been able to classify countries annually in terms of the least and most corrupt. The method of 'list and

shame', so strongly opposed by many governments, especially in developing countries, has nevertheless gained currency and wide usage in the literature. Those countries that are favourably ranked even see the score as an indicator of their attractiveness to investment. Those ranked low are, of course, expected to work harder to improve their embarrassing scores. The Corruption Index is therefore a useful tool acting as both deterrent and incentive to governments and the private sector alike.

However, it is also true that, if one were to ask both Freedom House and Transparency International about their experiences developing their indexes, they would tell of the disputes and difficulties they went through before they achieved the recognition they have today. Indeed, their measures have had to be improved over time, reflecting inputs from their critics.

Another important recent experience in developing a global index for classification of countries on the basis of their development performance comes from the UNDP. In 1990 the UNDP introduced what was then the highly controversial Human Development Index (HDI). Tired of the reductionist nature of the traditional gross national product (GNP) as a measure of development, the UNDP brought together a group of leading social scientists under the leadership of the late Maboub ul Haq, former minister of finance of Pakistan, to develop a new index for measuring development rather than just national economic growth. They brought together the element of GNP that is income, measures of health and indicators on education and literacy into the composite index that we now know and which is widely accepted as the best measure of human development/progress. As in the previous cases, the HDI was very controversial at the beginning, and some United Nations member states rejected it out of hand. Today, many take the HDI literally and forget that, like all other aggregate measures, it is an estimate that has its own deficiencies.

This background constitutes the basis and justification for the development of a global index for measuring the impact of democracy assistance beyond project level. Such an index will help the donors/development assistance agencies to do the business of evaluation of democracy assistance differently. The approach has to be different if the results and impact are to be appreciated. The following sections make proposals for the way forward in developing such an index.

Proposals for measures towards the development of a global index for measuring the impact of democracy assistance

The papers prepared for the IDEA/Sida workshop on Methods and Experiences of Evaluating Democracy Support tell a profound story of the issues and lessons from the field of democracy assistance on the ground. As is pointed out above, one emerging issue is that there is no consensus yet on the common tools to be used in evaluating democracy assistance programmes. In fact toolmaking is still a work in progress. Many exciting tools and methods are emerging and being developed in the

field. Some of these have a narrow focus—to empower the project target group to carry out democracy development in their communities and their countries (see e.g. International IDEA 2002; chapter 7 by Harry Blair and chapter 4 by Sandra Elena and Héctor Chayer in this volume). Others are experiments intended to equip donors to use a common tool to evaluate the impact of democracy assistance. Still others have the broader goal of developing generic, general quantitative and quality indexes and assessments measuring and comparing countries' progress in democracy (as is apparent in the case of USAID). Clearly, the goal of democracy assessment initiatives has been defined in different ways that range from the empowerment of citizens to monitor the quality of their country's democracy (International IDEA 2002), the promotion of good governance (as with the United Nations Economic Commission for Africa, for instance), and peer review in support of good governance (as with the New Partnership for Africa's Development, or NEPAD).

There have been several other instances of related work pursued at individual country, regional and international levels, some of it focusing on the development of good governance indicators while others have used opinion surveys to measure democracy/good governance at the country level. Examples can be found at the World Bank, Eurostat, Paris 21, the 'democracy barometers' (Afrobarometer, Latinobarometer, Eurobarometer and East Asia Barometer), and the Lokniti programme in South Asia. Clearly, these different methodologies are still evolving and are in need of better coordination and collaboration. A major area of concern, however, is that the assessment methodologies are not being linked up with the evaluation methodologies of the development assistance agencies. In this section we propose that first we need a common understanding of democracy assistance programmes. Why have so many development agencies, intergovernmental organizations, international, regional and national NGOs and political institutes/foundations all of a sudden devoted so much time, resources and energies to democracy assistance? The answer is simple: all this effort is intended to strengthen and sustain democracy around the world. If this is so, then the major questions must be: What are the tools required to attain this goal? Should those tools be the same for all democracy promoters? If so, how can that be made possible? If the tool or tools are available, will they be able to help individual financiers of democracy programmes to determine the impact of their individual programmes on the wider goal of democracy promotion? In this writer's view, the answers to these questions should form the focus of the debates donors have in the future. If the goal(s) is/are clear, then the assessment methodologies should not pose a great deal of difficulty. Hence the following case study.

The 'State of Democracy'

Beginning in 1999, International IDEA started a project dubbed the State of Democracy by bringing together a group of leading researchers and academics

from the developed and developing countries to design and test a methodology for democracy assessment. The members of the initial team came from Botswana, India, Italy, Kenya, Lesotho, Peru, Poland, Russia, Spain and the UK, among other countries. The aims of the State of Democracy project were as follows:

• to development a comprehensive methodology for democracy assessment;
• to use the methodology as a tool that governments, citizens and democracy support agencies can use to evaluate the progress of democracy in their countries or areas of democracy assistance; and
• to use the methodology to generate a regular publication on the state of democracy around the world, thereby sharing the lessons of democracy development around the world.

The experts, led by a team of academics based at the University of Leeds in the UK and coordinated and financed by International IDEA, did indeed design and test a comprehensive methodology on the state of democracy. The methodology started by defining what democracy is, and outlined the components of democracy, including the basic freedoms—basic civil, political and cultural rights; the institutional framework of democracy such as the constitution; the rule of law; political institutions, including parliament, political parties and local government; the media; civil society; and the role of external players or what was called 'democracy beyond the state'. Each section, for instance, on the rule of law, was followed by a set of searching questions on the basis of which they could determine whether the rule of law was entrenched in a particular country. The topic could be different—for example, the 'role of civil society'—but again the searching questions were designed to help the assessor(s) to determine whether in a particular country there was a recognition of the role civil society could play in promoting democracy.

This methodology was then tested in the field over a period of two years. Country case studies were conducted in Bangladesh and South Korea, Kenya and Malawi, Peru and El Salvador, and New Zealand and Italy. Relatively independent studies using elements of the same methodology were also conducted in the UK, Sweden, Australia and Canada. Most of the assessors were joint teams of academics and civil society leaders. The assessment involved some desk studies and discussions with key players in the democracy field in each country, including the speaker of the parliament, political party leaders, the head of the EMB, judges, independent scholars, leaders of civil society and the media, and the donor representatives within the country. On the basis of the desk research and interviews the assessors were able to produce a comprehensive report. The requirements for the pilot assessment in each country were that:

• each report should be subjected to review by independent reviewers based within the country and knowledgeable on the subject matter; and

- an in-country workshop of a cross-section of stakeholders should be held to both validate the report and increase awareness of the assessment tool.

