

RESEARCH REPORT

# EVALUATIONS **AND** LEARNING IN RULE OF LAW ASSISTANCE

ELIN COHEN WITH CONTRIBUTIONS FROM KRISTINA SIMION



**FOLKE  
BERNADOTTE  
ACADEMY**

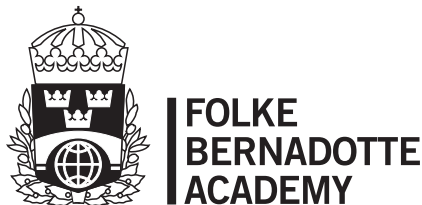
**THE FOLKE BERNADOTTE ACADEMY** is a Swedish government agency dedicated to enhancing the quality and effectiveness of international conflict and crisis management, with a particular focus on peace operations. The overall objective is to contribute to lasting peace and development. FBA functions as a platform for cooperation between Swedish agencies and organisations and their international partners. Its main areas of responsibility are:

- › Civilian personnel contribution to international peace operations
- › Education, training and exercises
- › Policy, research and development
- › National and international cooperation and coordination
- › Funding of civil society peace projects

FBA has a preparedness to offer good offices for conflict resolution initiatives, such as talks between parties to a conflict. Within its mandate, it serves as a national point of contact with international organisations, including the UN, EU, AU, OSCE and NATO.

FBA coordinates the International Forum for the Challenges of Peace Operations and is an active member of the International Association of Peacekeeping Training Centres (IAPTC), Europe's New Training Initiative (ENTRi), and the European Security and Defence College (ESDC).

The agency is named after Count Folke Bernadotte, the first official UN mediator, and its mission reflects Sweden's commitment to international peace, security and development.



**ISBN: 978-91-980487-1-1**

**Author:** Elin Cohen with contributions from Kristina Simion

**© Folke Bernadotte Academy April 2013**

**[www.folkebernadotteacademy.se](http://www.folkebernadotteacademy.se)**

## **LIST OF ACRONYMS AND ABBREVIATIONS**

<b>AusAID</b>	Australian Agency for International Development
<b>CIDA</b>	Canadian International Development Agency
<b>Danida</b>	Danish International Development Agency
<b>DFID</b>	Department for International Development (UK)
<b>DG program</b>	Democracy and Governance program (USAID)
<b>EC</b>	European Commission
<b>ECG</b>	Evaluation Cooperation Group
<b>IEG</b>	Independent Evaluation Group (IEG) of the World Bank
<b>MRE</b>	Management Response System
<b>NGO</b>	Non-governmental organization
<b>Norad</b>	Norwegian Agency for Development Cooperation
<b>NZAID</b>	New Zealand Agency for International Development
<b>OCS</b>	Observatory for Social Change (Observatoire du Changement Social)
<b>OECD/DAC</b>	Organization for Economic Cooperation and Development/ Development Assistance Committee
<b>PRSP</b>	Poverty Reduction Strategy Papers
<b>SADEV</b>	Swedish Agency for Development Evaluation
<b>Sida</b>	Swedish International Development Cooperation Agency
<b>TOR</b>	Terms of Reference
<b>UN</b>	United Nations
<b>UNDP</b>	United Nations Development Programme
<b>UNEG</b>	United Nations Evaluation Group
<b>UNHCR</b>	United Nations High Commissioner for Refugees
<b>USAID</b>	United States Agency for International Development

# TABLE OF CONTENTS

	<b>FOREWORD</b>	
	<b>EXECUTIVE SUMMARY</b>	
1.	<b>INTRODUCTION</b>	9
2.	<b>EVALUATING RULE OF LAW PROGRAMMING</b>	11
2.1.	Defining Rule of Law	11
2.2.	Rationales for Conducting Evaluations	12
2.3.	Demonstrating Results in Rule of Law	13
2.4.	Intuitive Projects: the Problem of Attribution	14
2.5.	Ongoing Change and Confounding Factors: The Challenge of Assessing Impact of Rule of Law Activities	15
3.	<b>CONDUCTING EVALUATIONS</b>	18
3.1.	Joint Evaluations	20
3.2.	Partner Countries Involvement in Evaluations	22
4.	<b>LEARNING FROM EVALUATIONS</b>	24
4.1.	Finding Evaluation Reports	24
4.2.	Absorbing Knowledge from Evaluations	25
4.3.	Post-evaluation Follow-up	27
4.4.	Spotlight Series	28
4.5.	Success Stories and Admitting Failure	29
5.	<b>CONCLUDING RECOMMENDATIONS</b>	31
6.	<b>BIBLIOGRAPHY</b>	33

**THE RULE OF LAW** is an essential foundation of international conflict and crisis management. Experience gained from past decades clearly demonstrates that establishing the foundations for peace, security and development, whilst ensuring justice and fighting impunity, in the immediate post-conflict period and during long-term consolidation, is not possible without the rule of law. Also, for peacebuilding and statebuilding goals, which are reflected in Swedish policy on peace and security, the need to address injustices and increase people's access to justice has been acknowledged.

The Folke Bernadotte Academy (FBA) promotes the rule of law as a means for conflict prevention, security, human rights protection, legal empowerment and other objectives as outlined in the Swedish Government's collective policy on the rule of law. The FBA also provides support for the professionalization of the international field of rule of law assistance.

Recently, the sustainability and effectiveness of rule of law projects has been questioned. The emerging consensus reveals that rule of law programmes need to be evaluated more consistently as the rule of law field has an uneven record of evaluating, disseminating and implementing lessons learned.

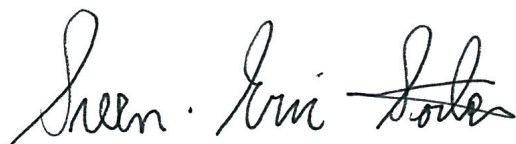
This report concludes that evaluations and thoughtful reflections on past experiences are not keeping pace with the expansion of the rule of law field. It highlights the importance of disseminating and incorporating lessons learned into new or on-going program development as one of the more pressing issues that need to be addressed within the rule of law field. It represents an example of the FBA's work on contributing to improvements and the professionalization of rule of law assistance in order to promote lasting peace, security and development.

The importance of evaluating aid projects is stressed by the Swedish Government through requests to its agencies asking them to show what results come out of aid funded projects and programmes. This report reflects the FBA's recognition of the importance of continuous evaluations and quality assurance as a systematic tool for analysing results. This is an integral part of its mission.

Moreover, this report links in with and reflects the conclusions from the High Level Forum on Aid Effectiveness in Busan in 2011; specifically regarding the importance of sustainable results as well as transparency and accountability. Busan highlighted that evaluations are important but that they need to respect the ability and capacity of the host countries to administer the requirements of multiple donor evaluations. It also highlighted the need for coordination among donors.

Thus, the current report is both timely and much needed and will hopefully provide knowledge, inspiration and the incentive for future evaluations of rule of law projects and programmes.

The Author, Elin Cohen (JSD), is a visiting scholar at the University of Washington and an independent consultant with extensive rule of law experience from Eastern and Western Africa. She is a member of the FBA Rule of Law Research Working Group. Kristina Simion (LL.M.) is a Rule of Law Officer at the FBA, currently completing her doctoral studies at the Australian National University.



Mr. Sven-Erik Söder  
Director General  
Folke Bernadotte Academy

# EXECUTIVE SUMMARY

The rule of law field is growing as operations, actors involved, and budgetary spending are all increasing. Rule of law projects range from strengthening public institutions and reforming the judiciary to training ‘barefoot’ lawyers. Projects are motivated by the overarching principle that the rule of law is necessary for economic development and the protection of democracy. The rule of law is also recognised as essential to ending or preventing violent conflict, supporting sustainable peace and statebuilding, and promoting security and the protection of human rights.

Many observers are questioning whether rule of law projects are actually accomplishing their goals and even if they were to be attained, whether the impact of these projects can be sustained. So as to learn from past experiences, there is an emerging consensus that rule of law programmes need to be evaluated. Evaluations assess whether programmatic objectives have been met and whether specific activities have been effectively implemented. Evaluations are tools for international development agencies, peacekeeping and peacebuilding missions, implementing organizations, and partner countries to acquire knowledge about lessons learned and good (as well as not so good) practices; however, evaluations and thoughtful reflections of past experiences are not keeping pace with the expansion of the field, which has an uneven record of evaluating, disseminating, and implementing lessons learned.

Several factors contribute to why constructive use of lessons learned is not keeping pace with the expansion of the rule of law field:

- › *Differing understandings of what the rule of law entails;*
- › *Numerous rule of law actors operating in the field;*
- › *An uneven quality of the methodological approaches used to evaluate rule of law projects or programs;*
- › *Inadequately prepared and thought out evaluations produced on tight timelines;*
- › *Limited joint evaluations between donors and national counterparts;*
- › *Difficulty in accessing evaluation reports with valuable lessons learned;*
- › *Evaluation reports used to justify or discontinue on-going programming, but rarely to gain knowledge of what works or what does not work in the field; and*
- › *Lessons learned from problematic projects are often ignored or forgotten.*

As numerous rule of law initiatives have taken place around the globe in the last decades, there has been an increase of knowledge and experience in the field. Evaluations therefore play an important role by transmitting information about past experiences. Disseminating and incorporating lessons learned into new or on-going program development remains one of the more pressing issues that need to be addressed by the rule of law community.

## SUMMARY OF KEY CONCLUSIONS OF THE REPORT:

- › **Review past evaluations.** This review should highlight both what worked well, as well as what did not work so well. To conduct such a review, donors would have to be consulted to gain access to evaluation reports and find out more about the project that might not have made it into the evaluation. Such an engagement with donors would provide an opportunity to solicit input and interest in the anticipated publication, which would increase the chances that the publication will be consulted and lessons learned shared.
- › **Conduct an empirical case study.** Such a case study should explore issues on how contractors, local and international NGOs, or partner countries engage with evaluations and acquire knowledge about lessons learned as the specific project design and implementation of particular programs is carried out by contractors and NGOs.
- › **Incorporate funding for evaluations in projects and programs award allocations.** These funds should help projects and programs gain knowledge of what has worked or not worked in the field, when new rule of law assistance programs and projects are conceived and designed, and make sure lessons learned from problematic projects are not ignored or forgotten.
- › **Make underlying assumptions of a program explicit.** These assumptions should be made explicit during the planning stages so that subsequent evaluations of the same program can link the project activities in a coherent and logical way to the outputs, outcomes and overall objectives of the program.
- › **Create a database of evaluations.** These databases should make evaluation reports with valuable lessons learned more accessible as the currently inadequate ways evaluation reports are collected and disseminated seriously limit the possible use of lessons learned.
- › **Use evaluations to demonstrate results of rule of law programs.** The use of these evaluations should help grant-awarding institutions demonstrate the results of the rule of law programs and projects and avoid programs and projects that receive continued support without knowing whether development assistance has been carried out in an accountable and effective manner.
- › **Ensure that evaluations are taken into consideration in decision-making forums and policy development.** Such a strategy should ensure that past lessons are systematically taken into consideration in decision-making forums and that policy is based on lessons learned. This strategy recognizes that transmitting lessons learned into policy is inherently difficult with so many levels of government agencies and intergovernmental and non-governmental contractors involved in evaluating rule of law programs.
- › **Enhance the active use of evaluation findings and recommendations.** Encourage the use of evaluation findings when designing new projects and programs.

