

Holzappel, Sarah

**Working Paper**

## Boosting or hindering aid effectiveness? An assessment of systems for measuring agency results

Discussion Paper, No. 31/2014

**Provided in Cooperation with:**

German Institute of Development and Sustainability (IDOS), Bonn

*Suggested Citation:* Holzappel, Sarah (2014) : Boosting or hindering aid effectiveness? An assessment of systems for measuring agency results, Discussion Paper, No. 31/2014, ISBN 978-3-88985-657-9, Deutsches Institut für Entwicklungspolitik (DIE), Bonn

This Version is available at:

<https://hdl.handle.net/10419/199444>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

d.i.e

Deutsches Institut für  
Entwicklungspolitik



Years | 1964 – 2014

German Development  
Institute

Discussion Paper

31/2014

# Boosting or Hindering Aid Effectiveness?

## An Assessment of Systems for Measuring Agency Results

*Sarah Holzapfel*

Boosting or hindering aid effectiveness?  
An assessment of systems for measuring  
agency results

Sarah Holzapfel

Bonn 2014

Discussion Paper / Deutsches Institut für Entwicklungspolitik  
ISSN 1860-0441



Die deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available at <http://dnb.d-nb.de>.

ISBN 978-3-88985-657-9

**Sarah Holzapfel**, Gesellschaft für Internationale Zusammenarbeit (GIZ), Advisor, & German Development Institute / Deutsches Institut für Entwicklungspolitik (DIE), Associate Researcher

E-mail: [sarah.holzapfel@die-gdi.de](mailto:sarah.holzapfel@die-gdi.de)

© Deutsches Institut für Entwicklungspolitik gGmbH  
Tulpenfeld 6, 53113 Bonn  
 +49 (0)228 94927-0  
 +49 (0)228 94927-130  
E-Mail: [die@die-gdi.de](mailto:die@die-gdi.de)  
[www.die-gdi.de](http://www.die-gdi.de)

# Contents

## Abbreviations

<b>Summary</b>	<b>1</b>
<b>1 Introduction</b>	<b>5</b>
<b>2 Background to the study</b>	<b>6</b>
2.1 The results agenda	6
2.2 Challenges of results measurement	10
<b>3 Criticisms of the results agenda and adverse effects</b>	<b>12</b>
3.1 Competing accountability demands	12
3.2 Costs and adverse effects of accountability for results	15
<b>4 Standard indicators: differences between bilateral and multilateral donor organisations</b>	<b>17</b>
4.1 Summary of results measurement systems and use of results information	18
4.2 Levels of the results chain and the inclusion of qualitative results	19
4.3 Attribution vs. contribution	22
4.4 Performance assessment	24
<b>5 Limitations and risks of using standard indicators for results reporting</b>	<b>26</b>
5.1 Limitations	26
5.2 Risks and adverse effects	28
<b>6 Conclusion and policy recommendations</b>	<b>29</b>
<b>Bibliography</b>	<b>33</b>
<b>Annex</b>	<b>39</b>

## **Figures**

Figure 1: Results chain	8
Figure 2: Competing accountability demands: impossible geometries?	13
Figure 3: Performance measurement	15
Figure 4: Reporting on results in agency results frameworks	19

## **Tables**

Table 1: Development agencies reviewed: use of standard indicators for agency-wide performance measurement	17
Table 2: Levels of the results chain (reporting on contributions to long-term development goals)	22
Table 3: Use of baselines and targets: comparison of development agencies	26

## **Tables in Annex**

Table A: List of interviewees	41
-------------------------------	----

## Abbreviations

ADB	Asian Development Bank
AfDB	African Development Bank
AusAID	Australian Agency for International Development
BP	Busan Partnership for Effective Development Co-operation
CESR	Center for Economic and Social Rights
CIDA	Canadian International Development Agency
CPIA	Country Policy and Institutional Assessment
DAC	Development Assistance Committee
DFID	Department for International Development
DG DevCo	Directorate-General for Development and Cooperation
DOTS	Development Outcome Tracking System
EC	European Commission
EU	European Union
GNI	Gross National Income
IDA	International Development Association
IDB	Inter-American Development Bank
IDG	IFC Development Goal
IEG	Independent Evaluation Group
IFC	International Finance Corporation
IPC	Integrated Food Security Phase Classification
MDB	Multilateral development bank
MDG	Millennium Development Goal
M&E	Monitoring and evaluation
MfDR	Managing for Development Results
NPM	New public management
ODA	Official development assistance
OECD	Organisation for Economic Co-operation and Development
OHCHR	Office of the United Nations High Commissioner for Human Rights
PPP	Purchasing power parity
PBA	Programme-based approach
RBM	Results-based management
UK	United Kingdom
UN	United Nations
UNDG	United Nations Development Group
UNDP	United Nations Development Programme
USAID	United States Agency for International Development





## Summary

Amid rising criticism of aid effectiveness coupled with tight budgets in many traditional donor countries at a time of economic crisis, donor agencies are under pressure to deliver more value for money and to provide evidence of the positive effects of development cooperation. In response to these pressures, more and more development agencies are adopting agency results frameworks for monitoring and managing their progress in pursuing their strategic objectives and for reporting on performance.

So as to provide the public with a snapshot of their contributions to overarching development goals, they use a selected set of standard indicators to aggregate results across interventions and countries. While greater accountability for results thanks to the use of standard indicators can sharpen the focus on results and bolster public support for development cooperation, there are certain drawbacks in that the emphasis lies on providing accountability to taxpayers and donors only, and also in that it may have an adverse effect on aid effectiveness. For example, there is a risk that the accountability demands of donors and taxpayers may push development agencies to focus on short-term results to the detriment of long-term results that are not immediately visible and often harder to achieve. Moreover, it is not easy to measure results in development cooperation; certain unintended effects may arise if the indicators chosen to assess performance are imperfect measures or only reflect partial aspects of an underlying objective.

This paper analyses and compares the results reporting practices of ten bilateral and multilateral donor agencies in order to assess their implications for aid effectiveness. It examines the limitations, risks and adverse effects of results measurement systems and makes recommendations for designing indicator systems that minimise risk.