The pilot reports from all the countries in the sample were published separately and in a summarized publication by International IDEA (2003). Clearly, some common lessons emerged, but some shortcomings of this tool were also identified.

The lessons were that democracy was indeed stalling in many countries and that, whereas many countries had adopted the institutions of democracy, the practice of democracy was poor due to a number of factors. The institutions were weak, political awareness was low, mobilization efforts were not sustained and civil society leaders were co-opted into government, while the political opposition was weak and poorly resourced (International IDEA 2003). Basic freedoms and the independence of the EMB and the judiciary were still among the major issues of democracy in such countries such as Kenya, Peru, Bangladesh and, interestingly, even Italy.

On the shortcomings of the State of Democracy methodology itself, several issues emerged.

- The objectivity of the assessors became an issue: some were known critics of the state concerned.
- The political standing of assessors was questionable, that is, they were not sufficiently influential people in the country.
- The selection of evidence to support the arguments was seen as subjective.
- The data were often outdated and in some areas inaccessible.
- The assessment was seen as too detailed and cumbersome.
- The report looked more like the output of academic research than an advocacy tool that civil society and the media could use to enhance their democracy advocacy work within a country.
- The fact that the assessment was qualitative and did not yield a composite measure of the state of democracy made it less appealing to donors and civil society alike.

The State of Democracy methodology, which was then published in French, Russian, Spanish and Arabic for wider consumption, has not gained the same measure of popularity (or controversy) that has been generated by the Freedom House, Transparency International and UNDP tools described above. There are two reasons for this. First, International IDEA was always a reluctant player in designing and promoting this methodology. No annual reports were published. Second, the absence of an index which would make it possible to classify countries into those that are strongly democratic and those that are only weakly democratic proved a major drawback.

The way forward

The initiation of a global democracy assessment methodology by International IDEA has been a step in the right direction. What is needed now is a global democracy assessment/evaluation methodology that is a tool that citizens can use to assess democracy progress in their country and at the same time will serve as an evaluation tool which donors and beneficiaries can use to evaluate the impact of their projects on the political system. Such a tool needs to be simple and widely accessible. It should be both quantitative and qualitative. The way forward that this writer proposes is for International IDEA to develop the State of Democracy methodology into a quantifiable tool whereby scores will be allocated to performance in the rule of law, freedom of speech, media freedom and so on, leading to a composite index that will then be calculated along lines similar to the TI Corruption Index, the UNDP's HDI and Freedom House's indexes.

As we learned above, it is rather optimistic to expect development agencies to coordinate among themselves and adopt a common tool. One brave organization must emerge and drive the process. It is the credibility of the tool that will popularize it, rather than political consensus from the players who must a priori decide that they need such a tool. International IDEA is well placed and has made a start on a potentially useful evaluation tool. The State of Democracy methodology is more comprehensive and focused on participation by democracy supporters and promoters. It has the potential to surpass several of the existing tools in this area—in fact some NGOs in South Africa, for example, the Institute for Democracy in South Africa (IDASA), have began the process of allocating numbers to an International IDEA-type methodology.

References and further reading

Åbo Human Rights Institute, Report of the Turku Expert Meeting on Indicators, 10–13 March 2005

Alcantara, Anthony O., 'Globe Telecom Says Capacity to Hit 4.5 m.', *Philippine Daily Inquirer*, 28 November 2000 Amnesty International, Irish Section, 'Our Rights, Our Future: Human Rights Based Approach in Ireland Principles, Policy and Practice', www.ihrnetwork.org

Åslund, Anders and McFaul, Michael (eds), *Revolution in Orange: Origins of Ukraine's Democratic Breakthrough* (Washington, DC: Carnegie Endowment for International Peace, 2006)

Avritzer, Leonardo, *Democracy and the Public Space in Latin America* (Princeton, N.J.: Princeton University Press, 2002)

Axelrod, Robert, *Från konflikt till samverkan* [The evolution of cooperation] (Stockholm: SNS förlag, 1987)

Barnes, C. and Abdullaev, K., 'Introduction: From War to Politics', in K. Abdullaev and C. Barnes (eds), *Politics of Compromise: The Tajikistan Peace Process*, Accord series, Issue 10, Conciliation Resources (2001), www.c-r.org/our-work/accord/tajikistan/introduction.php

Bemelmans-Videc, Marie-Louise, Rist, Ray C. and Vedung, Evert (eds), *Carrots, Sticks and Sermons: Policy Instruments and Their Evaluation* (New Brunswick, N.J: Transaction Publishers, 1998)

Bigdon, Christine and Korf, Benedikt, 'The Role of Development Aid in Conflict Transformation: Facilitating the Empowerment Processes and Community Building', in A. Austin, M. Fischer and N. Ropers (eds), *Transforming Ethno-political Conflict: The Berghof Handbook* (Berlin: Berghof Research Center for Constructive Conflict Management, 2004)

Bjornlund, Eric C., *Beyond Free and Fair. Monitoring Elections and Building Democracy* (Washington, DC: Woodrow Wilson Press Center, and Baltimore and London: Johns Hopkins University Press, 2004)

Blair, Harry, 'USAID and Democratic Decentralization: Taking the Measure of an Assistance Programme, in Peter Burnell (ed.), *Democracy Assistance: International Cooperation for Democratization* (London: Frank Cass, 2000), pp. 226–41

— 'Civil Society Strategies Assessment in the ANE Region: The Philippines', Occasional Papers Series (Washington, DC: USAID, Governance and Democracy Center, 2001)

— 'Research and Practice in Democratization: Cross-Fertilization or Cross Purposes?', in Edward R. McMahon and Thomas A. P. Sinclair (eds), *Democratic Institution Performance: Research and Policy Perspectives* (Westport, Conn.: Praeger, 2002), pp. 175–92

— 'Jump-Starting Democracy: Adult Civic Education and Democratic Participation

in Three Countries', *Democratization*, 10/1 (2003), pp. 53–76

— 'Assessing Civil Society Impact for Democracy Programmes: Using an Advocacy Scale in Indonesia and the Philippines', *Democratization*, 11/1 (2004), pp. 77–103

Bob, Clifford, *The Marketing of Rebellion* (Cambridge: Cambridge University Press, 2005)

Boix, Carles, 'Policy analysis y gestión de servicios', Paper presented at 'Jornadas sobre la modernización de las administraciones públicas', Madrid, 1992