- › **Incorporate research series on past donor experiences in professional training of rule of law practitioners.** This research series should help improve the learning from existing evaluations and avoid learning gaps within organizations as practitioners rarely find the time to read lengthy evaluation reports.
- › **Support a research series.** This research should review and synthesize past donor experiences in particular areas of the rule of law such as judicial training and customary justice.



# 1. INTRODUCTION


The rule of law field is growing in terms of number of operations, actors involved, and budgetary spending. Rule of law projects range from strengthening public institutions and reforming the judiciary to training ‘barefoot’ lawyers. Projects are motivated by the overarching principle that the rule of law is necessary for economic development and the protection of democracy. Rule of law is also recognised as essential to ending or preventing violent conflict, supporting sustainable peace, supporting statebuilding by ensuring responsive and legitimate States, and promoting security and the protection of human rights.

Recently many people have questioned whether rule of law projects are actually accomplishing their goals and even if their goals were attained, whether the impact can be sustained. There is an emerging consensus that rule of law programmes need to be evaluated so future projects can gain from past experiences. It is important that the impact of the intervention, whether it is large or small, be assessed.

Evaluations are meant to function as an accountability check to assess whether programmatic objectives have been met and whether specific activities have been implemented effectively. Evaluations are tools for international development agencies, peacekeeping and peacebuilding missions, implementing organizations, and partner countries to acquire knowledge about lessons learned and good (as well as not so good) practices; however, evaluations and thoughtful reflections of past experiences are not keeping pace with the expansion of the field. The rule of law field exhibits an uneven record of evaluating, disseminating, and implementing lessons learned.

Multiple factors contribute to why constructive use of lessons learned from past experiences is not keeping up with the expansion of the rule of law field. With multiple understandings of what the rule of law entails and the numerous rule of law actors operating in the field, it is difficult to grasp what type of rule of law activities are taking place, what has already been tested, what worked, and what did not work. There is an uneven quality of the methodological approaches employed in evaluating rule of law projects and programs. Many rule of law projects aspire to bring about broad overarching objectives, such as preventing conflict, establishing peace and security, and improving democratic governance. The specific activities undertaken to advance these goals are often difficult to link to a specific outcome. Evaluations too often assume that a particular result can be attributed to the project that is being evaluated without examining alternative explanations that might have contributed to the change. In addition, rule of law projects are often designed to encourage change or support on-going change in a society. Thus, there are a host of other social, legal, economical, and political factors and activities that simultaneously take place. Unless there is an effort from the start of the project to consider and account for these confounding factors, it is difficult to assess the effectiveness of a particular rule of law program.

There is growing recognition that quality evaluations require planning by staff who understand the basic principles of evaluations and that the evaluators need to be professional, independent, and given realistic time frames. Following the Paris Declaration on Aid



Effectiveness, donors and partner countries have shown a growing interest in conducting joint evaluations to align the evaluations to the partner country's priorities. Despite this awareness and stated interest, evaluations are still often thrown together as an afterthought with tight timelines and few evaluations are truly joint evaluations.

Another pressing issue is that evaluation reports with valuable lessons learned are difficult to access and few people read them. Evaluations that receive more buy-in from a wider group of actors are more likely to be read. The inadequate ways evaluation reports are collected and disseminated seriously limits the possible use of lessons learned. Even when evaluation reports are read or discussed, they are primarily used to justify or discontinue on-going programming and are rarely used to gain knowledge of what has worked or not worked in the field when new rule of law assistance programs and projects are conceived and designed.

Although several actors in the assistance community have instituted post-evaluation systems to stimulate an active use of an evaluation's findings and recommendations, the uptake has been slow and the actual application in decision-making processes limited. Several donors have launched evaluation summary series to make the main findings more widely known, but the various efforts to highlight experiences have so far been ineffective, unsustainable, or difficult to locate. There is a tendency to focus on success stories and to ignore problematic issues.

## 2. EVALUATING RULE OF LAW PROGRAMMING

### 2.1. DEFINING RULE OF LAW

The rule of law concept<sup>1</sup> is used by international agencies, politicians, lawyers, and other practitioners within the international field. Its meaning is debated and often remains unclear. The fundamental principle of the rule of law is that States should be governed by laws and that no one is above the law. These laws should respect fundamental democratic principles including legality, predictability, proportionality, transparency, equality, and human rights. Rule of law should enhance democratic governance and human rights so as to advance economic development. As such, rule of law is essential when managing crisis, post-conflict reconstruction, and peacebuilding. Because rule of law is such a broad concept, it is often bundled with other concepts such as ‘rule of law and peace and security’ or ‘rule of law and access to justice’. Thus, the different ways of grouping rule of law and the lack of a common understanding of the meaning of rule of law makes it difficult to evaluate the effect of rule of law initiatives. For example, a cross-national study on the effects of US foreign assistance to democracy building found that 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.”<sup>2</sup> However, it is possible to compare specific aspects of rule of law such as the delivery of legal aid services or education of judges and prison personnel across programs, countries, and donors. For such a comparison to be fruitful, it is important to specify clearly how broader concepts are understood within particular projects. For example, what does a peacebuilding and conflict resolution program seeking to support the process of conflict transformation really want to accomplish? Moreover, such a comparison must remain grounded within each particular political, economic, cultural, and social context to remain meaningful in practice. Logically, the understanding of what the rule of law entails will vary between different locations, actors, donors, partners, and recipients, but as long as such differences are taken into consideration, meaningful comparisons can be made.

To measure the advancement of rule of law, various indexes have recently been developed, notably the World Justice Forum Rule of Law Index and the Worldwide Governance Indicators. However, this increased interest of the development of rule of law indicators has also been accompanied by recognition that these indicators may have some practical limitations. Faced with difficulties defining and measuring rule of law, indicator projects, often separate the rule of law concept from other related concepts such as democracy. Attempts to standardize rule of law performances have been criticized for over simplifying complex issues and not capturing the pluralistic local arrangements formed by both formal and informal legal institutions. Similarly, observers have criticized the typical means – expert panels and public polls – of collecting data about indicators. Nevertheless, if rule of law indexes can overcome methodological shortcomings, indicators could provide information about complex issues that need further study.

1. The 2004 UN document, Rule of Law and Transitional Justice in Conflict and Post-Conflict Societies describes the rule of law as "...a principle of governance in which all persons, institutions and entities, public and private, including the State itself, are accountable to laws that are publicly promulgated, equally enforced and independently adjudicated, and which are consistent with international human rights norms and standards. It requires, as well, measures to ensure adherence to the principles of supremacy of law, equality before the law, accountability to the law, fairness in the application of the law, separation of powers, participation in decision-making, legal certainty, avoidance of arbitrariness and procedural and legal transparency." This description of the concept is also used by, for example, the EU policy, Justice Components for CSDP Missions from 2011.

2. Sarles 2007; Burnell (ed.) 2007.

Although the evaluation of individual projects is important, the value of evaluation is substantially augmented when findings from one project are compared with another project. Comparison makes it easier to identify relationships between interventions and outcomes and what intervening variables may affect the outcome of the project. Because the understanding of the rule of law field is so fragmented, it is important to compare and contrast approaches and results. For a comparison to be meaningful, it needs to take into account the ‘the thick description’ of local arrangements and culture and specific characteristics that cannot easily be reduced to scores or numbers. Nevertheless, evaluations that systematically compare findings from multiple projects in various countries or settings are beneficial, although they remain rather unusual.

## 2.2. RATIONALES FOR CONDUCTING EVALUATIONS

An evaluation is “a systematic and objective assessment of an on-going or completed project, program or policy, its design, implementation and results.”<sup>3</sup> The Swedish International Development Cooperation Agency (Sida) evaluation manual refers to evaluations as ‘a reality test’ of the results of programmatic actions in relation to prior objectives, standards, and expectations.<sup>4</sup> Evaluations serve two primary functions: to find out whether the responsible organization has carried out the project in a satisfactory manner considering the prevailing circumstances and to find out whether the intended results were achieved. In addition, an evaluation should produce substantive knowledge about what worked, what did not work, and how the reviewed activities can be improved, either as a direct continuation of the project or in an all-together different program. While the first function is more narrowly focused on the accountability of a particular project, this function stresses evaluations as a tool to improve organizational performance and learning.<sup>5</sup>

It is important to identify what the purpose of an evaluation is rather than just embarking on a symbolic evaluation because one is expected to do evaluations in modern organizations. Identifying the purpose of the evaluation up-front will inform the design of the evaluation, its methods, and who will serve on the evaluation team. Because evaluations that lack clear objectives are less informative, they are less likely to be used in future programs and projects. To the extent evaluations are used, they are typically used as a source of information in conjunction with other sources to inform programmatic and financial decisions. That is, these evaluations ask the following questions: Was the program effective? and Should this program or a similar program be funded again? Although some studies suggest that evaluations are used as a stalling tactic or to legitimize on-going programs, a more disturbing finding is that many evaluations are never read or considered.

Monitoring is the tracking of activities in relation to the targets throughout the duration of a project. By closely monitoring activities and their sequencing, the implementation plan can be adjusted to encourage progress. However, as rule of law projects tend to run on fairly short project cycles and unexpected events occur, the follow-up needs to start at an early stage and continue throughout the duration of the project. Monitoring is an important tool to ensure that progress is being made during the implementation of the project and serves as a basis for the evaluators to track decisions.

3. Organization for Economic Cooperation and Development / Development Assistance Committee (OECD/DAC) 2002.

4. Molund and Schill 2007.

5. It has for example been recognized by the Swedish Agency for Public Management that evaluations should be used in a more systematic manner to analyse long term results, covering ten years and beyond. Swedish Agency for Public Management 2011.

## WORLD BANK, 2008

### **Implementation Completion and Results Report (IRD-44010) on a Loan in the Amount of US \$33.00 Million Equivalent to the Republic of Guatemala for a Judicial Reform Project.**

An evaluation of a judicial reform program in Guatemala found that a well executed assessment and training exercise to accelerate processing times in the appellate courts was not followed up. Instead, as soon as the training was completed, the appellate bench was replaced. The consultant left the assessment and training material “in three very fat volumes now residing in the PCU [Project Coordinating Unit], but the PCU could not provide evidence of even informal adoption of his suggestions”.

## 2.3. DEMONSTRATING RESULTS IN RULE OF LAW

Following the end of the Cold War, there was a renewed surge in rule of law initiatives in development and crisis management and post-conflict reconstruction. Initially, there was a sense of urgency, and new projects were fuelled by anecdotal evidence of success. Many of these programs and projects received continued support without the donors requiring an evaluation.

Over the past two decades, development agencies have recognized that the general public and taxpayers are demanding credible assessments of whether aid ‘works’. Similarly, academic literature has begun to question the effectiveness of the rule of law enterprise.<sup>6</sup> In response to these concerns, more actors have formally developed monitoring instruments to track whether development assistance has been carried out in an accountable and effective manner.

The Paris Declaration on Aid Effectiveness (2005) and its subsequent documents require local ownership, donor coordination, and mutual accountability as a way to produce and measure results.<sup>7</sup> Increasingly, interested parties have focused on whether development assistance produces ‘progress’ in a particular field and various indicators have been developed to measure this advancement. There is also an increased acknowledgment that new programs need to be based on lessons learned from previous projects as “knowledge based action produces better results than a stab in the dark or uninformed good intentions.”<sup>8</sup>

## BUSAN AND THE NEW DEAL

**The fourth high level meeting on Aid Effectiveness was held in Busan, in December 2011. The meeting was the last in a series of meetings which included Rome 2003, Paris 2005 and Accra 2008.**

In Busan, existing commitments from Paris and Accra were confirmed and further developed, with a stronger focus on the country level, including commitments concerning performance frameworks.