There are several differences between donor agencies' approaches to agency-wide results measurement. These affect whether the standard indicators used for reporting on aggregate results are useful as both an accountability and a management tool. They also influence the likelihood of adverse effects occurring and determine whether results measurement drives or hinders aid effectiveness. There are differences in terms of the levels of the results chain at which standard indicators are defined, in relation to how the results of interventions are reported that focus on qualitative changes, and in terms of how the issue of attribution vs. contribution is addressed and how performance is assessed (i.e. with the aid of baseline and target values).

The analysis shows that the data provided by donor agencies on agency-wide results resulting from the use of standard indicators is generally only of limited informational value and does not provide an adequate basis for holding the donor agencies to account. The majority of standard indicators are formulated at output level or at the level of short-term outcomes, and it is difficult to assess whether these contribute to longer-term development results. Moreover, results frameworks capture only a very small proportion of total results. Usually, only a very small number of indicators, i.e. between 20 and 40, are used so as to preserve the snapshot character and not to overwhelm the target audience. In order nonetheless to capture the results of a variety of interventions, the indicators tend to be very broadly defined and do not adequately reflect qualitative differences between results. Moreover, the results of interventions designed to support policy reforms, institution-building or capacity-building in partner countries are not adequately reflected by results frameworks because it is difficult to standardise and quantify such results.

Using standard indicators for accountability and management purposes in spite of their limitations may have certain adverse effects. For example, by focusing on selected results that can be measured with standard indicators, donor agencies run the risk of neglecting outputs and outcomes that are not reported at an organisational level. Donor agencies may also end up prioritising activities that deliver a particularly high volume of results with the fewest resources or that reach an especially high number of beneficiaries. Efforts to deliver value for money may also play a role in reducing the use of partner-country systems if results can be achieved more efficiently by donor-implemented projects.

Additionally, the practice of attributing development results, as is followed by most of the agencies reviewed, clashes with the principle of ownership by partner countries because it does not acknowledge the contributions made by partner countries and other development partners to the measured results. The practice of attributing results can also have an adverse effect on donor harmonisation and the use of country systems. This is because attribution is generally more challenging where interventions are implemented and financed by more than one development partner.

Another risk arises if agencies set agency-wide targets for results measured by standard indicators from the top down. If the demand for support from partner countries does not correspond with the priorities reflected by results targets, it may be more difficult to align support with partner countries' priorities. Moreover, agency-wide targets generally increase the risk of adverse effects because they raise the pressure on development agencies to deliver results.

To reduce the risk of adverse effects resulting from agency-wide results measurement and to ensure that accountability mechanisms drive rather than hinder aid effectiveness, I suggest that donor agencies adopt the following recommendations in designing indicator systems for measuring agency-wide results:

- Donor agencies should also report on the results of interventions that seek to bring about qualitative changes through policy reform, capacity-building and institution-building at agency level. Although the indicators used for this purpose may be imperfect given the difficulty of standardising such results, this would reduce the risk of overemphasising quantitative and short-term results.
- Donor agencies should report on the results achieved with their support (i.e. to which they have contributed) rather than directly attribute results to their own engagement. The attribution of results is methodologically challenging and often not feasible when reporting on agency-wide results. In addition, attribution may be seen as contrary to the principle of country ownership and as hindering donor harmonisation.
- Targets should not be set for standard indicators at agency level because these may conflict with efforts to align aid to partner countries and may raise the risk of adverse effects.
- Donor agencies should harmonise the definitions of standard indicators and reporting methodologies. This would reduce the cost of coordination among donor agencies carrying out joint projects and programmes. It would also reduce the overall cost of data collection and monitoring and lessen the burden on partner countries' monitoring systems.

Given the limitations and risks inherent to agency-wide results measurement systems identified in this paper, I would also suggest that donor agencies should explore complementary options or

alternatives to standard indicators in order to meet their reporting requirements. For example, donor agencies are advised to invest more in rigorous impact evaluations, to increase the transparency of individual interventions and to incorporate beneficiary feedback more systematically into their monitoring and evaluation efforts.



## 1 Introduction

The international aid system has pushed for greater accountability and a clearer results orientation during the past two decades. Despite being seen in the past as “*angelic deliverers of urgently needed sustenance*” (Ramalingam 2013, 106), aid agencies today are under increasing public scrutiny and must demonstrate evidence of the positive effects of development cooperation. There are numerous reasons for this. Huge aid flows over the past 50 years have failed to produce substantial economic growth and reduce poverty in developing countries. On the contrary, in sub-Saharan Africa (where the biggest aid recipients are located), per capita incomes actually fell between the 1970s and the mid-1990s due to low economic growth coupled with high population growth (Lancaster 1999, 478).

Against this background, many observers have criticised aid as being ineffective or even as acting as a brake on economic growth (Doucouliagos / Paldam 2009; Lensink / White 2011; Easterly 2007).<sup>1</sup> Aid is often said to support corrupt and incompetent governments and to hamper much needed reforms and democratisation (Nuscheler 2008, 6; Radelet 2006, 9). On the donor side, a lack of transparency, the excessive fragmentation of aid, and ineffective aid practices such as tied aid and food aid, have all been singled out for criticism (Easterly / Pfutze 2008, 29). The recent economic crisis, which caused many traditional donor countries to tighten their aid budgets, has increased pressure on development agencies to make aid more effective and to allocate resources more efficiently (OECD/DAC 2008a, 6–7; Eyben 2013, 23).

In response to concerns about the effectiveness of aid and in order to demonstrate ‘value for money’,<sup>2</sup> a growing number of bilateral and multilateral development agencies have adopted results frameworks for monitoring and managing progress in pursuing strategic objectives and for reporting on performance. These results frameworks involve the use of a selected set of standard indicators for aggregating intervention results across countries in order to provide the public with an overview of their contributions to overarching development goals. The idea is that the higher level of accountability produced by the use of standard indicators will help to sharpen the focus on results, boost aid effectiveness and bolster public support for development cooperation.