Bollen, Kenneth, Paxton, Pamela and Morishima, Rumi, 'Research Design to Evaluate the Impact of USAID Democracy and Governance Programs', Report prepared for USAID in September 2003, commissioned by the Social Science Research Council under John Tirman

Bossuyt, Jean et al., 'Thematic Evaluation of the EC Support to Good Governance: Evaluation for the European Commission, Final Report', Contract no. EVA/80-208 (Brussels, June 2006)

British Department for International Development (DFID), *Appraisal of the Shire Highlands Sustainable Livelihood Programme Oxfam, Malawi*, Report by Hanne Lund Madsen (London: DFID, 2001)

Brodeur, Jean et al., *Five Year Review of the Organization and its Activities of the International Centre for Human Rights and Democratic Development* (Montreal: Gested International, 1993)

Brumberg, Daniel, *Liberalization vs Democracy: Understanding Arab Political Reform*, Carnegie Endowment for International Peace Working Papers no. 37 (Washington, DC: Carnegie Endowment for International Peace, May 2003), <http://www.ceip.org>

Building Unity for Continuing Coconut Industry Reform (BUCO), 'Clippings', 15 July 1998

Burnell, Peter, 'Political Strategies of External Support for Democratization', *Foreign Policy Analysis*, 1/3 (2005), pp. 361–84

— *Promoting Democracy Backwards*, Working Paper no. 28 (Madrid: Fundación para las Relaciones Internationales y el Diálogo Exterior, 2006)

— (ed.), *Globalising Democracy: Party Politics in Emerging Democracies* (Milton Park and New York: Routledge, 2006)

— 'From Evaluating Democracy Assistance to Appraising Democracy Promotion', *Political Studies*, forthcoming 2007/8

Buscaglia, Edgardo and Dakolias, María, *Comparative International Study of Court Performance Indicators: A Descriptive and Analytical Account* (Washington, DC: World Bank, Legal Department, Legal and Judicial Reform Unit, 1999)

Business and Human Rights Project, <http://www.humanrights.dk>

Calumpita, Ronnie E., 'Farmers Deny Compromise Agreement on Coco Levy Fund', *Manila Times*, 8 February 2006

Canada, International Centre for Human Rights and Democratic Development Act 1988, http://laws.justice.gc.ca/en/I-17.3/text.html

Canadian Office of the Inspector General (OIG), *Five Year Review of Rights & Democracy* (1998–2003) (Ottawa: OIG, 2003)

Carothers, Thomas, *Assessing Democracy Assistance*: The Case of Romania (Washington, DC: Carnegie Endowment for International Peace, 1996)

— *Aiding Democracy Abroad: The Learning Curve* (Washington, DC: Carnegie Endowment for International Peace, 1999)

— 'The End of the Transition Paradigm', *Journal of Democracy*, 13/1 (2002), pp. 1–21

— *Essays on Democracy Promotion* (Washington, DC: Carnegie Endowment for International Peace, 2004)

— 'The Backlash against Democracy Promotion', *Foreign Affairs*, 85/2 (March/April 2006), <http://www.carnegieendowment.org> (2006a)

— *Confronting the Weakest Link: Aiding Political Parties in New Democracies* (Washington, DC: Carnegie Endowment for International Peace, 2006) (2006b)

— *Promoting the Rule of Law Abroad: In Search of Knowledge* (Washington, DC: Carnegie Endowment for International Peace, 2006) (2006c)

— and Ottaway, Marina (eds), *Uncharted Journey: Promoting Democracy in the Middle East* (Washington, DC: Carnegie Endowment for International Peace, 2005)

Chandrasekaran, Rajiv, 'Philippine Activism, at Push of a Button', *Washington Post*, 10 December 2000

— 'Philippine Government Collapses', *Washington Post*, 20 January 2001

Church, C. and Shouldice, J., *The Evaluation of Conflict Resolution Interventions. Part II: Emerging Practice and Theory*, INCORE International Conflict Research (Londonderry: University of Ulster and United Nations University, 2003)

Ciurlizza, Javier and Acosta, Gladys, *Democracy in Peru: A Human Rights Perspective* (Montreal: Rights & Democracy, 1997), <http://www.dd-rd.ca>

'The Civicus Civil Society Index', http://www.civicus.org

Colletta, N. J. and Cullen, M. L., *Violent Conflict and the Transformation of Social Capital: Lessons from Cambodia, Rwanda, Guatemala, and Somalia*, Social Capital Working Paper Series (Washington, DC: World Bank, Social Development Department, June 2003)

Coomans, F. and Kanninga, M. (eds), *Extraterritorial Application of Human Rights Treaties* (Antwerp: Intersentia, 2004)

Crawford, Gordon, 'Promoting Democracy from Without: Learning from Within (Part I)', *Democratization*, 10/1 (2003), pp. 77–98 (2003a)

— 'Promoting Democracy from Without: Learning from Within (Part II)', *Democratization*, 10/2 (2003), pp. 1–20 (2003b)

Dahl, Robert A., *On Democracy* (New Haven, Conn.: Yale University Press, 1998)

Dahl-Østergaard, Tom, *Lessons Learned on the Use of Power and Drivers of Change*

Analyses in Development Co-operation, Review Commissioned by the Organisation for Economic Co-operation and Development (OECD) Development Assistance Committee (DAC) Network on Governance (Paris: GOVNET, 2005)

Danish International Development Agency (Danida), *Danish Support to Promotion of Human Rights and Democratization. Evaluation, Vol. 1: Synthesis Report* (Copenhagen: Ministry of Foreign Affairs, 1999) (1999a)

— *Danish Support to Promotion of Human Rights and Democratization. Evaluation, Vol. 3: Elections* (Copenhagen: Ministry of Foreign Affairs, 1999) (1999b)

— *Danish Support to Promotion of Human Rights and Democratization. Evaluation Vol. 5: Empowerment and Participation* (Copenhagen: Ministry of Foreign Affairs, 1999 (1999c)

— *The Danish NGO Impact Study: A Review of Danish NGO Activities in Developing Countries, Overview Report* (Copenhagen: Ministry of Foreign Affairs, 1999 (1999d)

Danish Ministry of Foreign Affairs, Evaluation Department, *Peer Assessment of Evaluation in Multilateral Organizations: United Nations Development Programme*, by Mary Cole et al. (Copenhagen: Ministry of Foreign Affairs of Denmark, 16 December 2005) (January 2006)

D'Souza, Dilip, *The Narmada Dammed: An Inquiry into the Politics of Development* (New Delhi: Penguin Books, 2002)