It was established that a global action plan will be developed to strengthen partner countries capacity to measure and evaluate effects as well as strengthen result-based governance in partnership countries. A few relevant output and outcome indicators will be identified based on the partnership countries own goals and priorities. In addition, common baselines, evaluations and audits will be built into the review and follow up of results.

6. See for example Trubek and Santos 2006; Carothers 2006; Golub 2006; Bergling 2006; and Taylor 2010.

7. Regardless of the commitments of the Paris Declaration on Aid Effectiveness (2005) and the Accra Agenda for Action (2008) results and value for money in fragile states has been modest. See A NEW DEAL for engagement in fragile states 2012.

8. Hammergren 2002.

The New Deal, a new framework for engagement in fragile and conflict-affected states, was also presented in Busan. It proposes key peacebuilding and statebuilding goals to address injustices and increase people's access to justice.

The New Deal focuses on achieving better results in fragile states and managing resources more effectively. It also stresses the importance of tracking progress at the global and the country level and presents a plan for the implementation of peace- and statebuilding objectives to measure development.

## 2.4. INTUITIVE PROJECTS: THE PROBLEM OF ATTRIBUTION

Many rule of law projects aspire to bring about broad overarching objectives, such as conflict prevention, improved democratic governance, and economic growth. To reach these overarching objectives, a program might seek to speed up case management, to increase access to alternative dispute resolution, or to professionalize the legal profession. To attain these outcomes, donors have funded programs that support the revision of legislation, business licensing processes, and regulatory agencies. In addition, these donors have computerized and refurbished courthouses and held training sessions for various actors in the legal profession. The links between the overarching objectives, the outcomes, the outputs, and the proposed activities involve a number of steps that are hard to trace and each step is difficult to attribute to the next. For example, to improve access to justice, a judicial reform project in Guatemala sponsored numerous workshops to sensitize judges and staff to the needs of indigenous populations, publication of material in indigenous languages, and the hiring/training of interpreters. However, the project did not attempt to gauge how these training sessions and publications might have affected access to justice. In their evaluation, the evaluators summed up the problem as follows: “[W]e know that these activities occurred, but absent any effort to evaluate impact, [we] cannot attempt to assess what it [the impact] was.”<sup>9</sup>

There is an “undeniable common sense appeal”<sup>10</sup> that well-functioning and predictable judicial and administrative institutions should attract investment. Many observers have questioned whether there is a causal link between legal reforms and increased investment. For example, Amanda Perry-Kessaris notes that the World Bank’s Independent Evaluation Group (IEG) calls upon the World Bank’s Doing Business team to make its reform analysis more meaningful and, in Perry-Kessaris’ words, “stop assuming that its work was having the intended consequences”.<sup>11</sup> Evaluation reports assuming that the achievement of the direct output results (training of 75 prosecutors) was directly connected with the fulfilment of the overarching objective (free and fair trials) is rather common. Some evaluations are more critical about this ‘common sense’ practice than others. A World Bank evaluation<sup>12</sup> of three judicial reform projects in South America found that all three projects showed weak linkages between the objectives and the proposed activities. The evaluation pointed out that the most apparent illustration of this flaw related to the construction and refurbishing of courthouses and other legal institutions, which received considerable funds in the three projects. However, the program documents lacked any kind of discussion concerning the impact renovated courthouses might have had on the overarching program objectives – creating a more effective, accessible, and credible judicial system. Instead, the evaluations criticize the projects for assuming that the links between actions and objectives appeared *almost intuitive*.

9. World Bank 2008.

10. Carothers 2003.

11. Cohen, Fandl, Perry-Kessaris and Taylor 2011.

12. World Bank 2010.



**Implementation Completion and Result Report (Cr. 3384-CE), on a Credit in the Amount of SDR 13.6 Million (US\$ 18.2 Million Equivalent) to the Democratic Socialist Republic of Sri Lanka for a Legal and Judicial Reform Project.**

In the World Bank funded project in Sri Lanka, the overall objective was to promote foreign investment and economic growth by making the legal and judicial framework “more efficient, transparent and responsive to the needs of the public at large and private sector in particular”.

**The expected outcomes of the reform were:**

- (a) To modernize the legislative framework that impacts private sector activity;
- (b) To improve the administration, monitoring, and regulatory functions of the Company Registry; and
- (c) To build capacity of the judiciary and other institutions providing dispute resolution services.

The project ultimately contributed to the drafting of a new Companies Law, but the modernization of the company registry was delayed. Judges, lawyers, and law school faculty were trained in commercial law and commercial mediation, courthouses were refurbished, and case tracking systems were introduced in a selected number of courthouses. The project set forth a number of performance indicators, including the number of successful mediations, increased number of judgments written with the use of information technology, and increased number of professionals trained in commercial law.

Although we learn from the evaluation report that 24 courthouses were refurbished, 180 judges were writing their judgments using computers, and the legal professionals liked the training they obtained, these easily enumerated statistics do not address the fundamental questions motivating the project: How do these activities affect foreign investment and economic growth and how do we know whether specific legal reforms aimed at improving the investment climate affected the overarching objective?

## **2.5. ONGOING CHANGE AND CONFOUNDING FACTORS: THE CHALLENGE OF ASSESSING IMPACT OF RULE OF LAW ACTIVITIES**

As rule of law interventions are implemented in a social, economic, and legal environment where there are a host of other factors at play (e.g., the internal political process, conflict, external pressure, civil society activities, and other donor projects), it is often difficult to gauge the effectiveness of a rule of law program.<sup>13</sup> Considering these challenges, it is important to formulate realistic, specific, and at times more modest goals. It is equally important to acknowledge that there are many other factors at play that could influence the outcome of a project.

Projects that have overly optimistic and vague programmatic objectives are difficult to evaluate. These unrealistic and vague objectives contribute to the perception that rule of law and democracy programs are difficult to evaluate. For example, an initiative supporting a constitutional review process where the project’s stated goal is to pass a new constitution is likely to fail as the review process might still be on-going when the project comes to an end and other confounding factors are likely to affect the outcome of the review process.

<sup>13</sup>. See for example Special EU Programmes Body 2007.

A more appropriate, modest, realistic, and attainable objective is needed. For example, the objective might be to increase participation in the review process. Thus, the perception that rule of law interventions are difficult to evaluate can be altered if the objectives are realistic and confounding factors are considered.

Observable changes celebrated as programmatic successes are often difficult to attribute to a particular program. The importance of other confounding factors that might very well have affected the overall goal is often diminished. For example, a review of the United States Agency for International Development's (USAID) justice program in Latin America found a significant decline of certain human rights abuses. Although the change might have been initiated by USAID-funded projects, confounding factors (e.g., the end of civil war, external political pressure, and programs funded by other donors) could also have contributed to this change.<sup>14</sup> Apparently, rule of law evaluations often do not consider these types of alternative explanations when a desirable outcome is observed. A study reviewing a sample of 25 evaluations of the USAID's democracy and governance (DG) programming found that only two evaluations carefully considered whether alternative explanations could have contributed to the observed results.<sup>15</sup>

## RANDOMIZED IMPACT EVALUATIONS

**In the rule of law field, relatively few evaluations are randomized impact evaluations where the outcomes for those who participated in a program are compared with those that did not participate in the program.**

One example of a randomized impact evaluation is a World Bank evaluation<sup>16</sup> of the effects of legal aid on the wellbeing of poor women and children in Ecuador. Using a mixed-method approach, the evaluation found that women who used the legal aid centres were better off legally, economically, and subjectively than those that had not received legal aid services. Such a rigorous evaluation requires more time than more commonly featured 'fly-in' evaluations, which are typically conducted over two to five weeks. These evaluations also require more methodological expertise and are more expensive.

An impact evaluation is much more doable if preparations for such an evaluation are done at the initial stages of the program design. For example, the NGOs in Ecuador that provided legal services were required to keep records of all clients they served in the program. This information helped the evaluation team design a survey, establish a control group, and select participants for the evaluation.

To address the all too common issue of improper assumptions and vague objectives, donors are starting to require projects use a logical framework approach (logframe) where specific objectives are identified at the initial stage of the project and a clear plan is established for how these objectives will be attained.

By making underlying assumptions explicit, a program's plan, and later the evaluation of the same program, should link the project activities (e.g., training of paralegals in underserved areas) in a coherent and logical way to the expected outputs (e.g., more skilled legal advisors in underserved areas), outcomes (e.g., more people in the underserved areas claimed legal rights as a result of obtained legal services), and overall objective (e.g., greater access to justice in underserved areas). Nevertheless, the use of logframes is uneven and is at times only partially used.

14. Hammergren 2003.

15. Bollen, Paxton and Morishima 2005.

16. World Bank 2010.



Several conditions explain this uneven pattern of establishing a clear link between project activities and expected outcomes. For some projects, institutional memory is compromised because of high turnover of staff or a high reliance on short-term contractors. In addition, rule of law projects often respond so quickly to change there is little time to plan and develop a logframe. Thus, at times, these 'planning deficiencies' probably reduce the overall effectiveness of the projects.<sup>17</sup>

### LOGICAL FRAMEWORK APPROACH

**The logframe approach is based upon the theory-based evaluation method, which seeks to make underlying hypotheses guiding a program explicit.<sup>18</sup> The thorough preparation of a logframe helps planners ensure:**

- (a) that activities are directed towards clearly stated goals;
- (b) that important assumptions about external factors are taken into account;
- (c) that the goals can be logically derived from the activities; and
- (d) that indicators are identified for monitoring.

17. Poate, Riddell, Chapman and Curran 2000.

18. Poate, Riddell, Chapman and Curran 2000.

### 3. CONDUCTING EVALUATIONS

Over the last two decades, many international evaluations have been produced. To guide this work, major donors have developed step-by-step evaluation manuals and are regularly updating their evaluation policies and practices. Most donors align their evaluation policies and guidelines with international standards developed by the Organization for Economic Cooperation and Development/Development Assistance Committee (OECD/DAC), United Nations Evaluation Group (UNEG), or the Multilateral Financial Institution's Evaluation Cooperation Group (ECG).