While everyone is of course in favour of disseminating results, there is mounting criticism of the ‘results agenda’. In practice, the latter is often geared more towards providing accountability to funders than towards using results information for management and learning purpose so as to improve aid effectiveness (Natsios 2010, 2–3; Barder 2012). In addition, many commentators have pointed out that the focus on results may have adverse effects and could conflict with the commitments made as part of the aid effectiveness

---

1 Doucouliagos and Paldam (2009) found that aid had been ineffective in fostering economic growth. This finding was, however, challenged in a later study by Mekasha and Tarp (2013), who re-examined the main hypothesis put forward by Doucouliagos and Paldam. After expanding Doucouliagos and Paldam’s meta-analysis to better reflect the statistical and data challenges, Mekasha and Tarp found that aid had had a positive, significant impact on growth.

2 Value for money may be defined as “*the optimum combination of whole-life cost and quality (or fitness for purpose) to meet the user’s requirement. It can be assessed using the criteria of economy, efficiency and effectiveness*” (Jackson 2012, 1).

agenda to strengthen country ownership, harmonise with other donors and align aid with partner countries' priorities (Sjöstedt 2013; Hudson / Jonsson 2009).

Various studies have addressed the risks inherent to an excessive focus on results and accountability in development cooperation (see for example Ramalingam 2013; Cook et al. 1995; Vielajus et al. 2009; Vähämäki / Schmidt / Molander 2011; Natsios 2010; Savedoff 2011; Eyben 2013; Sjöstedt 2013). To date, however, no studies have been published specifically on the implications of agency-wide results measurement systems. This paper aims to address this research gap by analysing the implications of using standard indicators for reporting on an agency's aggregate contributions to overarching development goals.

There are two main reasons why it is important to understand the effects of agency-wide results measurement systems:

- (1) First, the pressure to account for results shapes development cooperation practices and is one of the determinants in deciding which results are prioritised by donor agencies.
- (2) It is important to be aware of the potential unintended effects of agency-wide results measurement systems in order to minimise the risks involved and to ensure that accountability mechanisms help boost aid effectiveness.

For the purpose of this analysis, I drew on the current literature on accountability, public-sector performance measurement and aid effectiveness. Based on a combination of desk research and a series of semi-structured interviews, I compared the experiences of ten selected multilateral and bilateral donor agencies with standard indicators in reporting on their aggregate contributions in results frameworks.

This paper is formatted as follows. The next chapter sets out the background to the study. It summarises the results agenda and accountability mechanisms in development cooperation and discusses the challenges encountered in measuring results. Chapter 3 explains the potential risks and adverse effects of results measurement and reporting, and analyses the competing accountability demands of aid funders and recipients. Chapter 4 explores the differences among donor agencies' approaches to the use of standard indicators as part of results frameworks for measuring agency-wide performance. Chapter 5 discusses the limitations of and the risks associated with the use of standard indicators in the light of the theoretical background set out in chapter 3. Finally, a number of conclusions are drawn and policy recommendations made in chapter 5.

## **2 Background to the study**

### **2.1 The results agenda**

The results agenda in development cooperation can be seen as an attempt to make aid more effective and to make development actors more accountable by measuring and reporting on development results (Shamash / Burall / Whitty 2013, 12). There are two main driving factors behind the results agenda.

The first is the adoption of the Millennium Development Goals (MDGs) in 2000. For the first time, the international community set global goals, targets and indicators for reducing the many dimensions of extreme poverty for the period to 2015 (OECD/DAC 2008a, 7; MfDR n.d.). The adoption of the MDGs created a need to measure progress towards the goals. Massive investments were made in data collection and in statistical capacity-building in developing countries. At the International Conference on Financing for Development in Monterrey in 2002, the development community acknowledged that a substantial increase in aid was needed to help developing countries achieve the MDGs. At a time when aid effectiveness was the butt of growing criticism, the donor countries' pledge to raise their official development assistance (ODA) to 0.7% of their gross national income (GNI) (UN 2002; Health and Development Networks n.d.) meant that they would have to answer more critical questions about whether funds had been efficiently spent and whether they had helped to achieve certain global development goals (Nuscheler 2008, 5).

Donor agencies therefore also pledged to increase aid effectiveness and took various steps to enhance their results-based management and to measure, monitor and report on results throughout the development process (MfDR n.d.; Kabba et al. 2002). In a series of High Level Fora on Aid Effectiveness in Paris (2005), Accra (2008) and Busan (2011) that followed the International Conference on Financing for Development, donor agencies undertook to 'manage for results'. This means that aid implementation should be guided by the desired results and that information obtained from monitoring should be used to improve decision-making (OECD/DAC 2005/2008, 7; 19–20; BP 2011).

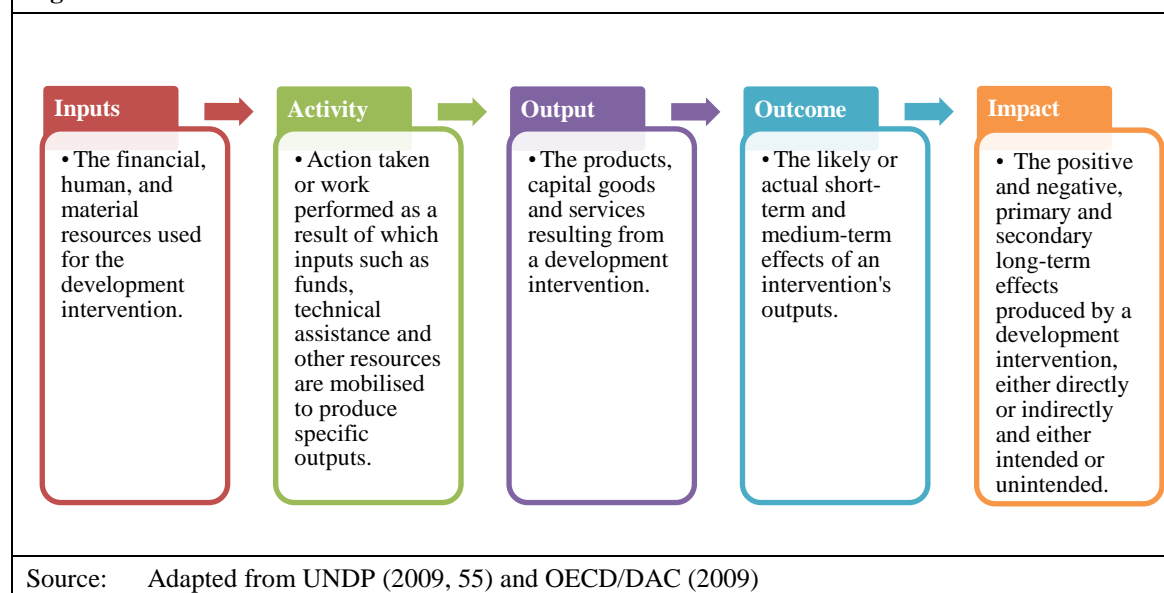
The second driving factor is the efforts made by the members of the Organisation for Economic Co-operation and Development (OECD) to improve public-sector performance by introducing a system of new public management (NPM). NPM was first introduced in the 1980s and is characterised by a shift in emphasis from process accountability (i.e. have inputs been used in accordance with the rules?) towards accountability for results (i.e. has the money been spent efficiently and effectively?) (Saltmarshe / Ireland / McGregor 2003, 23; Eyben 2013, 13). One of the main reasons for this change was that process accountability trapped public-sector managers in inflexible structures and routines that stifled innovation and experience (Posner 2006, 81). Moreover, it was argued that, if accountability mechanisms focused too much on inputs and processes as opposed to results, public-sector organisations might lose sight of their overall goals (Schacter 1999, 1).