Dwivedi, Ranjit, 'Resisting Dams and "Development": Contemporary Significance of the Campaign against the Narmada Projects in India', *European Journal of Development Research*, 10/2 (1998), pp. 135–83

Eade, Deborah (ed.), *Development and Advocacy: Selected Essays from Development and Practice* (Oxford: Oxfam, 2002)

Ehmann, Claire, Morriss, Sharon and Alimkulova, Mahabat, 'Mid-term Evaluation of the Central Asian Community Action Investment Program, MSDSP', Dushanbe, October 2003 (internal report of the Aga Khan Foundation)

Erdmann, Gero, 'Hesitant Bedfellows: The German *Stiftungen* and Party Aid in Africa', in Peter Burnell (ed.), *Globalising Democracy: Party Politics in Emerging Democracies* (Milton Park and New York: Routledge, 2006), pp. 181–99

European Centre for Development Policy Management, 'Institutional Evaluation of the Netherlands Institute for Multiparty Democracy, Final Report', Maastricht, December 2005

Finkel, Steven, 'Can Democracy be Taught?', *Journal of Democracy*, 14/4 (2003), pp. 137–51

— et al., 'Effects of US Foreign Assistance on Democracy Building: Results of a Cross-National Quantitative Study: Final Report', USAID, Vanderbilt University and University of Pittsburg, 2006, available on the USAID website at http://usaid.gov

Fisher, William F. (ed.), *Toward Sustainable Development: Struggling Over India's*

Narmada River (Armonk, NY: M. E. Sharpe, 1995)

Foro de Estudios sobre la Administración de Justicia, www.foresjusticia.org.ar

Forss, Kim, *Finding Out about Results from Projects and Programmes Concerning Democratic Governance and Human Rights: A Study Commissioned by Sida* (Stockholm: Swedish International Development Cooperation Agency (Sida), September 2002)

Friends of River Narmada, http://www.narmada.org

Fuller, Stephan and Mirimanova, Natalia, 'Final Evaluation Report, Community Action Investment Program', for the Mountain Societies Development Support Programme, Dushanbe, 2005

Gallie, Walter Bryce, 'Essentially Contested Concepts', *Proceedings of the Aristotelian Society,* 56 (1956), pp. 167–98

Gandhi, Ajay, 'Developing Compliance and Resistance: The State, Transnational Social Movements and Tribal Peoples Contesting India's Narmada Project', *Global Networks*, 3/4 (2003), pp. 481–95

Gaventa, John, *Triumph, Deficit or Contestation? Deepening the 'Deepening Democracy' Debate*, Working Paper no. 264 (London: Overseas Development Institute (ODI), 2006)

German Development Institute (GDI), *Evaluation of EC Positive Measures in Favour of Human Rights and Democracy (1991–1993)* (Berlin: GDI, 1995)

Gillies, David, *Human Rights and Democratic Governance: A Framework for Analysis and Donor Action* (Montreal: Rights & Democracy, 1993)

— and Makau wa Mutua, *A Long Road to Uhuru: Human Rights and Political Participation in Kenya* (Montreal: Rights & Democracy, 1993), www.dd-rd.ca

Golub, Stephen, *Beyond Rule of Law Orthodoxy: The Legal Empowerment Alternative*, Working Paper no. 41 (Washington, DC: Carnegie Endowment, 2003)

Gomez, Jim, 'Efforts to Oust Arroyo Likely to Persist', *Washington Post*, 26 February 2006

Gready, Paul, *Reinventing Development? Translating Rights-based Approaches from Theory into Practice* (London: Zed Books, 2005)

Green, Andrew and Kohl, Richard, 'Challenges of Evaluating Democracy Assistance: Perspectives from the Donor Side', *Democratization*, 14/1 (2007), pp. 151–65

Gregorio-Mendel, Angelita, 'Coalition Assessment, USAID Civil Society Programme: Building Unity for Continuing Coconut Industry Reform (BUCO) Programme', Manila, Urban Integrated Consultants, Inc., 1998

Guha, Ramachandra, 'The Arun Shourie of the Left', *The Hindu*, 26 November 2000 (2000a), www.hindu.com/2000/11/26/stories/13260411.htm

— 'Perils of Extremism', *The Hindu*, 17 December 2000 (2000b), www.hindu.com/2000/12/17/stories/1317061b.htm

Gupta, Rajiv A., 'River Basin Management: A Case Study of Narmada Valley Development with Special Reference to the Sardar Sarovar Project in Gujarat,

India', *Water Resources Development*, 17/1 (2001), pp. 55–78

Haarhuis, Carolien. Klein and Leeuw, Frans L., 'Fighting Governmental Corruption: The New World Bank Programme Evaluated', *Journal of International Development*, 16/4 (2004), pp. 547–61

Halfani, M. and Nzomo, M., *Towards a Reconstruction of State–Society Relations: Democracy and Human Rights in Tanzania* (Montreal: Rights & Democracy, 1995), www.dd-rd.ca

Hammergren, Linn, 'Assessments, Monitoring, Evaluation, and Research: Improving the Knowledge Base for Judicial Reform Programs', 2002, www.pogar.org/publications/judiciary/linn1/evaluation.html#foot10

Harding, Luke, 'Roy Goes to Prison and Agonises over Fine or Serving Longer Term: India Jails Booker Winner for One Day', *The Guardian* (London), 7 March 2002, website.lineone.net/~jon.simmons/roy/020306c.htm

Horton, Douglas et al., 'Evaluation, Learning and Change in Research and Development Organizations: Concepts, Experieces, and Implications for the CGIAR', International Service for National Agriculture Research Discussion Paper no. 03-2 (February 2003), www.isnar.org

Humanitarian Accountability Project (HAP), 'Overview Report', 2005, www.hapinternational.org

Hyman, Gerald, Keynote speech delivered at the Democracy Partners Conference, December 2002

International Development Research Centre (IDRC), *Evaluating Governance Programmes: Report of a Workshop* (IDRC: Ottawa, 1999), www.idrc.ca

International IDEA, *Handbook on Democracy Assessment* (The Hague: Kluwer Law International, 2002)

— *The State of Democracy: Democracy Assessments in Eight Nations Around the World* (The Hague: Kluwer Law International, 2003)

Jarstad, Anna, *International Assistance to Democratization in Bosnia and Herzegovina*, Kosovo and Macedonia: Synthesis Report (Uppsala: Uppsala University, Department of Peace and Conflict Research, 2005)

Jilani, Hina, *Human Rights and Democratic Development in Pakistan* (Montreal: Rights & Democracy, 1998), <http://www.dd-rd.ca>