#### OECD/DAC EVALUATION PRINCIPLES

**To assess development interventions, OECD/DAC has developed a set of evaluation principles adopted by many organizations. These guiding principles for what evaluations of development interventions should cover are:**

- a) Does the intervention effectively achieve its objectives?
- b) What are the intended and unintended impacts of the development intervention?
- c) Does the intervention consider its relevance in relation to the needs and priorities of the target group?
- d) Is the intervention sustainable after the completion of the project?
- e) Are the financial and human resources efficiently used and can they be justified by the project's result?<sup>19</sup>

Producing effective evaluations of international assistance in conflict areas is challenging. There is often limited time to plan the interventions, there is often constant change in the conflict, and there is often a high concentration of donor-funded projects that approach the conflict from different perspectives including defence, development, and humanitarian aid. In addition, there is often no existing baseline data and data collection is constrained by the security situation. For example, an audit found that 40% of the Department for International Development's (DFID) monitoring and evaluation efforts in post conflicts and fragile States had been disrupted by security concerns.<sup>20</sup> Given the complexity of peace and security operations, the OECD/DAC has developed a separate guide for evaluating conflict preventions and peacebuilding activities.<sup>21</sup>

Evaluations are carried out on different levels and in various formats: many donor agencies have centralized evaluation units that are independent from the programmatic work. These centralized evaluation units tend to evaluate a broader strategy or sector so as to contribute more general knowledge to a development approach or policy. In addition, agencies' regional and programmatic units also undertake or commission evaluations, but these evaluations tend to focus on individual programs or projects. As few donors have the capacity to carry out their own evaluations, donors have increasingly depended on independent contractors to conduct evaluations<sup>22</sup>. Recently, it has become more common to conduct joint evaluations

19. Molund and Schill 2007.

20. Department for International Development 2010.

21. Organization for Economic Cooperation and Development / Development Assistance Committee 2008.

22. See for instance the work of Channel Research, <http://www.channelresearch.com/>.

and donors have involved partner countries in the evaluation process, although the practical implications of these strategies have been limited.

Good evaluation practice requires an evaluation team to operate independently from the program team to ensure objectivity and impartiality. Several forces might compromise an evaluation's independence: external or internal pressure to not disclose certain findings; the withholding of program documents or sources from the evaluators; the evaluators' self-censorship to not offend colleagues; and the evaluators' concern that findings might negatively impact future job prospects. To ensure independence, several donor agencies have separated the central evaluation office from the rest of the organization. For example, the Independent Evaluation Group at the World Bank reports directly to the board of directors. However, many donor agencies have decentralized project evaluations for field staff although there is some concern that this "moves the responsibility for the quality of a large portion of evaluations in a donor's portfolio to individuals who generally have limited evaluation training".<sup>23</sup> Although evaluations are often carried out by or in collaboration with consultants, the program officers need to have a good understanding of evaluation techniques to draft the statement of work. Thus, realistic timeframes and budgets as well as detailed scope of works outlining the purpose of the evaluation are key components to an effective evaluation.

Anecdotal evidence from evaluation contractors and USAID staff suggest that a statement of works is often lacking in these areas.<sup>24</sup> To address similar concerns, the New Zealand Agency for International Development (NZAID), in a review of its Terms of References (TOR), found that the objectives, the evaluation questions, the methodology sections, and the reporting requirements were weak or unclear.<sup>25</sup> Hence, the agency developed clearer guidelines for developing TOR. Nevertheless, the evaluators' skills and competences, both in evaluation methods and the subject areas, directly affect the quality of an evaluation. Although somewhat dated, a review of team members carrying out evaluations for USAID between 1998 and 1999 found that three quarters of the evaluators had no theoretical or practical evaluation expertise.<sup>26</sup> Recently, this lack of expertise has been addressed by providing more training for evaluators and by establishing more regional evaluation networks. To address continued concern about evaluators' competences, the United Nations Evaluation Group has adopted a set of core qualifications to professionalize evaluation practices within the UN system.

Evaluations carried out by central evaluation units appear on the whole to be of better quality than evaluations commissioned by regional or programmatic donor units. For example, the Ministry of Foreign Affairs of the Netherlands has reported that the agency plans to reduce the number of field office evaluations, as the quality of these evaluations is very low.<sup>27</sup> One possible explanation for this observation is that the staff members in the centralized units are better qualified to oversee an evaluation and a core concern of their job is to produce good quality evaluations.<sup>28</sup>

Studies examining the content of evaluations commissioned by decentralized departments within aid agencies have found that the quality of these evaluations is uneven. A study reviewing evaluations commissioned by Sida found that the applied methodology and evaluation analysis were commonly so weak that the conclusions and recommendations should be

23. Hageboeck 2009.

24. Hageboeck 2009.

25. USAID 2009; and NZAID 2008.

26. Clapp-Wincek and Blue 2001.

27. Hildenwall, Sjöberg and Wikström Öbrand 2008.

28. "The vast majority (32 of 36) of the [OECD/DAC] Network members' centralised evaluations are carried out by consultants or by staff and consultants together. Four Network members (EBRD, IFC, SADEV and WB) rely solely on their own staff to carry out evaluations." Hildenwall, Sjöberg and Wikström Öbrand 2008.

called into question.<sup>29</sup> A study reviewing a sample of 25 evaluations commissioned by the USAID's democracy and governance (DG) program found that the evaluations needed major improvement.<sup>30</sup> The review found that most evaluations provided insufficient information of sources used to assess the reliability of the information underpinning the evaluators' findings. Furthermore, the review found that evaluation reports frequently failed to provide detailed information, beyond notions such as 'strengthening civil society', about what activities had actually occurred. When the reports were more detailed, they tended to focus on immediate output of the undertaken activities (e.g., 200,000 how-to-vote brochures with illustrations were produced) rather than what outcome these activities actually had (e.g., did the voter brochures affect voter participation?). As noted earlier, it was highly unusual for the evaluations under review to consider alternative explanations for observed results, including the presence of other donor programs. The review concludes that the 'fly-in' approach employed to conduct the USAID DG evaluation, where a team interviews staff and representatives of the local people for a couple of weeks, is of value but these methods alone are not strong enough to support claims of program effectiveness.

### 3.1. JOINT EVALUATIONS

Over the last 15 years, the number of joint evaluations has significantly increased. The increased interest in coordinating joint evaluations has been propelled by the broader development agenda focusing on donor coordination and sector-wide approaches, aid effectiveness, and results.<sup>31</sup> Joint evaluations are particularly useful when there is a high concentration of donor activities or when there is a need to evaluate effectiveness of assistance funded through basket or general budget support. Moreover, joint evaluations can also facilitate evaluations of more controversial development issues or mitigate evaluation fatigue in host countries.

The first major joint evaluation was carried out in the wake of the genocide in Rwanda. This joint evaluation reviewed the planning, coordination, and execution of emergency relief, conflict management, and reconstruction efforts.<sup>32</sup> The joint evaluation was spearheaded by the Danish International Development Agency (Danida), one of the largest proponents for joint evaluations, and included an unprecedented number of bilateral and multilateral donors (30). The evaluation required an extraordinary amount of time (15 months) and resources (among other things 52 consultants were contracted). Since then, joint evaluation guidelines have been developed, and several other larger-scale joint evaluations have followed in areas where major humanitarian, disaster, or peacebuilding interventions have taken place, for instance in Southern Sudan or post-Tsunami relief efforts. Thus, joint evaluations, such as the one for Southern Sudan, present a holistic picture of what efforts worked well or not so well. For example, by reviewing all donor-funded programs, the evaluation found that community reconciliation and peacebuilding efforts were isolated events that lacked links to national initiatives and were characterized by poor monitoring and follow-up.<sup>33</sup>

29. Hanberger and Gisselberg 2006.

30. Bollen, Paxton and Morishima 2005.

31. See for example Swedish Agency for Public Management 2012.

32. Eriksson, et al. 1996.

33. Bennett, et al. 2010.

Although there seems to be a general consensus that joint evaluations are effective, the majority of evaluations remain single donor evaluations. A review of close to 700 evaluations found that 75% of the evaluations were single donor reviews, 7% were joint evaluations with another donor, and 15% were joint evaluations with a partner country (the level of involvement of partner countries is discussed in more detail below).<sup>34</sup> Country or sector evaluations are usually conducted by the centralized evaluation units, but most project evaluations are carried out by the field missions. However, few field missions carry out joint evaluations because of a lack of resources and lack of ‘evaluation knowledge’. Thus, these issues combined with other factors such as cost, logistical hurdles, and language barriers mean that the use of joint evaluations remains limited.

To facilitate coordination among the centralized evaluation units of donors, the DAC secretariat has launched an on-line platform to share evaluation plans. According to an OECD survey, staff at the centralized evaluation units commonly share their own evaluation plans through this platform and consistently review other members’ evaluation plans for future evaluations.<sup>35</sup> However, actual cooperation among donors is still limited, and when it occurs, it takes place primarily between a smaller group of like-minded donors, such as the Scandinavian countries, the Netherlands, and DFID.

Because joint evaluations require significant coordination between the parties involved, they tend to become more expensive and time-consuming than single donor evaluations. For example, the execution of the joint conflict prevention and peacebuilding evaluation of Southern Sudan (excluding the time and resources it took to plan the evaluation) took one year to complete and involved more than 20 consultants. Donors need to spend time negotiating the objectives, the methods applied, as well as the selection of the evaluation team. Moreover, the relevance of the topic under review might be time-sensitive for some donors, so the timing of the evaluation process becomes more challenging. To delegate responsibility for the planning of an evaluation, partners in a joint evaluation have to trust each other. Finally, geographical distances, language barriers, and domestic public procurement requirements might further complicate the ability of donors to collaborate on joint evaluations.

There are different approaches to joint evaluations. For example, a joint evaluation assessing anti-corruption support conducted evaluations of individual donor projects and programs in five different countries in order to compare the support across donors and countries. A different strategy is to review already existing evaluations commissioned by multiple donors in a particular subject area. Such an approach was used in a joint evaluation to assess donor support to internally displaced persons within the broader context of human rights and humanitarian law. The review of the existing evaluation reports was supplemented with meetings and interviews with evaluators and the representatives from the evaluation and program units of the participating donor organizations. Synthesis reports distilling the key programmatic and methodological lessons learned from evaluation reports, across donors but within a particular subject area, are highly useful as few donor officials read other donor agencies’ evaluations. These sorts of efforts constitute a promising avenue for analysing lessons learned on a broader scale. When officials from various donor agencies are engaged in the compilation of a synthesis evaluation, more donors might actually read the report.

---

34. OECD 2010.

35. OECD 2010.

### 3.2. PARTNER COUNTRIES INVOLVEMENT IN EVALUATIONS

Two of the guiding principles in the Paris Declaration on Aid Effectiveness are local ownership and mutual accountability between donors and partner countries. However, evaluations have traditionally been one of the more heavily donor-driven activities in development cooperation. Evaluations have typically been tailored towards the donors' need to demonstrate accountability and impact, while the partner country's needs might not have been met or even solicited. Instead, recipients of development assistance have at times perceived multiple donors conducting overlapping evaluations as an irritating distraction from their work. One Danida officer expressed the concern this way: "In Uganda there are so many studies and evaluations produced every year that the government finds it difficult to keep track of/ internalize them all".<sup>36</sup>

To strengthen the partner countries involvement in the evaluation process, many donors have recently reviewed their evaluation policies to align them with their Paris Declaration commitments to jointly assess aid effectiveness with partner countries.<sup>37</sup> Although many donors in theory are committed to involve their development partners in the evaluation process, in practice this commitment varies: how, who, and to what extent they involve stakeholders from the partner countries is inconsistent. The purpose of involving representatives from a partner country is to develop an evaluation that is also relevant for the host country. Therefore, relevant stakeholders from the partner country need to be involved during the planning phase leading to an evaluation to ensure that the inquiry is appropriate. Although there are some good examples of joint partner-donor evaluations of, for example, country programs', multiple donors', and multiple partners' thematic evaluations, evaluations where the partner country is truly involved are still unusual. In fact, a recent survey among the DAC members' central evaluation units indicated that only 22% of the DAC representatives thought that their own evaluations were relevant for the partner countries.<sup>38</sup>

In general, joint evaluations, whether they are between donors or between donors and partner governments, are time consuming, resource demanding, and require a strong commitment from all parties. For instance, a joint partner-donor evaluation in Tanzania took about six months longer to complete than a similar unilateral evaluation would have taken. Because representatives of the partner country sometimes lack evaluation competence, capacity building might be a required initial component of the exercise. However, more often it might be a matter of finding the right office or person within the government who is qualified to champion the joint evaluation. For example, in Benin Danida found a partner in the government office overseeing the development of the Poverty Reduction Strategy Papers (PRSP) to carry out a country evaluation. An official at the Danish embassy in Benin noted that "[w]ithout the involvement of OCS (Observatory for Social Change) we could have cancelled the evaluation mid-stream without one single stakeholder in Benin taking notice".<sup>39</sup>


In the study referred to previously, evaluation officials from the DAC members' centralized evaluation departments stated that 15% of their evaluations were joint evaluations with a partner country. However, most of the 'partner participation' did not involve representatives

36. Jensen, Rikke and Eriksson 2009.

37. See for example Swedish Agency for Public Management 2012.

38. 72% thought that the evaluations were "somewhat" of relevance, while 6% thought that it was "not at all" relevant to the partner countries. OECD 2010.