While NPM may be seen as a global phenomenon, there are substantial disparities among countries in terms of the rate of adoption of NPM practices and the impact of the NPM reforms. Anglophone countries (e.g. Australia, New Zealand, the United Kingdom and the United States) are leading the reform agenda (Bach / Bordogna 2011, 2290; Halligan 2010, 84) and have also been quickest to apply NPM to development cooperation agencies.

One important element of NPM that was widely adopted in development cooperation in the 1990s is results-based management (RBM). RBM is a life-cycle approach to planning, monitoring and evaluation with longer-term development outcomes and impacts as its main focus (Meier 2003, 6; UNDP 2009, 7–10). During the planning stage, objectives are formulated, indicators are identified for measuring progress and targets are set for each indicator. Once a project, programme or strategy is launched, progress is monitored, and results are reviewed and reported. Evaluations can complement monitoring by providing a

more rigorous and ideally independent assessment of progress. The information obtained from monitoring and evaluation (M&E) is important as a basis for taking corrective action and for generating lessons and recommendations (UNDP 2009, 7–10; UNDG 2011, 2).

**Figure 1: Results chain**



At the centre of RBM stands the results chain (see Figure 1). This is a logical and sequential model of the steps that need to be taken in order to achieve the desired objectives or results. It starts with inputs such as the funding allocated to an intervention, followed by the activities undertaken to achieve the desired goals. The resources invested result in direct outputs which contribute to short-term and medium-term outcomes and long-term impacts (Meier 2003, 6–7; OECD/DAC 2009, 42–43).

The purpose of introducing RBM in development cooperation is to improve the efficiency and aid effectiveness by:

- (1) using results information for internal learning and management; and
- (2) raising public accountability for the results of development cooperation (Binnendijk 2000, 9; Meier 2003, 6; UNDP 2009, 7–10).

Quantitative information on results can be used to find out what does and does not work by measuring and comparing the performance of different delivery and performance options (Pidd 2005, 488).

Public accountability has two dimensions (Pidd 2005, 488):

- First, development agencies must demonstrate that development cooperation has delivered important results and that these results represent value for money.
- Second, it is sometimes argued that the availability of data on results allows donors to decide which development agency or project to support. For example, an increasing



number of bilateral donors<sup>3</sup> carry out multilateral performance assessments and base funding decisions on results information. This is done in order to demonstrate to their domestic constituencies that aid for multilateral organisations has been well spent (Obser 2007, 5). In the same way, data on results may allow beneficiaries to choose which services to use, given that alternatives available.

For most agencies, the introduction of RBM practices has meant going beyond their traditional focus on inputs, activities and outputs, to create a system of internal incentives that focuses on outcomes, and to develop new monitoring and reporting systems that measure results and performance (MfDR n.d.). RBM can be implemented at different organisational levels, i.e. project, country programme and agency level (Binnendijk 2000, 11). For a long time, however, RBM was confined to the project and country programme levels due to problems in aggregating development results across interventions and countries as a measure of agency performance (Binnendijk 2000, 79–80). It was not until the late 1990s that the first attempts were made by bilateral donor agencies (e.g. the United States Agency for International Development, the Department for International Development, the Australian Agency for International Development and Danida) to undertake strategic planning and performance measurement at agency level, often in response to government legislation and executive orders (Binnendijk 2000, 79). The United States, for instance, adopted the Government Performance and Results Act in 1993 obliging all government agencies, including the United States Agency for International Development (USAID), as from 1999, to prepare an annual performance plan covering each programme activity and to set performance indicators that would enable their performance in any given year to be compared with the goals set (Office of Management and Budget 1993; Cook et al. 1995, 1303). Multilateral agencies adopted agency-wide performance measurement systems around the same time, in response to pressure from donor countries to increase aid effectiveness and to become more accountable for results (IDA 2002, 1; Bester 2012, 8; Binnendijk 2000, 81).

Results frameworks are the most common format for reporting on agency-wide performance. A results framework is one of the core elements of RBM. It “*explains how the development objective is to be achieved, including causal relationships and underlying assumptions*” (OECD/DAC 2009, 43). At agency level, results frameworks are used to guide and measure the progress made by agencies in achieving their objectives (Roberts / Khattri 2012, 17).<sup>4</sup> The top of the framework sets out the agency’s overarching objectives (e.g. inclusive growth and the transition towards green growth<sup>5</sup>). To monitor progress at this level, country-level outcome and impact indicators are used for which data is available in international statistical databases (Binnendijk 2000, 81). An agency’s performance and its contribution to overarching objectives are measured at lower levels.

---

3 For example, the Department for International Development (DFID) assessed multilateral organisations in its multilateral aid review. The aim was to ensure that UK aid delivered the maximum value for money, based on evidence of its contribution to development results (DFID 2011, 112–113).