Johns Hopkins University, Center for Civil Society Studies, www.jhu.edu/~cnp/

Jonsson, Urban, 'Human Rights Based Approach to Development Programming', UNICEF, Esaro, 2003, <http://www.unicef.org>

Kapoor, Ilan, *Indicators for Programming in Human Rights and Democratic Development: A Preliminary Study* (Gatineau, Quebec: Canadian International Development Agency, 1996), www.acdi-cida.gc.ca

— *Background Paper and Literature Review*, for the IDRC Evaluation Unit (Ottawa: International Development Research Centre (IDRC), March 1999), www.idrc.ca/fr/ev-10110-201-1-DO_TOPIC.html#ad

Kelley, Judith, *Ethnic Politics in Europe: The Power of Norms and Incentives* (Princeton, N.J: Princeton University Press, 2004)

van der Knaap, Peter, 'Theory-based Evaluation and Learning: Possibilities and Challenges', *Evaluation*, 10/1 (2004), pp. 16–34

Kumar, Krishna, *Promoting Independent Media* (Boulder, Colo.: Lynne Rienner, 2006)

Lean, Sharon F., 'Democracy Assistance to Domestic Election Monitoring Organizations: Conditions for Success', *Democratization*, 14/2 (2007), pp. 289–312

Leeuw, Frans L., 'Reconstructing Programme Theories and Problems to be Solved', *American Journal of Evaluation*, 24/1 (2003), pp. 5–20

Linz, Juan J., and Stepan, Alfred, *Problems of Democratic Transition and Consolidation: Southern Europe, South America, and Post-Communist Europe* (Baltimore, Md.: Johns Hopkins University Press, 1996)

Madsen, Hanne Lund, 'Evaluering af bistand til menneskerettigheder' [Evaluation of assistance to human rights], *Den Ny Verden* [The new world], Vol. 1 (Copenhagen: Centre for Development Research, 1998)

— 'Assessing the Impact of Human Rights Work: Key Elements for a Methodology', Paper prepared for the Novib Workshop on Measuring Impacts of Human Rights Interventions, The Hague, May 2001 (unpublished)

— 'Characteristics of Human Rights Indicators', Paper presented at the Danida seminar on Human Rights, Democratization and Decentralisation, November 2003 (unpublished)

— *Consultative Review of FIAN International* (Stockholm: Swedish International Development Cooperation Agency (Sida), 2004)

Magno, Leo, 'The Inquirer, 15 Years Later', *Philippine Daily Inquirer*, 28 January 2001

Mangahas, Mahar, 'From Juetenggate to People Power 2: The SWS Surveys of Public Opinion', Presentation at Quezon City, 16 February 2001, <http://www.sws.org.ph>

Mathieson, David and Youngs, Richard, *Democracy Promotion and the European Left: Ambivalence Confused?,* Working Paper no. 29 (Madrid: Fundación Para Las Relaciones Internationales y el Diálogo Exterior, 2006)

Matute, Eileen, 'The Coco Levy Controversy: A Fight against Poverty, a Struggle for Justice', *Community and Habitat* (published by the Philippine Rural Reconstruction Movement), Issue 9 (2001), pp. 90–101

Mokhiber, Craig G., 'Toward a Measure of Dignity: Indicators for Rights-Based Development', *Statistical Journal of the United Nations Economic Commission for Europe*, 18/2–3 (2001), pp. 125–283

Møller, Lars and Jackson, Jack, 'Journalistic Legwork that Tumbled a President: A Case Study and Guide for Investigative Journalists', WBI Working Paper, Report

no. 28734, World Bank Institute, 2002

Molund, Stefan and Schill, Göran, *Looking Back, Moving Forward* (Stockholm: Swedish International Development Cooperation Agency (Sida), 2004)

Mydans, Seth, 'Political Turmoil Again Thwarts Progress in Philippines', *New York Times*, 26 February 2006 (2006a)

— 'For Philippine Military, Politics Remains a Crucial Mission', *New York Times*, 5 March 2006 (2006b)

La Nación (Buenos Aires), 16 December 2005 and 20 March 2006

Naciri, Rabia et al., *Développement démocratique et action associative au Maroc* (Montreal: Rights & Democracy, 2004), <http://www.dd-rd.ca>

Narayan, Deepa, 'Conceptual Framework and Methodological Challlenges', in Deepa Narayan (ed.), *Measuring Empowerment: Cross-Disciplinary Perspective*s (Washington, DC: World Bank, 2005), pp. 3–38

Netherlands Institute for Human Rights, *From Development of Human Rights to Managing Human Rights Development: Global Review of the OHCHR Technical Cooperation Programme, Synthesis Report* (The Hague: Netherlands Institute for Human Rights, 2003)

Netherlands Institute for Multiparty Democracy, *No Lasting Peace and Prosperity without Democracy and Human Rights: Report Commissioned by the European Parliament* (The Hague: Netherlands Institute for Multiparty Democracy, July 2005)

Norwegian Parliament, *Om norsk landbruk och matproduksjon: tilrådning fra Landbruksdepartementet* [White paper on agriculture and food production] (St. meld. 19, 1999/2000), 17 December 1999, <http://www.regjeringen.no/en/ministries/lmd/Documents/regpubl/stmeld/19992000/Stmeld-nr-19-1999-2000-.html?id=192695>

O'Brien, Paul and Jones, Andrew, *Human Rights and Rights Based Programming Training Manual* (Nairobi: Care, 2002)

Omvedt, Gail, 'An Open Letter to Arunadhati Roy', June 1999, reprinted by Friends of River Narmada, <http://www.narmada.org/debates/gail/gail.open.letter.html> (1999a)

— 'Dams and Bombs', *The Hindu*, 4 and 5 August 1999, reprinted by Friends of River Narmada, <http://www.narmada.org/archive> (1999b)

Organisation for Economic Co-operation and Development (OECD), *Evaluation of Programs Promoting Participatory Development and Good Governance: Synthesis Report* (Paris: OECD, 1997)

Ottaway, Marina, 'Social Movements, Professionalism of Reform, and Democracy in Africa', in Marina Ottaway and Thomas Carothers (eds), *Funding Virtue: Civil Society Aid and Democracy Promotion* (Washington, DC: Carnegie Endowment for International Peace, 2000), pp. 77–104

— and Carothers, Thomas, *Funding Virtue: Civil Society Aid and Democracy Promotion*

(Washington, DC: Carnegie Endowment for International Peace, 2000)