39. Jensen, Rikke and Eriksson 2009.



from the government and it was rare that partner countries were engaged in the planning, development, or follow-up phases of the evaluation process. Instead, the most common way donors incorporate country perspectives into the evaluation reports was to hire local consultants. Although local consultants add local knowledge to an evaluation, they are consultants and not representatives of the partner country. Furthermore, local consultants might be hesitant to pinpoint weaknesses in a program as they might be too closely linked to the program or wish to be hired for future assignments. An experienced rule of law evaluator pointed out that “locals [as the lead evaluator] are often too close to the program or otherwise politicized, they are often a bad bet”. The evaluator continued, “My preference is to involve them under the direction from someone from the outside”. If an evaluation’s goal is to promote joint accountability and joint learning, the relevant stakeholders from the partner country need to be involved from an early stage to ensure that the inquiry is aligned with the priorities of both parties.



## 4. LEARNING FROM EVALUATIONS

As numerous rule of law initiatives have taken place around the globe in the last decades, there has been an increase of knowledge and experience in the field. Harnessing, processing, transmitting, and absorbing this knowledge has however proven to be very challenging.<sup>40</sup> Some have argued that we are faced with an ‘evaluation gap’ as donors are not producing “enough impact evaluations and because those that are conducted are often methodologically flawed.”<sup>41</sup> There is some concern that the uneven quality of evaluation reports have diminished the interest in using and incorporating knowledge from past program experiences. Despite this valid criticism, there are numerous valuable evaluations and reviews. Still, transmitting knowledge into policy is inherently difficult with so many different levels of government agencies, intergovernmental, nongovernmental, and contractors involved in evaluating rule of law programs.<sup>42</sup> Many donor agencies have made their evaluation reports available on-line, but these on-line databases are difficult and time consuming to search and do not contain all evaluation reports. Thus, it is challenging to gain a good understanding of what activities have taken place in a particular subject area, to find out if the activity was evaluated, and then to obtain that evaluation. In fact, program officers within larger donor organizations are even finding it difficult to learn what similar programs or evaluations the agency might have conducted in different geographical or thematic areas.

### 4.1. FINDING EVALUATION REPORTS

Considering that there are so many organizations involved in rule of law assistance, it does not come as a surprise that there is no systematic way to share information among agencies. Moreover, each organization classifies rule of law differently (such as conflicts, justice, governance, human rights, and gender equality), which makes it difficult to compare efforts. What might seem less intuitive is that larger organizations do not know what their own branches or departments do within the field of rule of law. To overcome this first stumbling block (not learning from past experiences), there are some more recent internal and external efforts to map an organization’s rule of law activities. For example, the Folke Bernadotte Academy has surveyed all United Nations peace operations with a rule of law component in Africa from 1989 to 2010.<sup>43</sup> The empirical data collected put forward a descriptive account of how the rule of law has emerged as a key objective in peace operations, the main areas of rule of law assistance, and how the concept is put into practice. Another example is the World Bank’s justice sector project summaries (last updated in 2009), which provide an overview of the World Bank’s activities in the area. Ideally, such an overview would have been linked to the program documents that are available in the World Bank’s on-line database. Although such a link does not exist, a search in the database shows that most program documents are not available on-line. Of the close to 60 funded rule of law projects and programs in Africa that are listed in the summary, 12 have some kind of documentation in the on-line database; five of these are evaluations. Some of these projects are probably deemed too small to evaluate. Nevertheless, none of the background documents are made available, a situation that makes it

---

40. Carothers 2003.

41. Savedoff, Levine and Birdsall 2006.

42. Burnell 2007.

43. Zajac, Sannerholm, Möller, Simion and Hallonsten 2012.



more challenging to even learn that the project happened in the first place. Other organizations such as USAID have documented that organizations often fail to forward evaluation reports to their online database (the Development Assistance Clearinghouse) even though it is a mandatory requirement. An external review of USAID's evaluation practices was prompted by the discovery that the number of evaluations the agency carried out dropped sharply in the 1990s.<sup>44</sup> The report, following the review, made the following observation:

*[The] most straightforward explanation for this drop was that evaluation reports were not being sent in. During fieldwork, one evaluator brought back 22 reports from Africa that had not been submitted and there were others he couldn't carry.*

Thus, evaluation reports are not readily available, significantly complicating efforts to survey past experiences.

A second related issue that makes it challenging to locate past program documents is that documents are misclassified or that the database has inadequate capacity to delimit searches. For example, a search of the World Bank's project documents related to Judicial System Reform yields 183 document hits. A closer review of the hits reveals that about a quarter of the documents were filed twice, almost a third related to projects with more discreet or very limited relevance for judicial system reforms (such as an agricultural training and research project in Uganda or a hydropower rehabilitation project in Ukraine), and less than half of the documents concerned (33 projects) directly related to judicial sector reform. Most of the project documents that were submitted to the on-line database related to the initial stages of a project such as the procurement plan or project appraisal, which typically is of less interest for lessons learned. Because some documents were misfiled, they would also not show up in a search of the appropriate category. Thus, staff or other users need to be willing to spend considerable time sorting through these documents, which is a deterring factor for accumulating new knowledge. A recent study on learning among DFID staff found that employees considered the system for storing project documentation and evaluations to be difficult to search and lacked sufficient quality control.<sup>45</sup> A majority of the DFID staff found that they would look for reports they already knew about: "[. . .] otherwise most did not bother to search through and see what DFID had done on the subject".

#### 4.2. ABSORBING KNOWLEDGE FROM EVALUATIONS

Considering the international community's stated focus on demonstrating results and impact of development assistance, it seems counter intuitive that relatively little attention is paid to learning from existing evaluations. The issue of distributing and organizing evaluation reports already discussed above certainly contributes to the inadequate use of produced reports. Several larger donor organizations have recognized the learning gaps within their own organizations, so they have commissioned studies to assess organizational learning from evaluations and research reports. However, studies have found that employees of development agencies rarely find the time to read evaluation reports commissioned by their own agency and even less so by other agencies.<sup>46</sup> Some consider formal evaluation studies to be ineffective or too long and too technical to read. A study of the Norwegian Agency for Development Cooperation

44. Clapp-Wincek and Blue 2001.

45. Jones and Mendizabal 2010.

46. Ostrom, Gibson, Shivakumar and Andersson 2002.

(Norad) found that employees learned little from reading evaluation reports.<sup>47</sup> To the extent they absorb new knowledge, it related primarily to specific technical matters in the individual's field of expertise and daily work. In general, little knowledge of lessons learned from past projects and programs is transmitted by reading evaluation reports. According to one study, DFID employees found that they often had to 'reinvent the wheel' because information was not adequately transmitted during and following staff rotations. Consequently, more robust institutional memory is needed so as to pass on acquired knowledge.<sup>48</sup>

#### **SHARING OF INFORMATION THROUGH PERSONAL CHANNELS AND IN MORE INFORMAL SETTINGS**

**Sharing of information based on past professional experiences also takes place through personal channels and in more informal settings.**

A study<sup>49</sup> analysing Sida officers' learning patterns about sustainability found that one of the most important opportunities to learn was through informal discussions during lunch seminars, staff meetings, and coffee breaks. Only one department representative out of eleven mentioned evaluations as an opportunity for the department to discuss and learn about sustainability. The study also found that staff members who have moved to a new assignment maintained very limited contact with their former post and their accumulated knowledge, and insights about the projects or programs they previously managed were rarely used. Considering the frequent rotation of staff within a development agency, opportunities to informally discuss insights about a particular program during coffee breaks are limited.

For staff and other stakeholders with limited time to read evaluation reports, dissemination seminars or workshops are the preferred way to get information about the lessons learned. Face-to-face meetings and the possibility to have a dialogue about the findings is an important factor for promoting the use of the findings. Research has found that staff members are more likely to take in findings and recommendations conveyed in evaluation reports when they are involved (without interfering with the independence of an evaluation) or kept informed about the development of the evaluation process. Large joint evaluations where there are many partners and stakeholders involved during the evaluation process are anticipated by many and tend to get more attention. Danida commissioned a rather unusual follow-up assessment eight years after the very large joint evaluation 'The International Response to Conflict and Genocide: Lessons from the Rwanda Experience'. The assessment found that the evaluation report had been widely cited in academic research as well as in policy reports.<sup>50</sup> Although the assessment found that there were a number of areas in the humanitarian sector where the evaluation had a positive influence or impact, the assessment was on balance pessimistic. Many of the specific and detailed recommendations that had been put forward to prevent and suppress genocide and massive human rights abuses in the joint evaluation report remained unaddressed. Thus, even when evaluation reports are widely disseminated, read, and cited, many other factors influence whether the recommendations are considered and implemented.

When findings from evaluation reports are considered in a decision making process, they are reviewed alongside a host of competing factors, including timing, budgetary matters, and larger policy or political consideration. Employees at several agencies have found that when

47. Forss, Cracknell and Samset 1994.

48. Jones and Mendizabal 2010.

49. Ostrom, Gibson, Shivakumar and Andersson 2002.

50. Borton and Eriksson 2004.

evaluation reports are not synchronized with the agency's decision and policy cycle, the usefulness of the reports is diminished. Although donors to a greater extent strive to align their decision and funding cycles with the partner country's system, in practice, the timing of an evaluation is challenging when there are multiple stakeholders involved. Moreover, there is a certain level of path-dependence where the donors' political, geographical, and thematic interests and knowledge as well as formed relationships with partner countries and contractors are driving forces in the funding and programmatic decision-making processes.