4 Agency results frameworks are based on a private-sector approach to performance measurement known as the balanced scorecard. Balanced scorecards were originally introduced in the private sector as a tool for aligning a company’s short-term activities with its long-term objectives (Kaplan / Norton 2007, 1). They provide managers with a snapshot of a company’s health that goes beyond purely financial measures (Kaplan / Norton 1992, 71).

5 AfDB strategic goals (AfDB Bank Group 2013b, 7).

Standard indicators,<sup>6</sup> i.e. indicators that share a commonly accepted definition, method of measurement and interpretation (Holzapfel 2014, 26), are the most important instrument for reporting on contributions. Since these allow results to be aggregated across interventions and partner countries, they can be used to present a snapshot of an agency's contributions to longer-term development objectives.

Reporting on aggregate results by means of standard indicators is still not common practice in development cooperation, however. The multilateral development banks (MDBs) and Anglophone donor agencies are leading the results agenda and have the most advanced results measurement systems. However, aggregate results reporting is growing in significance among development agencies. For example, the United Nations Development Programme (UNDP) recently introduced a new results framework as part of its strategic plan for 2014-2017. This uses standard indicators to report on aggregate intervention results (UN 2013a). Likewise, EuropeAID is currently developing a results framework based on the MDB model (EC 2013). Various other donors (e.g. Switzerland and Germany) are considering introducing similar results measurement systems so that they can report on their aggregate contributions.

## 2.2 Challenges of results measurement

Compared with many other areas of public policy, the development cooperation industry was relatively slow to take up NPM philosophies such as RBM that focus on accountability for results (Saltmarshe / Ireland / McGregor 2003, 446). This can be explained by the difficulty of holding individual development actors to account for results and also of measuring performance in development cooperation.

Accountability may be defined as *“the mechanisms through which people entrusted with power are kept under check to make sure that they do not abuse it and that they carry out their duties effectively”* (de Renzio / Mulley 2006, 1). Accountability thus implies answerability, i.e. the obligation of accountors to report on their actions and to explain and justify their decisions, and enforcement, i.e. the ability of accountees to impose sanctions (Schedler 1999, 14–16). For accountability mechanisms to function, certain conditions must be met. The basis of answerability is transparency, i.e. the availability of information on compliance and performance. Answerability also depends on the capacity of accountors to analyse and review the available information (de Renzio / Mulley 2006, 1). Enforcement requires mechanisms to sanction poor performance or the abuse of power and adequate incentives for compliance (de Renzio / Mulley 2006, 1; Eyben 2008, 11). In addition, accountability works best if accountors have clearly defined duties and performance standards as these allow for the transparent and objective assessment of their behaviour (OHCHR / CESR 2013, 10). For example, clearly defined goals, targets and indicators as part of RBM at project, country programme and agency level can help to specify responsibilities. This will improve answerability and incentivise actors to perform better (OHCHR / CESR 2013, xii).

---

6 Some development agencies (e.g. the European Commission and the Millennium Challenge Corporation) prefer the term “common indicator” to “standard indicator”. Standardised key indicators for different sectors (e.g. education, infrastructure and climate change) are often referred to as “core sector indicators” (see World Bank 2013, 6).

The above conditions are not met in full by the accountability mechanisms used for reporting on results in development cooperation. There are various reasons for this:

- First, there is a structural problem with accountability mechanisms in development cooperation. Those who pay for aid cannot directly observe whether it has been spent effectively and those who receive aid often do not have the power or the means to hold donor agencies to account (Ramalingam 2013, 107-108). For this reason, and because aid usually only represents a very small fraction of a country's budget, there is less pressure to deliver than in other public sectors, such as health or education, which are of greater concern to taxpayers in donor countries (Saltmarshe / Ireland / McGregor 2003, 446).
- Second, while individual development agencies can be held responsible for specific tasks and outputs, they cannot be held responsible for the achievement of long-term development goals such as the MDGs (Easterly 2006, 3). Although a number of development agencies and other actors may contribute to them, the achievement of long-term development goals is mainly in the hands of partner countries. This creates problems of attribution when measuring and reporting on the results of individual interventions or actors (Saltmarshe / Ireland / McGregor 2003, 446). In addition, since the achievement of the MDGs is a collective responsibility of the international community, it creates only weak incentives for individual agencies to perform (Easterly 2006, 3).
- Third, complexity theory suggests that accountability for results in development cooperation is possible only to a limited extent due to the complexity of problems. Performance measurement and accountability for results work well if there is broad agreement on the nature of the problem and a common understanding of how it might be solved, and if the problem is bound in terms of the time and resources required for its resolution (Eyben 2005, 101; Chapman 2002, 36; Pidd 2005, 486). Problems in development cooperation are, however, only rarely of this kind and it is difficult to identify and agree on indicators to measure performance. For example, objectives such as "poverty reduction" or "sustainable and inclusive growth" are very broad and there is considerable uncertainty as to how improvements could be made and about the time and resources needed to achieve them. Whether accountability for results is practical in development cooperation also depends on the type of intervention as some are less complex than others. Interventions for delivering goods and services (e.g. immunisations, distribution of bed nets and the construction of schools and roads) are less complex and relatively easy to quantify and measure. By contrast, the results of interventions for building local self-sustaining institutions (e.g. through staff training, development of organisational procedures and institutional cultures) or for policy dialogue and reform (e.g. in the form of discussions with a variety of stakeholders) are more difficult to measure and are also often less visible (Natsios 2010, 4). They require a longer time horizon until results can be observed and their success depends on the cooperation of local institutions and actors. This makes them not only riskier and less predictable (Natsios 2010, 4) but also, because of the long time horizon and the difficulty of attributing results, less suitable for performance measurement.
- Fourth, it is especially difficult for development agencies to report on agency-wide performance. Unlike most public agencies, development agencies are usually active in a multitude of sectors such as healthcare, education, agriculture and infrastructure

and also perform a wide variety of interventions in each sector. This makes it extremely difficult to set up a performance measurement system that captures the diversity of results and at the same time provides the public with a snapshot of the overall contributions (Binnendijk 2000, 80).