Owen, John M. and Rogers, Patricia J., *Program Evaluation: Forms and Approaches* (London: Sage, 1999)

Palencia Prado, Tania and Holiday, D., *Towards a New Role for Civil Society in the Democratization of Guatemala* (Montreal: Rights & Democracy, 1996), <http://www.dd-rd.ca>

Parreño, Earl G. and Gaborni, Joel, 'Danding Gojuangco: The "Pacman" Returns', 4 parts, *Political Brief* 7, 9 (Manila: Institute for Popular Democracy, September 1999)

Pawson, Ray, 'Evidence-based Policy: The Promise of a "Realist Synthesis"', *Evaluation*, 8/3 (2002), pp. 340–58

Pazzibugan, Dona Z., Batino, Clarissa S. and Torrijos, Elena R., 'Coco Levy Ruling Just a Key Battle Won, Says Ombudsman', Inquirer News Service, 13 July 2003, <http://www.inq7.net/nat/2003/jul/13/text/nat_1-1-p.htm>

Piccone, Ted and Youngs, Richard (eds), *Strategies for Democratic Change: Assessing the Global Response* (Washington, DC: Democracy Coalition Project, and Madrid: Fundación para las Relaciones Internacionales y el Diálogo Exterior, 2006)

Poate, Derek et al., *The Evaluability of Democracy and Human Rights Projects* (Stockholm: Swedish International Development Cooperation Agency (Sida), 2000)

Pressman, Jeffrey and Wildavsky, Aaron, *Implementation*, 2nd edn (Berkeley, Calif.: University of California Press, 1979)

Putnam, Robert, *Making Democracy Work* (Princeton, N.J.: Princeton University Press, 1993)

Rebien, C. C., *Evaluating Development Assistance in Theory and Practice* (Aldershot: Avebury, 1996)

Reid, Ben, 'The Philippine Democratic Uprising and the Contradictions of Neoliberalism: EDSA II', *Third World Quarterly*, 22/5 (2001), pp. 777–93

Reygadas Robles Gil, Rafael and Soto Martínez, Maricela Adriana (eds), *Self-Made Citizens: Building Democracy Through Human Rights in Mexico* (Montreal: Rights & Democracy, 2003), <http://www.dd-rd.ca>

Rivas, E. and Gonzáles-Suárez, M., *Obstacles and Hopes: Perspectives for Democratic Development in El Salvador* (Montreal: Rights & Democracy, 1994), <http://www.dd-rd.ca>

River Path Associates, 'Review of the Westminster Foundation for Democracy, Final Report', 14 January 2005, <http://www.fco.gov.uk>

Roche, C., *Impact Assessment and NGOs: Learning for a Change?* (Oxford: Oxfam 1999)

Routledge, P., 'Voices of the Dammed: Discursive Resistance Amidst Erasure in the Narmada Valley, India', *Political Geography*, 22 (2003), pp. 243–70

Roy, Arundhati, 'The Greater Common Good', April 1999, reprinted by Friends of

River Narmada, <http://www.narmada.org/gcg/gcg.html>

Save the Children UK, *Global Impact Monitoring: Save the Children UK's Experience of Impact Assessment* (London: Save the Children, 2004)

— *In the Right Direction. Examples of the Impact of Save the Children's Work 2005: A Synthesis Report from Global Impact Monitoring Reports 2005* (London: Save the Children UK, 2005)

Schmitter, Philippe C. and Brouwer, Imco, *Conceptualizing, Researching, and Evaluating Democracy Promotion and Protection*, European University Institute (EUI) Working Paper SPS no. 99.9 (Florence: European University Institute, 1999)

Schmitter, Philippe C. and Karl, Terry Lynn, 'What Democracy Is … and Is Not', *Journal of Democracy*, 2/3 (1991), pp. 74–88

Schraeder, Peter J. (ed.), *Exporting Democracy Rhetoric vs Reality* (Boulder, Colo.: Lynne Rienner, 2002)

Schumpeter, Joseph, *Capitalism, Socialism and Democracy* (New York: Harper & Brothers, 1942)

Scott, James and Steele, Carie A., 'Assisting Democrats or Resisting Dictators? The Nature and Impact of Democracy Support by the United States National Endowment for Democracy, 1990–99', *Democratization*, 12/4 (2005), pp. 439–60

Spendolini, Michael J., *The Benchmarking Book* (New York: AMACOM, 1992), translated into Spanish by Carlos Villa (Bogota: Grupo Editorial Norma, 1992)

Spuches, M. G., 'Inventaire et classification des projets en développement démocratique, 1991–2001' [Inventory and classification of democratic develoment projects, 1991–2001], Working document, Rights & Democracy, Montreal, 2000

Stiglitz, Joseph, 'Participation and Development: Perspectives from the Comprehensive Development Paradigm', *Review of Development Economics*, 6/2 (2002), pp. 163–82

Swedish Agency for Development Cooperation (SADEV), 'Policy Brief: How to Trace the Results of Democracy Support', November 2006, <http://www.sadev.se>

Swedish International Development Cooperation Agency (Sida), 'The Challenge of Evaluating Support for Democracy and Human Rights', *SIDA Evaluations Newsletter*, 2/00 (2000)

—, Department for Cooperation with Non-Governmental Organisations and Humanitarian Assistance and Conflict Management, 'Sida's Policy for Civil Society', Stockholm, April 2004

Taylor, Charles and Muntarbhorn, Vitit, *Roads to Democracy: Human Rights and Democratic Development in Thailand* (Montreal: Rights & Democracy, 1994), <http://www.dd-rd.ca>

Thede, Nancy, 'Some Reflections on the No-Man's-Land Between Concept and Indicator', *Statistical Journal of the United Nations Economic Commission for Europe*, 18/2–3 (2001), pp. 259–73, <http://www.dd-rd.ca>

— *Democratic Development 1990–2000: An Overview* (Montreal: Rights &

Democracy, April 2002), <http://www.dd-rd.ca>

— et al., *The Democratic Development Exercise* (Montreal: Rights & Democracy, July 1996), <http://www.dd-rd.ca>

Theis, Joachim, 'Promoting Rights-based Approaches: Experiences and Ideas from Asia and the Pacific', Save the Children Sweden, Bangkok, 2004

UK Evaluation Society, 'Guidelines for Good Practice in Evaluations', 2003, <http://www.evaluation.org.uk>

United Nations, Economic and Social Council, *Report of the Special Rapporteur on the Right of Everyone to the Enjoyment of the Highest Attainable Standard of Physical and Mental Health*, by Paul Hunt, UN document E/CN.4/2006/48, 3 March 2006