Several studies have found that evaluations are most commonly used to legitimize on-going programs or take the more radical decision to phase out support. Although these are valuable means to employ knowledge from past experiences, a third avenue, which is less frequently mentioned, is to assess past lessons when new programs are designed. An illustrative case is the unofficial pressure within donor agencies to move and disburse money. A study conducted by Sida<sup>51</sup> found that 40% of the annual disbursements took place in the last two months of the fiscal year. Staff members were rushing to disburse funds at the end of the year due to the fear that extra funds would not be re-budgeted the following year. However, this pressure created a funding bias towards the renewal of on-going projects rather than a review of past performances of multiple other projects to initiate a new project or program based on lessons learned. Similarly, a study about learning from research and evaluations within DFID found that "considerations about on-going work and existing relationships, and the costs of a radical change of direction were too great for it to be worth drastically altering a program".<sup>52</sup>

#### 4.3. POST-EVALUATION FOLLOW-UP

Although all development agencies stress the importance of lessons learned, there is also a general, although less advertised, recognition that past lessons are not systematically taken into consideration in decision-making forums. To "deal with the weakest link in the evaluation system, i.e. the insufficient use of evaluations",<sup>53</sup> several donor agencies have instituted a Management Response (MRE) System. At Sida, where management responses are compulsory, only about half of all evaluations are followed by a MRE. A review of Sida's MRE system found that most MRE documents were incomplete and did not adequately devise a proper action plan to address the recommendations in the evaluation report. In addition, like other evaluation-related documents, MREs are difficult to access and they are therefore less likely to be used. The evaluation of Sida's MRE system found that only 15% of the MREs that were produced were available in Sida's database. Not surprisingly, MRE documents are rarely used in decision-making. Instead, their content has generally been forgotten, even by those who were involved. The Sida study found that the MRE system has a low status compared with other administrative routines. In addition the study found that because the employees had few incentives to complete a MRE, the MRE had limited value for transferring knowledge.

51. Ostrom, Gibson, Shivakumar and Andersson 2002.

52. Jones and Mendizabal 2010.

53. Hanberger and Gisselberg 2006.

## MANAGEMENT RESPONSE (MRE) SYSTEM

**The MRE system, aimed to promote action-oriented learning, is a written administrative procedure where the donor agency responds to an evaluation's findings and recommendations.**

In the MRE, the development agency is supposed to clearly articulate what actions it will take to address findings and recommendations from the evaluation in future decision-making processes.

The concept behind MREs – to reflect on findings and propose actions for future undertakings – is valuable. However, for a staff member, the tangible benefits of the MRE process are more apparent when the evaluation of a completed program is followed by a new initiative in a related thematic (and geographic) area. For example, an evaluation was conducted following the completion of a large Australian Agency for International Development (AusAID) program aimed at strengthening the capacity of Indonesian government and civil society institutions to promote legal reforms and the protection of human rights. In the management response, AusAID addressed the evaluation's concerns and recommendations. As AusAID was designing a new law and justice sector program in Indonesia, the MRE described how shortcomings and recommendations for improvements would be addressed in the new program. For the new program, measures were taken to ensure alignment with the Indonesian government's reform efforts and priorities to develop a gender strategy and increase gender expertise and employ a full-time monitoring and evaluation advisor throughout the full program period.<sup>54</sup> It is worth noting that AusAID provides a short description of each evaluation report and links it to the management response, which makes relevant documentation easier to find hence the likelihood of them being used greater.<sup>55</sup>

### 4.4. SPOTLIGHT SERIES

To address concerns that evaluation reports are too long and time consuming to read, several development agencies have started to produce evaluation briefs and thematic synthesis reports. Many evaluation reports are longer and contain more technical detail than most people might initially need; for example, the European Commission contracted a thematic evaluation of its support to justice and security sector reforms. The evaluation report is very comprehensive and covers almost a decade of European commission justice and security sector programming worldwide, but it is also rather long (Volume I, the main report, is 139 pages). To encourage greater internal and external use of evaluation reports, several agencies have started to produce and disseminate shorter evaluation summaries. For example, in 2011 the former Swedish Agency for Development Evaluation (SADEV) launched an evaluation brief series. One of the evaluation series' briefs provides a two-page summary of the main findings of a larger evaluation of Sida's support to justice in reconciliation processes. The brief refers those interested to the full report if they want to learn more about recommended measures to strengthen planning, implementation, and monitoring of justice and reconciliation

54. Australian Aid (undated. Likely 2010 or 2011).

55. See for example, [http://www.ausaid.gov.au/Publications/Pages/5364\\_9850\\_7825\\_4786\\_5011.aspx](http://www.ausaid.gov.au/Publications/Pages/5364_9850_7825_4786_5011.aspx).

programming. However, several of these evaluation spotlight initiatives have not been sustained and kept up-to-date. For example, in 2002 the Canadian International Development Agency (CIDA) launched an evaluation brief series called 'What We're Learning', but there are only five summaries available on-line and the most recent one is from 2007. Similarly, UNDP launched an 'Essentials' series synthesizing evaluation findings, but it has not been updated since 2004. In addition, the United Nations Office on Drugs and Crime's evaluation briefs series contain only four summaries between 2006 and 2008, and the World Bank produced an 'Influential Evaluation' initiative that only resulted in one publication. DFID produces a very useful evaluation summary series (EVSUM) that contains the main findings and recommendations as well as a summary of the management response. However, there is no dedicated page or portal for these evaluation summaries, making them difficult to find, defeating the purpose of producing evaluation briefs. The various efforts to highlight past experiences by making the evaluation findings easy to digest and access have so far fallen short of that goal.

#### 4.5. SUCCESS STORIES AND ADMITTING FAILURE

To learn from past experiences, it is important to examine what worked well as well as what did not work so well. In development assistance, there is often emphasis on 'success stories', and many agencies and contractors highlight inspiring anecdotal stories about achievements on their websites and promotional material.<sup>56</sup> USAID even has a template for contractors to submit success stories to be posted on the USAID website. Following the development agencies' focus on results-based management, there is a great interest in showcasing positive results while failures are not highlighted. Some concern has been raised that the desire to produce good results is driving what projects are being evaluated and what results are presented in the final edited reports:

*[W]here 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 be 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.<sup>57</sup>*

In the 1990s, USAID allocated mission budgets according to reported progress; however, this practice caused the field missions to avoid evaluations, as they did not want to draw attention to less successful projects and risk receiving reduced budgets. Although USAID discontinued the attempted practice to link budget allocations to performance, there is still a persistent desire to showcase success stories while failures are kept out of the limelight.

A study examining learning at DFID found that employees learned little from failed programs and few people wanted to be associated with such a program.<sup>58</sup> The study found that DFID rarely drew on lessons learned from what did not work or from difficulties implementing certain programs. To acknowledge a failed project or that the underlying assumptions for the project were flawed often equals professional suicide. Therefore, most criticisms of programs come from observers in academia or think tanks not directly involved in program implementation. Observers from the outside do not have the same access to programmatic details and intimate familiarity with working routines, but a couple of organizations have

56. Jones and Mendizabal 2010.

57. Burnell 2007.

58. Jones and Mendizabal 2010.

recently made their own ‘failures’ public. Engineers without Borders – Canada has produced an annual failure report since 2008 where they highlight mistakes and failures and analyse how they can learn and improve. Other examples include the Centre for Court Innovation’s reflections on their challenges, missteps, and failures in improving the criminal justice system, and the Peace Dividend Trust, a peacebuilding and humanitarian non-profit contractor, who produced a failure report on their own organizational failures.

Evaluations are also valuable if they include constructive analysis of what went wrong and what steps could be taken to mitigate an identified issue in a future project. The evaluation of a small joint pilot project to strengthen security in a refugee camp in Guinea provides an unusual but very interesting example of an evaluation outlining how little impact and how fatally wrong almost everything went in the project. The evaluation provides detailed insights into the many faults and oversights that occurred. Thus, the evaluation of the project constitutes a good lesson learned of what did not work and what issues need to be addressed or all together avoided in a three-way partnership (UNHCR, Canada, and Guinea) in a humanitarian intervention.

#### **JOINT EVALUATION BY THE GOVERNMENT OF CANADA AND UNHCR OF THE DEPLOYMENT OF RCMP OFFICERS TO REFUGEE CAMPS IN GUINEA**

**To strengthen the protections of refugees, the Government of Canada and United Nations High Commissioner for Refugees (UNHCR) collaborated on a project where two Canadian police officers at a time were deployed to train its Guinean counterparts. Although this was a small pilot, the administrative arrangements of the collaboration between Canada and UNHCR became very involved and when an agreement could finally be reached on the terms of references for the mission, the political and security environment had changed in Guinea.<sup>59</sup>**

By the time the Canadian police officers were deployed, their Guinean counterparts, who during the planning of the project were police officers, were now largely illiterate civilians without any police or security training. Thus, the primary focus of the deployment “to hone the technical skills of the BMS [the security force] in basic procedures, community policing methods and humanitarian issues, became a distant objective for these irregular security personnel”. As a result, the planned training modules were no longer relevant. In addition, there was limited buy-in from the Guinean government, and the logistical arrangements and communication between the Canadian police officers and the UNHCR was irregular and cumbersome.

Due to the changed circumstances, the training the Canadian police could offer was no longer critical to the UNHCR office, which therefore lost interest in their services. The Canadian police officers became discouraged and felt “that their time would be better used back in Canada”. Following the end of the mission, the Guinean leadership for refugee security ignored or was even hostile towards the few Guineans who had been trained by the Canadian Police. In addition, “a few loose sheets found in a dusty cabinet”, the concrete imprints of the project, were close to non-existent, and left most people involved in the planning and execution of the mission frustrated.

Although pretty much nothing went as intended and the mission could “only be described as a disappointing, although a well-intentioned demonstration of international solidarity”, the evaluators believed that a similar endeavour was worth trying again.

59. United Nations High Commissioner for Refugees Evaluation and Police Analysis Unit 2005.



## 5. CONCLUDING RECOMMENDATIONS

The numerous rule of law initiatives that have taken place around the globe in the last decades have resulted in an increase of knowledge and experience in the field. Evaluations therefore play an important role in transmitting information about past experiences. The mixed quality of existing evaluations and the shortcomings in transmitting knowledge about past projects and evaluations however contribute to the weak uptake of lessons learned. To improve this transmission of knowledge, donors need to formulate more realistic, specific, and at times more modest goals.

Many rule of law programs support complex changes within a legal system and to expect that drastic changes will be observed as a result of the intervention is often to hope for too much. What is important is that the impact of the intervention, whether it is large or small, is assessed. The perception that rule of law interventions are difficult to evaluate could be moderated if the objectives are adjusted to be less over-reaching and confounding factors are considered.

Disseminating and incorporating lessons learned into new or on-going program development remain one of the more pressing issues that need to be addressed by the rule of law community. Rule of law programming is scattered over many agencies and implementers and it is very challenging to gain a good overview of what activities are taking place or have been tried in the past. This report focuses primarily on practices within certain unilateral or multilateral donor agencies. There is little documentation on how contractors, local and international NGOs, or partner countries engage with evaluations and acquire knowledge about lessons learned.

### THE PEACEBUILDING COMMUNITY COULD PLAY AN IMPORTANT ROLE IN HELPING BRIDGE THE KNOWLEDGE GAP BY DOING THE FOLLOWING:

- › **Review past evaluation.** This review should highlight both what worked well, as well as what did not work so well. To conduct such a review, donors would have to be consulted to gain access to evaluation reports and find out more about the project that might not have made it into the evaluation. Such an engagement with donors would provide an opportunity to solicit input and interest in the anticipated publication, which would increase the chances that the publication will be consulted and lessons learned shared.
- › **Conduct an empirical case study.** Such a case study should explore issues on how contractors, local and international NGOs, or partner countries engage with evaluations and acquire knowledge about lessons learned as the specific project design and implementation of particular programs is carried out by contractors and NGOs.
- › **Incorporate funding for evaluations in projects and programs award allocations.** These funds should help projects and programs gain knowledge of what has worked or not worked in the field, when new rule of law assistance programs and projects are conceived and designed, and make sure lessons learned from problematic projects are not ignored or forgotten.