### **3 Criticisms of the results agenda and adverse effects**

The results-based management practices of development agencies have been criticised by a growing number of academics and development practitioners as having three potentially adverse effects. First, development agencies might be pulled in different directions by a range of stakeholders all demanding accountability. Second, there might be a conflict of interests between aid funders and recipients (ODI 2011; Hudson / Jonsson 2009, v). Third, there are certain intrinsic drawbacks to using performance measurement and accountability for results.

#### **3.1 Competing accountability demands**

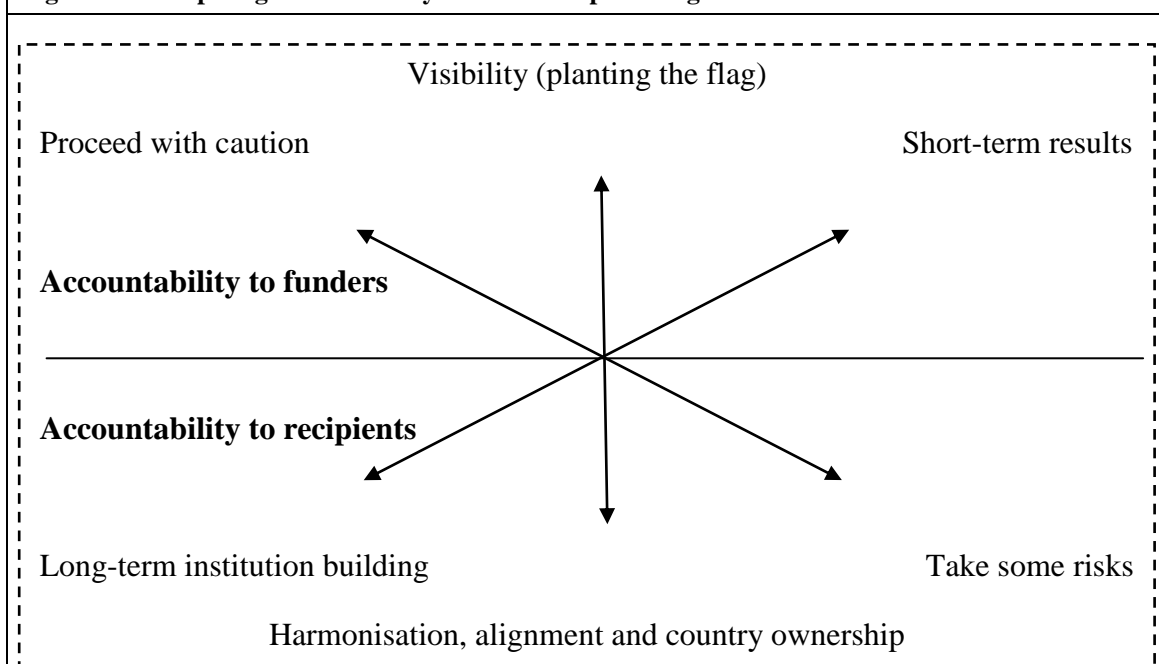
Accountability mechanisms in development cooperation are complex because development involves multiple stakeholders, many of whom are accountable to different actors. Basically, there are two type of accountability: “upward” accountability to funders and “downward” accountability to clients (Edwards / Hulme 1996, 967). While donor agencies are accountable to their respective governments, shareholders, parliaments and citizens as well as to partner-country governments and aid beneficiaries, partner-country governments have to answer to donor agencies and their citizens (OECD/DAC 2005/2008, 8; Schacter 2001, 1). These different levels of accountability in the aid system are not necessarily complementary and may conflict with one another (Klingebiel 2012, 2). The accountability demands of funders may push development agencies in directions that are not fully in line with the interests of partner countries (Hudson / Jonsson 2009, 17). They may also conflict with commitments made as part of the aid effectiveness agenda to promote country ownership, align efforts with the priorities set by developing countries, implement common arrangements and simplify procedures (harmonisation) (Hudson / Jonsson 2009, 17; Sjöstedt 2013, 144).<sup>7</sup>

De Renzio (2014) put forward the concept of “impossible geometries” (see Figure 2) to describe the competing accountability demands made by aid recipients and funders.

---

<sup>7</sup> Development agencies have to report to partner countries on the progress made against these commitments, as part of mutual accountability commitments (BP 2011, 3; OECD/DAC 2005/2008, 8).

**Figure 2: Competing accountability demands: impossible geometries?**



Source: Adapted from de Renzio (2014)<sup>8</sup>

Different tensions may arise for development agencies from competing accountability demands. There is a risk that accountability demands from funders could encourage donor agencies to pursue short-term results to the detriment of longer-term results that are not immediately visible and often harder to achieve (Barder 2012; Smith 1995, 289). For instance, short-term goals, such as an increase in the net enrolment rate in primary schools, can often be achieved relatively easily by paying for the delivery of services. However, this may conflict with the long-term goal of strengthening local institutions so that they can provide and maintain such services after donor support ends (Savedoff 2011, 7).

Similarly, pressure by funders to demonstrate quick, visible wins may lead to resources being shifted to projects and programmes where results are easy to measure and produce, i.e. low-hanging fruit. Consequently, those interventions, sectors and regions may be neglected in which results are not as easy to measure, take longer to unfold or are more difficult to attain. Some commentators have argued that the shift may be to the detriment of those areas where aid is most needed or is most effective in the long term (Cook et al. 1995, 1305; Vielajus et al. 2009, 63; Vähämäki / Schmidt / Molander 2011, 38; Natsios 2010, 7–10). Natsios (2010, 3) even goes so far as to claim that *“those development programs that are most precisely and easily measured are the least transformational, and those programs that are most transformational are the least measurable.”*

<sup>8</sup> In the original figure, de Renzio (2014) places the concept of “division of labour” opposite to “planting the flag”. I would argue that there is an even greater conflict with the principles of harmonisation, alignment and country ownership. The division of labour still allows donor agencies to be visible (albeit in fewer areas). The implementation of the principles of harmonisation, alignment and the country ownership, on the other hand, requires donor agencies to lower the flag, while giving greater recognition to the results of joint efforts and the partner country’s achievements.