United Nations, General Assembly, *In Larger Freedom: Towards Security, Development, Freedom and Human Rights for All*, UN document A/59/2005, 21 March 2005

United Nations, Office of the High Commissioner for Human Rights (OHCHR), *Draft Guidelines on a Rights Based Approach to Poverty Reduction Strategies* (Geneva: OHCHR, 2003)

— *Frequently Asked Questions About the Human Rights Based Approach* (Geneva: OHCHR, 2005)

— *What Is a Rights Based Approach to Development?*, no date, <http://www.unhchr.ch/development/approaches-02.html>

United Nations Children's Fund (UNICEF), *Consolidation and Review of the Main Findings and Lessons Learned of the Case Studies on Operationalising HRBAP in UNICEF* (New York: UNICEF, 2004)

United Nations Development Group (UNDG), *Common Country Assessment and United Nations Development Assistance Framework: Guidelines for UN Country Team Preparing CCA and UNDAF in 2004* (New York: UNDG, 2003)

United Nations Development Programme (UNDP), *Evaluation of the Human Rights Strengthening Project (Hurist)* (New York: UNDP, 2004)

— *Indicators for Human Rights Based Approaches to Development in UNDP Programming: A Users' Guide* (Washington, DC: UNDP, 2006)

— *Governance Indicators: A User's Guide* (Oslo: UNDP, no date), <http://www.undp.org/oslocentre/cross.htm>

United States Agency for International Development (USAID), *A.I.D.'s Experience with Democratic Initiatives: A Review of Regional Programs in Legal Institution Building*, A.I.D. Program Evaluation Discussion Paper no. 29 (Washington, DC: USAID, 1990)

— 'Measuring the Impact of Democracy and Governance Assistance: Summary Report of a USAID/Clingendael Institute Workshop, The Hague, March 11, 2005', 28 April 2005, available on the USAID website at <http://www.usaid.gov>

Universalia, *Five-Year Review of the International Centre for Human Rights and Democratic Development: 1993–1998* (Montreal: Universalia, 1998)

Vachudova, Milada, *Europe Undivided: Democracy, Leverage and Integration after Communism* (Oxford: Oxford University Press, 2005)

Windfuhr, Michael (ed.), *Beyond the Nation State: Human Rights in Times of Globalization* (Heidelberg: FoodFirst Information and Action Network (FIAN), 2005)

Wood, John R., 'India's Narmada River Dams: Sardar Sarovar Under Siege', *Asian Survey*, 33/10 (1993), pp. 968–84

World Bank, Operations Evaluation Department, 'Pakistan: The Aga Khan Rural Support Program. A Third Evaluation', Report no. 15157-PAK, 11 December 1995,

Yates, Emma, 'Booker-winner Arundhati Roy Jailed', *Guardian Unlimited*, 6 March 2002, <http://website.lineone.net/~jon.simmons/roy/020306a.htm>

Youngs, Richard (ed.), *Survey of European Democracy Promotion Policies 2000–2006* (Madrid: Fundación para las Relaciones Internacionales y el Diálogo Exterior, 2006)

de Zeeuw, J. and Kumar, K., *Promoting Democracy in Postconflict Societies* (Boulder, Colo.: Lynne Rienner, 2006)

About the authors

Harry Blair is presently Associate Chair, Senior Research Scholar and Lecturer in Political Science at Yale University. Previously he held professorial posts at Bucknell, Colgate, Cornell and Rutgers universities, as well as serving on several extended tours at the United States Agency for International Development (USAID). After focusing for many years on rural development and environmental concerns mainly in South Asia, more recently he has concentrated on democratization issues, principally civil society and decentralization in South-East Asia, Latin America and South-East Europe. His most recent articles have dealt with civil society, the environment, post-conflict and rural development, all in relation to governance issues.

Peter Burnell is a Professor of Politics in the Department of Politics and International Studies, University of Warwick, England.

Héctor Chayer has been Executive Director of the Forum for Studies on Judicial Administration (Foro de Estudios sobre la Administración de Justicia, FORES), a non-governmental organization (NGO) based in Buenos Aires, Argentina, since 2000. He is a legal expert with many years of experience in the judicial reform sector, particularly in the development of public policy in the areas of information technology and court administration, including field and desk research. He has also worked on the development, implementation and assessment of judicial reform projects in Argentina, Bolivia, Peru, Nicaragua and Uruguay, and has served as a consultant for the World Bank, Checchi/USAID and IFES (formerly the International Foundation for Election Systems), among others. He has given numerous lectures and published several papers on the subject of judicial reform, including the National Plan for Judicial Reform in Argentina, as well as over 25 articles on technology, court administration and judicial reform.

Sandra Elena is Coordinator of International Programmes at FORES. She is a lawyer and a political scientist, earning an Ll.M at the American University in Washington, DC, in 2002. She has been involved in the rule of law, human rights, civil society management, anti-corruption and political and legislative reform fields since 1990. Her areas of expertise include field and desk research, draft legislation, civil society coalition building, project management, contract negotiation, and programme design and implementation. While working in the United States she participated in long-term projects funded by USAID, the World Bank and the Inter-American Development Bank (IDB) in Argentina, Mexico, Peru, Haiti, Honduras, Nicaragua, Pakistan, India and Albania, among others. She also works as an adviser in the Judicial Council of Buenos Aires City.

Hanne Lund Madsen holds Master's degrees in international development studies, geography, and international human rights law. During 18 years of professional work she has serviced the United Nations, national governments, bilateral donors, international human rights organizations, global networks and local NGOs in the effort to promote democratization, human rights and accountable governance. She combines the disciplines, practices and approaches of both the development and the human rights community, and has been involved in developing strategies and impact assessments of human rights and democracy support and in the operationalization of rights-based approaches within development programming and poverty reductions strategies. She has long experience in designing evaluations and has acted as team leader on a broad range of evaluations in East and Southern Africa, South Asia, the Pacific, Central America and the Balkans.

Natalia Mirimanova is an independent researcher and practitioner in the field of conflict transformation. She has extensive experience throughout Russia, the South Caucasus and Central Asia, and in Moldova, Ukraine, the Balkans and Eastern Europe. Since 1993 she has designed and implemented university and training programmes on conflict analysis and resolution, cross-conflict initiative dialogues and problem-solving workshops in protracted conflict settings, and interdisciplinary research projects in the fields of state formation conflicts; facilitated dialogue and cooperative initiatives between the government, business and civil society sectors; and assisted NGOs and community groups in the design and implementation of civic advocacy campaigns. Natalia has published several books, articles and training manuals on conflict transformation, democracy building in transitional societies and the mass media, and written several television features and news stories. She received her PhD from the Institute for Conflict Analysis and Resolution, George Mason University, USA.