- › **Make underlying assumptions of a program explicit.** These assumptions should be made explicit during the planning stages so that subsequent evaluations of the same program can link the project activities in a coherent and logical way to the outputs, outcomes and overall objectives of the program.
- › **Create a database of evaluations.** These databases should make evaluation reports with valuable lessons learned more accessible as the currently inadequate ways evaluation reports are collected and disseminated seriously limit the possible use of lessons learned.
- › **Use evaluations to demonstrate results of rule of law programs.** The use of these evaluations should help grant-awarding institutions demonstrate the results of the rule of law programs and projects and avoid programs and projects that receive continued support without knowing whether development assistance has been carried out in an accountable and effective manner.
- › **Ensure that evaluations are taken into consideration in decision-making forums and policy development.** Such a strategy should ensure that past lessons are systematically taken into consideration in decision-making forums and that policy is based on lessons learned. This strategy recognizes that transmitting lessons learned into policy is inherently difficult with so many levels of government agencies and intergovernmental and non-governmental contractors involved in evaluating rule of law programs.
- › **Enhance the active use of evaluation findings and recommendations.** Encourage the use of evaluation findings when designing new projects and programs.
- › **Incorporate research series on past donor experiences in professional training of rule of law practitioners.** This research series should help improve the learning from existing evaluations and avoid learning gaps within organizations as practitioners rarely find the time to read lengthy evaluation reports.
- › **Support a research series.** This research should review and synthesize past donor experiences in particular areas of the rule of law such as judicial training and customary justice.



## 6. BIBLIOGRAPHY

Accra Agenda for Action, 2008

Armytage, Livingston, 2006, *Monitoring Performance of Legal and Judicial Reform in International Development: Early Lessons from Port Moresby & Phnom Penh*, Presentation at the International Bar Association Chicago Showcase, Judicial Reform: Economic Development and the Rule of Law, 18 September 2006  
<http://www.educatingjudges.com/Hyperlinks/MonitoringImpactIBA.pdf>  
(Last accessed February 26, 2013)

Australian Agency for International Development, undated (likely 2010 or 2011), *Indonesia-Australia Legal Development Facility Independent Completion Report-Management Report*, [http://www.usaid.gov/au/Publications/Pages/5364\\_9850\\_7825\\_4786\\_5011.aspx](http://www.usaid.gov/au/Publications/Pages/5364_9850_7825_4786_5011.aspx)  
(Last accessed February 26, 2013)

Bandstein, Sara and Erik Hedblom, 2008, *Reaping the Fruits of Evaluation? An Evaluation of Management Response Systems within Aid Organizations*, SADEV Report 2008:7, Karlstad, Sweden: Swedish Agency for Development Evaluation

Bennett, Jon, S Pantuliano, W Fenton, A Vaux, C Barnett and E Brusset 2010, *Aiding the Peace - Multi-donor Evaluation of Support to Conflict Prevention and Peace building Activities in Southern Sudan 2005-2010*, The ITAD Ltd., United Kingdom

Bennett, Jon, William Bertrand, Clare Harkin, Stanley Samarasinghe and Hermantha Wickramatillake, 2006, *Coordination of International Humanitarian Assistance in Tsunami-affected Countries*, Joint Evaluation 2006:3, Stockholm: Swedish International Development Cooperation Agency

Bergling, Per, 2006, *Rule of Law on the International Agenda: International Support to Legal and Judicial Reform in International Administration, Transition and Development Co-operation*, Intersentia, Antwerpen

Berman, Greg and Aubrey Fox, 2010, *Trial and Error in Criminal Justice Reform: Learning from Failure*, Washington DC: Urban Institute Press

Bollen, Kenneth, Pamela Paxton and Rumi Morishima, 2005, 'Assessing International Evaluations: An Example From USAID's Democracy and Governance Program', *American Journal of Evaluation*, Vol. 26, No. 2, pp. 189-203

Borton, John, Margie Buchanan Smith and Ralf Otto, 2005, *Support to Internally Displaced Persons: Learning from Evaluation. Synthesis Report of a Joint Evaluation Programme*, Sida Joint Evaluation 2005:1, Stockholm: Swedish International Development Cooperation Agency

Borton, John and John Eriksson, 2004, *Assessment of the Impact and Influence of the Joint Evaluation of Emergency Assistance to Rwanda*, Copenhagen: Evaluation Department, Ministry of Foreign Affairs, Denmark

Botero, Juan Carlos, Robert L. Nelson & Christine Pratt, 2011, 'Indices and Indicators of Justice, Governance, and the Rule of Law: An Overview', *The Hague Journal on the Rule of Law*, Vol. 3, Issue 2, pp. 153–169

Burnell, Peter, 2007, 'Methods and Experiences of Evaluating Democracy Support: a Moving Frontier', pp. 15–46 in Peter Burnell (ed.), *Evaluating Democracy Support - Methods and Experiences*, Sida Joint Evaluation 2007:3, Stockholm: International Institute for Democracy and Electoral Assistance and Swedish International Development Cooperation Agency

Burrows, Sally, 2009, *Closing the Learning Loop – Harvesting Lessons from Evaluations: Report of Phase 1*, Rome: World Food Programme

Busan Outcome Document, High Level Forum on Aid Effectiveness, 2011

Cahn, Miranda, 2010, *NZAID Programme 2009 Evaluations and Reviews - Annual Report on Quality*, Wellington: New Zealand Ministry of Foreign Affairs and Trade, [http://www.aid.govt.nz/webfm\\_send/45](http://www.aid.govt.nz/webfm_send/45) (Last accessed February 26, 2013)

Carothers, Thomas (ed.), 2006, *Promoting the Rule of Law Abroad: In Search of Knowledge*, Carnegie Endowment for International Peace, Washington

Carothers, Thomas, 2003, 'Promoting the Rule of Law Abroad: the Problem of Knowledge', *Carnegie Endowment for International Peace Working Paper*, Nr 34, Washington, DC: Carnegie Endowment for International Peace

Cissner, Amanda B. and Donald J. Farole, Jr., 2009, *Avoiding Failures of Implementation-Lessons from Process Evaluations*, New York: Center for Court Innovation

Clapp–Wincek, Cynthia and Richard Blue, 2001, *Evaluation of Recent USAID Evaluation Experience*, Washington DC: United States Agency for International Development

Cohen, Elin, Kevin Fandl, Amanda Perry-Kessaris and Veronica Talyor, 2011, 'Truth and Consequences in Rule of Law: Inferences, Attribution and Evaluation,' *The Hague Journal on the Rule of Law*, Vol. 3, Issue 1, pp. 102–129

Council of the European Union, 2011, *EU Concept for CSDP Justice Missions (within the Rule of Law framework)* (18173/10)

Dakolias, Maria and Bruce Owen, 2003, *Impact of Legal Aid: Ecuador*, Washington DC: The World Bank

Davis, Kevin, Benedict Kingsbury and Sally Merry, 2012, 'Indicators as a Technology of Global Governance', *Law & Society Review*, Vol 46

Davis, Kevin and Michael B. Kruse, 2007, *Taking the Measure of Law: The Case of the Doing Business Project*, *Law and Social Inquiry*, Vol. 32, Issue 4, pp. 1095–1119

Department for International Development, 2010, *Working Effectively in Conflict-affected and Fragile Situations*, Briefing Paper I: Monitoring and Evaluation, <http://www.dfid.gov.uk/Documents/publications1/governance/building-peaceful-states-I.pdf> (Last accessed February 26, 2013)

Department for International Development, 2009, *Building the Evidence to Reduce Poverty – The UK's Policy on Evaluation for International Development*, London: Department for International Development

Department for International Development, 2006, *Evaluation of DFID's Policy and Practice in Support of Gender Equality and Women's Empowerment*, EVSUM EV669 report, London: Department for International Development, <http://www.dfid.gov.uk/Documents/publications1/evaluation/ev669-summary-report.pdf> (Last accessed February 26, 2013)

Edgren, Gus, 2004, *Stronger Evaluation Partnerships: The Way to Keep Practice Relevant*; Sida Studies in Evaluation 04/01, Stockholm: Swedish International Development Cooperation Agency

Elena, Sandra and Héctor Chayer, 2007, 'Progress and Myths in the Evaluation of the Rule of Law: a Toolkit for Strengthening Democracy', pp. 95-118 in Peter Burnell (ed.), *Evaluating Democracy Support – Methods and Experiences*, Sida Joint Evaluation 2007:3, Stockholm: International Institute for Democracy and Electoral Assistance and Swedish International Development Cooperation Agency

Engineers without Borders – Canada, 2011, *Failure Report 2011*, <http://legacy.ewb.ca/en/whoweare/accountable/failure.html> (Last accessed February 26, 2013)

Eriksson, John, Howard Adelman, John Burton, Hanne Christensen, Krishna Kumar, Astri Suhrke, David Tardif-Douglin, Stein Villumstad and Lennart Wohlgemuth, 1996, *The International Response to Conflict and Genocide: Lessons from the Rwanda Experience, Synthesis Report*, Steering Committee of the Joint Evaluation of Emergency Assistance to Rwanda, Sida Joint Evaluation 1996:1, Stockholm: Swedish International Development Cooperation Agency

European Commission, 2011, *Thematic Evaluation of European Commission Support to Justice and Security System Reform*, [http://ec.europa.eu/europeaid/how/evaluation/evaluation\\_reports/reports/2011/1295\\_vol1\\_en.pdf](http://ec.europa.eu/europeaid/how/evaluation/evaluation_reports/reports/2011/1295_vol1_en.pdf) (Last accessed March 19, 2012)

Finkel, Steven, 2003, 'Can Democracy be Taught?', *Journal of Democracy*, Vol. 14, Issue 4, pp. 137-151

Forss, Kim, Basil Cracknell and Knut Samset, 1994, 'Can Evaluation Help an Organization to Learn?', *Evaluation Review*, Vol. 18, pp. 574-591

Fox, Aubrey and Emily Gold (eds.), 2010, *Daring to Fail: First-person Stories of Criminal Justice Reform*, New York: Center for Court Innovation

Freeman, Ted and Sheila Dohoo Faure, 2003, *Local Solutions to Global Challenges: Towards Effective Partnership in Basic Education*, The Hague: Netherlands Ministry of Foreign Affairs

Gilmore, Scott, 2010, *PDT's Failure Report 1.0*, posted on December 15, 2010, <http://buildingmarkets.org/blogs/blog/2010/12/15/pdts-failure-report-1-0/>, (Last accessed February 26, 2013)

Ginsburg, Tom, 2011, 'Pitfalls of Measuring the Rule of Law', *The Hague Journal on the Rule of Law*, Vol. 3, Issue 2, pp. 269–280

Golub, Stephen, 2006, 'A House Without a Foundation,' in Thomas Carothers (ed.), *Promoting the Rule of Law Abroad: The Search for Knowledge*, Carnegie Endowment for International Peace, Washington

Hageboeck, Molly, 2009, *Trends in Development Evaluation Theory, Policy and Practice*, Washington DC: United States Agency for International Development

Hammergren, Linn, 2003, 'International Assistance to Latin American Justice Programs', pp. 290–335 in Erik Jensen and Thomas Heller (eds.), *Beyond Common Knowledge: Empirical Approaches to the Rule of Law*, Palo Alto: Stanford University Press