A focus on accountability to funders may also produce more risk aversion and less innovation in the aid system. For instance, for the case of USAID, Natsios (2010, 35–36) shows that pressure to produce rapid and measurable results has created a culture of risk aversion at USAID. Staff members are increasingly reluctant to innovate and experiment because a failed programme or a negative audit report may end their careers. At country level, aid effectiveness principles such as “local ownership” and “use of country systems”<sup>9</sup> are being compromised to reduce the risk of money being squandered (Wood et al. 2011, 24; Natsios 2010, 3). This is especially evident in the debate on budget support, which is often said to be an instrument with potentially high returns but high risks. On the one hand, budget support gives partner countries greater ownership, lowers transaction costs, reduces fragmentation and is often said to be more effective than traditional modes of delivering aid (Koeberle / Stavreski 2006, 3). On the other hand, it increases fiduciary risks in countries with weak public financial management systems because donors are less able to monitor and control how funds are used (Shand 2006, 27).

There is also a conflict between the need felt by many donor agencies to be visible and report on results they can claim for themselves (i.e. “planting the flag”) and commitments to greater harmonisation with other donors, country ownership and alignment (Vähämäki / Schmidt / Molander 2011, 23). Programme-based approaches (PBAs),<sup>10</sup> in which donors make funds or other inputs available to an – ideally coherent – programme of development activities across a particular sector or area led by the partner country (OECD 2011a), are key to implementing the three above commitments. However, donor agencies have made only limited efforts to increase the use of common arrangements and procedures.<sup>11</sup> This has been due in part to “*donor headquarter insistence on their distinctive channels for reasons of visibility*” (Wood et al. 2011, 27). Another reason is that donor agencies are hesitant to rely on common reporting systems, in particular on partner countries' monitoring systems, because it creates problems of attribution (Wood et al. 2011, 36; Ministry of Finance Nepal 2010, 37). Attribution may be defined as “*the extent to which observed development effects can be attributed to a specific intervention or to the performance of one or more partners, taking account of other interventions, (anticipated or unanticipated) confounding factors, or external shocks*” (OECD/DAC 2009, 21) and is generally more challenging in PBAs. Because donors pool their resources and work together to achieve development goals, there is no direct link between the inputs provided by a particular donor and the results delivered by the joint efforts of donors and partner countries (Schacter 2001, 9).

---

9 According to the Paris Declaration for Aid Effectiveness (OECD/DAC 2005/2008, 14), the term country systems includes, but is not restricted to “*national arrangements and procedures for public financial management, accounting, auditing, procurement, results frameworks and monitoring*”.

10 According to the OECD/DAC (2008b, 2), PBAs share the following characteristics: (1) leadership by the host country or organisation; (2) a single, comprehensive programme and budget framework; (3) a formalised process for donor coordination and harmonisation of donor procedures for reporting, budgeting, financial management and procurement; (4) efforts to increase the use of local systems for programme design and implementation, financial management, monitoring and evaluation.

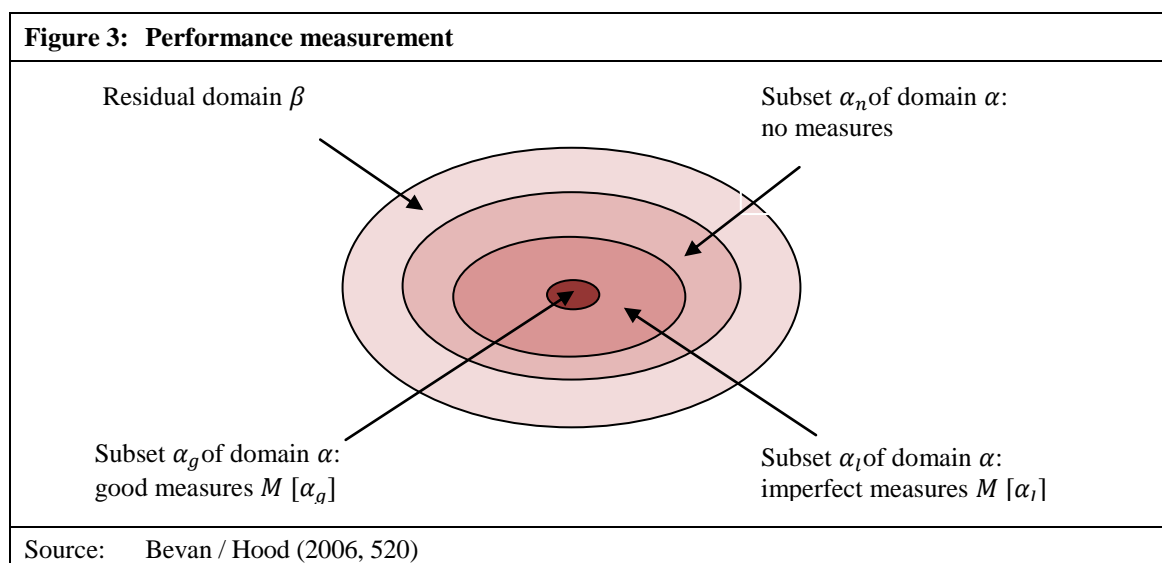
11 The Paris Agenda for Aid Effectiveness stipulates that 66% of aid flows in the period to 2010 should be provided in support of PBAs (OECD/DAC 2005/2008, 10). However, the target was not met: only 45% of aid flows in 2010 were delivered in the form of PBAs (OECD 2011b, 19).

### 3.2 Costs and adverse effects of accountability for results

There are two main types of performance measurement costs. The first is the cost of setting up and maintaining a performance measurement system. Many researchers have criticised this type of cost as being excessively high and crowding out the resources and time available for planning and implementing development interventions (Vähämäki / Schmidt / Molander 2011, 24; Natsios 2010, 5). The second type consists of the cost of unintended and often undesirable outcomes that may lower performance (Pidd 2005, 483). Unintended outcomes may arise if there is a lack of congruence between the agent's goals, as defined by the performance measurement scheme, and those of the principal (Smith 1995, 283).

The risk of unintended effects is especially high if targets are set for performance indicators. Bevan and Hood (2006, 520) show that two assumptions underlie governance by targets. The first is synecdoche, which means that it does not matter if there are measurement problems and that a part that is measurable may be taken as standing for the whole (Bevan / Hood 2006, 520–521). The second is that performance targets change the behaviour of individuals and organisations for the better, while the risk of ‘gaming’ can be kept low. Gaming in this context is defined as a “reactive subversion” of performance measurement (Bevan / Hood 2006, 521).