Patrick Molutsi is currently the Executive Secretary of the Tertiary Education Council (TEC) in Botswana. Before joining the TEC, he worked for International IDEA in Stockholm, Sweden, from 1999 to 2003 as Director of Field Programmes. Between 1980 and 1999, he was a lecturer and senior lecturer at the University of Botswana, Department of Sociology. In 1996–9 he supervised two international projects funded by the United Nations Population Fund (UNFPA) on population and development and the Centre of Excellence in Public Administration and Management, sponsored by the German Technical Cooperation Agency (GTZ). Dr Molutsi received his first degree at the University of Botswana. He has a Postgraduate Diploma in population studies (Ghana), an MPhil and a DPhil from the University of Oxford, both in the sociology of development, and a Diploma in public international law (Wolverhampton/Holborn College). Dr Molutsi has done research and written on a wide range of topics, including democracy, education, elections, rural development,

human development, poverty and governance. He has been a consultant to numerous government departments and development cooperation agencies, as well as local and international NGOs, on various issues.

Margaret Sarles is Chief of the Strategic Planning and Research Division, Democracy and Governance, at USAID, Washington, DC, USA.

Fredrik Uggla is a member of the Department for Evaluation and Audit at the Swedish International Development Cooperation Agency (Sida), Stockholm, Sweden.

Michael Wodzicki is Coordinator for Democratic Development at Rights & Democracy (formerly known as the International Centre for Human Rights and Democratic Development), an independent non-partisan body created by the Canadian Parliament and based in Montreal. For three years he worked as a policy adviser to a Canadian member of Parliament and Cabinet minister, and subsequently worked with the Organization for Security and Co-operation in Europe (OSCE) in Belgrade, Serbia, managing a programme which, in partnership with Serbia's National Assembly, sought to increase citizen awareness about the role of the parliament and parliamentarians. He holds a Master's degree in social and political studies from the University of Edinburgh.

About International IDEA

The International Institute for Democracy and Electoral Assistance—International IDEA—is an intergovernmental organization that supports sustainable democracy worldwide. Its objective is to strengthen democratic institutions and processes.

What does International IDEA do?

International IDEA acts as a catalyst for democracy building by providing knowledge resources and policy proposals or by supporting democratic reforms in response to specific national requests. It works together with policy makers, governments, UN agencies and regional organizations engaged in the field of democracy building.

International IDEA provides:

1 *assistance with democratic reforms* in response to specific national requests;
2 *knowledge resources,* in the form of handbooks, databases, websites and expert networks; and
3 *policy proposals* to provoke debate and action on democracy issues.

Areas of work

International IDEA's key areas of expertise are:

1 *Electoral processes.* The design and management of elections has a strong impact on the wider political system. International IDEA seeks to ensure the professional management and independence of elections, the best design of electoral systems, and public confidence in the electoral process.
2 *Political parties.* Polls taken across the world show that voters have little confidence in political parties even though they provide the essential link between the electorate and government. International IDEA analyses how political parties involve their members, how they represent their constituencies, and their public funding arrangements, management and relationship with the public.
3 *Constitution-building processes.* A constitutional process can lay the foundations for peace and development or plant seeds of conflict. International IDEA provides knowledge and makes policy proposals for constitution building that is genuinely nationally owned, sensitive to gender and conflict-prevention dimensions, and responds effectively to national priorities.
4 *Democracy and gender.* If democracies are to be truly representative, then women—who make up over half of the world's population—must be able to participate on equal terms with men. International IDEA develops comparative analyses and tools to advance the participation and representation of women in political life.
5 *Democracy assessments.* Democratization needs to be nationally driven. The State

of Democracy methodology developed by International IDEA allows people to assess their own democracy instead of relying on externally produced indicators or rankings of democracies.

Where does International IDEA work?

International IDEA works worldwide. It is based in Stockholm, Sweden, and has offices in Latin America, Africa and Asia.

Which are International IDEA's member states?

International IDEA's member states are all democracies and provide both political and financial support to the work of the institute. They are: Australia, Barbados, Belgium, Botswana, Canada, Cape Verde, Chile, Costa Rica, Denmark, Finland, Germany, India, Mauritius, Mexico, Namibia, the Netherlands, Norway, Peru, Portugal, South Africa, Spain, Sweden, Switzerland and Uruguay. Japan has observer status.

About Sida

Sida, the Swedish International Development Cooperation Agency, is a government agency under the Ministry of Foreign Affairs. Sida's goal is to contribute to making it possible for poor people to improve their living conditions. Sida works independently within the framework laid sown by the Swedish Government and Parliament. They decide in the financial limits, the countries Sweden – and thus Sida – shall cooperate with and which direction the cooperation shall have. Today Sida has extensive cooperation with some 70 countries.

The overall goal for the Swedish policy for global development is to contribute to equitable and sustainable global development. The goal of Swedish development cooperation is to contribute to an environment supportive of poor people's own efforts to improve their quality of life. This is well in line with the international commitment to halve the proportion of people living in absolute poverty in the world by 2015. It emphasises that poor people themselves have the power to change and develop their communities if they are given the opportunity.

Swedish development cooperation shall promote, and be characterised by, the following central component elements:

Fundamental values:
- respect for human rights
- democracy and good governance
- equality between women and men

Sustainable development:
- sustainable use of natural resources and protection of the environment
- economic growth
- social development and social security

Other component elements:
- conflict management and security
- global public goods

Sida's role

Sweden's partner countries are responsible for working to end poverty in their respective countries. Sida contributes money, advice and competence to the process.

We can contribute to development and change by financing, analysing, conducting dialogues and preparing contributions, projects and programmes. At a practical level the work can take many forms, but common to all contributions is that they must be judged from the perspective of how they will affect poor people.

Sida support

Sida finances thousands of small and large contributions that often span several years. Sida's support ranges from money to bolster government budgets to financing of specific projects. They might concern education, health care, small businesses, housing, legal rights, research, infrastructure and trade agreements. Part of it also supports humanitarian assistance.

Sida's own personnel rarely work directly with the specific activities. These are carried out by experts, politicians, voluntary organisations and others, both from Sweden and the partner countries.

Sida is primarily responsible for Sweden's bilateral assistance but also manages part of Sweden's multilateral assistance.

Index

Index

Index

Index