Hammergren, Linn, 2002, *Assessments, Monitoring, Evaluation, and Research: Improving the Knowledge Base for Judicial Reform Programs*, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.197.1114> (Last accessed February 26, 2013)

Hanberger, Anders and Kjell Gisselberg, 2006, *Sida's Management Response System*, Sida Studies in Evaluation 06/01, Stockholm: Sida's Department for Evaluation and Internal Audit

Hildenwall, Viktoria, Peter Sjöberg and Inger Wikström Öbrand 2008, *Challenges, Opportunities and Approaches for Increasing Joint Donor Programming of Evaluations*, SADEV Report 2008:8, Karlstad, Sweden: Swedish Agency for Development Evaluation

International Development and Law Organization, 2010, *Legal and Judicial Development Assistance: IDLO's Global Overview of Trends, Actors, Projects and Funding Strategies*, Rome

International Development and Law Organization, International Dialogue on Peacebuilding and Statebuilding, A NEW DEAL for engagement in fragile states

Jensen, Rikke Ingrid and John Eriksson, 2009, *Experiences with Conducting Evaluations Jointly with Partner Countries*, Evaluation Study 2009/3, Copenhagen: Danida's Evaluation Department

Jones, Harry and Enrique Mendizabal, 2010, *Strengthening Learning from Research and Evaluation: Going with the Grain*, London: Department for International Development

Kleinfeld, Rachel, 2005, *Competing Definitions of the Rule of Law: Implications for Practitioners*, Carnegie Papers No. 55, Washington, DC: Carnegie Endowment for International Peace

Mensa-Bonsu, Isaac and Margrethe Holm Andersen, 2010, 'Involving Developing Countries in the Evaluation of Development Cooperation: The Case of Joint Evaluation of Development Cooperation between Ghana and Denmark', *Evaluation*, Vol. 16, pp. 295-308

Messick, Rickard, 1999, 'Judicial Reform and Economic Development: A Survey of the Issues', *The World Bank Research Observer*, Vol. 14, Issue 1, pp. 117-136

Molund, Stefan and Göran Schill, 2007, *Looking Back, Moving Forward: Sida Evaluation Manual*, Stockholm: Swedish International Development Cooperation Agency, Department for Evaluation and Internal Audit

New Zealand Agency for International Development, 2008, *Annual Report on Quality*

Ngugi, Joel M., 2005, 'Policing Neo-Liberal Reforms: The Rule of Law as an Enabling and Restrictive Discourse', *The University of Pennsylvania Journal of International Law*, Vol. 26, pp. 513-599

Norwegian Agency for Development Cooperation, 2011, *Joint Evaluation of Support to Anti-Corruption Efforts 2002-2009*, Report 6/2011, Oslo, Norway: Norwegian Agency for Development Cooperation

Organization for Economic Cooperation and Development, 2010, *Evaluation in Development Agencies, Better Aid*, Paris: OECD Publishing

Organization for Economic Cooperation and Development, 2006, *Joint Evaluation of General Budget Support Synthesis Report*, Paris: Organization for Economic Cooperation and Development,  
<http://www.oecd.org/dataoecd/25/43/37426676.pdf> (Last accessed March 15, 2012)

Organization for Economic Cooperation and Development / Development Assistance Committee, 2008, *Guidance on Evaluating Conflict Prevention and PeaceBuilding Activities working draft for Application Period*, [http://www.oecd.org/secure/pdfDocument/0,2834,en\\_21571361\\_34047972\\_39774574\\_1\\_1\\_1\\_1,00.pdf](http://www.oecd.org/secure/pdfDocument/0,2834,en_21571361_34047972_39774574_1_1_1_1,00.pdf) (Last accessed February 26, 2013)

Organization for Economic Cooperation and Development / Development Assistance Committee, 2006, *Guidance for Managing Joint Evaluations*,  
<http://www.oecd.org/dataoecd/29/28/37512030.pdf> (Last accessed February 26, 2013)

Organization for Economic Cooperation and Development / Development Assistance Committee, 2002, *Glossary of Key Terms in Evaluation and Results Based Management*, Paris: Organization for Economic Cooperation and Development, <http://www.oecd.org/dataoecd/29/21/2754804.pdf> (Last accessed February 26, 2013)

Ostrom, Elinor, Clark Gibson, Sujai Shivakumar and Krister Andersson, 2002, *Aid, Incentives, and Sustainability: An Institutional Analysis of Development Cooperation*, Sida Studies in Evaluation 02/01, Stockholm: Swedish International Development Cooperation Agency

Paris Declaration on Aid Effectiveness, 2005

Parson, Jim, Besiki Kutateladze, Monica Thornton, April Band and Aminou Yaya, 2010, 'Justice Indicators for Post-Conflict Settings: A Review', *The Hague Journal on the Rule of Law*, Vol. 3, Issue 2, pp. 170-185

Peerenboom, Randall, 2002, 'Let One Hundred Flowers Bloom, One Hundred Schools Contend: Debating Rule of Law in China', *Michigan Journal of International Law*, Vol. 23, pp. 471-544

Perry-Kessaris, Amanda, 2011, 'Prepare your Indicators: Economics Imperialism on the Shores of Law and Development' *International Journal of Law in Context*, Vol. 7, Issue 4, pp. 401-421

Perry-Kessaris, Amanda, 2001, *Legal Systems as a Determinant of FDI, Lessons from Sri Lanka*, The Hague; Boston: Kluwer Law International

Poate, Derek, Roger Riddell, Nick Chapman and Tony Curran, 2000, *The Evaluability of Democracy and Human Rights Projects: A Logframe-related Assessment*, Sida Studies in Evaluation 00/3, Stockholm: Swedish International Development Cooperation Agency

Rosten, Keith, Natalija Djurickovic and Louise D. Williams, 2006, *Evaluation of the Justice System Reform Activity- Kosovo*, Washington DC: United States Agency for International Development

Savedoff, William, Ruth Levine and Nancy Birdsall, 2006, *When Will We Ever Learn? Improving Lives through Impact Evaluation*, Washington DC: Center for Global Development

Special EU Programmes Body, 2007, *A Monitoring and Evaluation Framework for Peace-Building*, [http://www.seupb.eu/Libraries/PEACE\\_III\\_Reports\\_Pubs/A\\_Monitoring\\_and\\_Evaluation\\_Framework\\_for\\_Peace-building.sflb.ashx](http://www.seupb.eu/Libraries/PEACE_III_Reports_Pubs/A_Monitoring_and_Evaluation_Framework_for_Peace-building.sflb.ashx) (Last Accessed February 26, 2013)

Stoddard, Abby, 2005, *Review of UNICEF's Evaluation of Humanitarian Action*, New York: United Nations Children's Fund

Swedish Agency for Public Management, 2012, *Utvärdering av svenskt bistånd: En översyn av utvärderingsverksamheten*, 2012:17

Swedish Agency for Public Management, 2011, *Styrning av svensk biståndspolitik: En utvärdering*, 2011:25



Swedish Agency for Development Evaluation, 2010, *Sida's Preparation of and Follow-up on Support to Justice in Reconciliation Processes*, SADEV Evaluation Brief 2010:1, Karlstad, Sweden: Swedish Agency for Development Evaluation

Swedish International Development Cooperation Agency, 2006, *Management Response to Returning Home- Evaluation of Sida's Integrated Area Programmes in Bosnia and Herzegovina*, <http://www.sida.se/Documents/Import/pdf/Management-Response.pdf> (Last accessed February 26, 2013)

Taylor, Veronica, 2011, *Regulating Rule of Law: Branding, Certification and Quality Assurance*, paper presented at Folke Bernadotte Academy's Rule of Law Research Group's Annual Meeting, Canberra December 6, 2011

Taylor, Veronica, 2010, 'The Rule of Law Bazaar', in Bergling, Ederlof & Taylor (eds.), *Rule of Law Promotion: Global Perspectives, Local Applications*, Iustus Förlag, Uppsala

Trubek, David and Alvaro Santos (eds.), 2006, *The New Law and Economic Development: A Critical Appraisal*, Cambridge

United Nations High Commissioner for Refugees Evaluation and Police Analysis Unit, 2005, *Joint Evaluation by the Government of Canada and UNHCR of the Deployment of RCMP Officers to Refugee Camps in Guinea*, EPAU/2005/013, New York: United Nations High Commissioner for Refugees

United Nations Secretary-General, 'Rule of Law and Transitional Justice in Conflict and Post Conflict Societies', S/2004/616, 23 August 2004

United States Agency for International Development, 2011, *Learning from Experience: USAID Evaluation Policy*, Washington DC: United States Agency for International Development

United States Agency for International Development, 2009, *Trends in Development evaluation theory Policy and Practice*, Washington DC: United States Agency for International Development

Weiss, Carol Hirschon, 1995, 'Nothing As Practical As Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families', pp. 65-92 in James I. Connell, Anne C. Kubisch, Lisbeth B. Schorr and Carol H. Weiss (eds.), *New Approaches to Evaluating Community Initiatives - Concepts, Methods, and Contexts*. Washington DC: The Aspen Institute

Wood, Bernard, J Betts, F Etta, J Gayfer, D Kabell, N Ngwira, F Sagasti and M Samaranayake, 2011, *The Evaluation of the Paris Declaration, Final Report*, Copenhagen: Danish Institute for International Studies, Available at: [www.oecd.org/dac/evaluationnetwork/pde](http://www.oecd.org/dac/evaluationnetwork/pde) (Last accessed February 26, 2013)

World Bank, 2010, *Clustered Project Performance Assessment Report, Ecuador Judicial Reform Project (Loan 4066); Guatemala Judicial reform Project (Loan 4401); Colombia Judicial Conflict Resolution Improvement Project (Loan 7081)*, Report No: 55277, Washington DC: The World Bank

World Bank, 2008, *Implementation Completion and Results Report (IRD-44010) on a Loan in the Amount of US \$33.00 Million Equivalent to the Republic of Guatemala for a Judicial Reform Project*, Report No: ICR0000623, Washington DC: The World Bank

World Bank, 2007, *Implementation Completion and Result Report (Cr. 3384-CE), on a Credit in the Amount of SDR 13.6 Million (US\$ 18.2 Million Equivalent) to Democratic Socialist Republic of Sri Lanka for a Legal and Judicial Reform Project*, Report No: 39538, Washington DC: The World Bank

World Bank Independent Evaluation Group, 2008, *Doing Business: An Independent Evaluation – Taking the Measure of the World Bank-IFC Doing Business Indicators*, Washington DC: The World Bank

World Bank, *Justice Sector Project Summaries-Africa Region*,  
Available at <http://go.worldbank.org/HKWFR43L0> (Last accessed February 26, 2013)

Zajac Sannerholm, Richard, Frida Möller, Kristina Simion and Hanna Hallonsten, 2012, *UN Peace Operations and Rule of Law in Africa 1989-2010: Data, Patterns and Questions for the Future*, Folke Bernadotte Academy Publications





**FOLKE  
BERNADOTTE  
ACADEMY**

SWEDISH AGENCY FOR PEACE, SECURITY AND DEVELOPMENT

**Folke Bernadotte Academy**

Sandö – Stockholm, Sweden  
Phone: +46 612 823 00

[www.folkebernadotteacademy.se](http://www.folkebernadotteacademy.se)