Figure 3 shows that setting performance targets involves prioritising among several aspects of performance and that indicators chosen to measure progress against targets often present only an inadequate and incomplete picture of total performance. Domains  $\alpha$  and  $\beta$  in Figure 3 represent two different aspects of total performance. Since performance measurement is costly, only a subset of total performance is usually measured. Here domain  $\alpha$  is prioritised over domain  $\beta$ , for which targets are not set and to which less importance is therefore attached. Selecting adequate indicators to measure performance is challenging and good measures may not always be available. In this Figure, only one subset  $\alpha_g$  of domain  $\alpha$  can be captured by good measures  $M[\alpha_g]$ , while another, subset  $\alpha_l$ , can be reflected only by imperfect measures  $M[\alpha_l]$ , which could be misleading. Subset  $\alpha_n$  is not measured by performance indicators because there are no data with which to assess performance.



The assumptions underlying target-based performance measurement are hence that omitting  $\beta$  and  $\alpha_n$  does not matter and that either  $M[\alpha_g]$  or  $M[\alpha_g] + M[\alpha_l]$  can be relied upon (Bevan / Hood 2006, 521). However, these assumptions may not always be valid and ‘gaming’ may produce unintended effects.

Smith (1995) identifies the two unintended effects of tunnel vision and measure fixation. Tunnel vision is defined “*an emphasis on phenomena that are quantified in the performance measurement scheme, at the expense of unquantified aspects of performance*” (Smith 1995, 284). Tunnel vision may occur where organisational objectives diverge from what is captured by the performance measurement system (Smith 1995, 283). Measure fixation is defined as an “*emphasis on measures of success rather than the underlying objective*” (Smith 1995, 290) and occurs because complex phenomena can be measured only imprecisely (Smith 1995, 283). If a performance measure does not capture all the aspects of an underlying objective, individuals or organisations are encouraged to pursue strategies that raise the values of the indicators they have to report upon rather than those which help to achieve the associated objective (Smith 1995, 290; Ramalingam 2013, 106). Pidd (2005, 487) argues that, in such cases, performance is “*virtual rather than real*”, because performance indicators do not reflect an organisation’s true objective.

Both problems are likely to arise in development cooperation. Development interventions or strategies usually have a large number of diverse objectives. It is impractical or impossible to quantify all of them. A selected set of key indicators must be chosen that reflect the most important objectives. Moreover, many aspects of development cooperation, such as the results of governance interventions, simply cannot be quantified adequately. Residual domain  $\beta$  and the subsets of domain  $\alpha$  in a performance measurement system, for which either imperfect or no measures exist ( $\alpha_l$  and  $\alpha_n$ ), are therefore both likely to be large.

Performance measurement may also have the unintended effect of incentivising individuals or organisations to manipulate data under their control so as to present themselves, their project or organisation in the best possible light (Smith 1995, 292; Delorme / Chatelain 2011, 3). Such data misrepresentation may have various adverse effects, such as the misallocation of resources (if resource allocation decisions are based on manipulated performance data) or the unequal treatment of actors (Smith 1995, 293).

Some commentators also argue that the use of performance measurement adversely influences the incentives and motivations of individuals or organisations. Fry (1995, 181) points out that management principles such as “what gets measured gets done” rely on the notion that the desire to do good is always externally motivated. However, this ignores the internal dimension of accountability, i.e. the “felt responsibility” of agents (Fry 1995, 181–182), which implies that staff are intrinsically motivated to serve their organisation’s ideas and strategies. The internal dimension of accountability implies that actors have a need to answer the question “Did we act as effectively as possible?” (Guijt 2010, 283) in order to uphold public trust (Ebrahim 2003, 194). Related to this, Bogart (1995, 165) argues that monitoring may have a negative effect on individuals’ motivation. People may work harder if they are monitored less closely because they want to show that they deserve trust.



#### 4 Standard indicators: differences between bilateral and multilateral donor organisations

There are certain differences between the results measurement systems used by bilateral and multilateral development agencies. These differences affect the usefulness of standard indicators as accountability and management tools. They also influence the likelihood of adverse effects occurring, and whether results measurement drives or hinders aid effectiveness. This chapter compares the experiences of ten selected agencies (see Table 1) in order to provide an overview of the variety of approaches and attempts. The agencies reviewed in this chapter were selected because they have considerable experience in agency-level results measurement. They include seven multilateral agencies and three bilateral agencies.

<b>Table 1: Development agencies reviewed: use of standard indicators for agency-wide performance measurement</b>	
<b>Multilateral agencies</b>	<b>Bilateral agencies</b>
<ul style="list-style-type: none"> <li>- African Development Bank (AfDB)</li> <li>- Asian Development Bank (ADB)</li> <li>- European Commission (Development and Cooperation Directorate-General – EuropeAID)<sup>12</sup></li> <li>- International Finance Corporation (IFC)Inter-American Development Bank (IDB)</li> <li>- United Nations Development Programme (UNDP)</li> <li>- World Bank</li> </ul>	<ul style="list-style-type: none"> <li>- Australian Agency for International Development (AusAID)</li> <li>- Department for International Development (DFID)</li> <li>- US Department of State / United States Agency for International Development (USAID)</li> </ul>
Source: Author	

Results measurement schemes differ in terms of:

- the levels of the results chain at which standard indicators are defined;
- how the results of interventions are reported that focus on qualitative changes;
- how agencies deal with the issue of attribution vs. contribution;
- how performance is assessed (i.e. baseline values and target values).

12 EuropeAID does not currently use standard indicators to report on its aggregate contributions to overarching development goals. However, it is in the process of developing a results framework for measuring and communicating development results delivered by EU-funded projects and programmes managed by EuropeAid. This is known as the “EU development and cooperation results framework”. The process is almost complete and several documents have been published describing the envisaged framework in detail (EC 2011, 11; EuropeAID 2013b; EC 2013).

#### 4.1 Summary of results measurement systems and use of results information

The majority of development agencies reviewed measure results at two levels within agency results frameworks<sup>13</sup> (see Figure 4). At level 1, internationally agreed standard indicators such as the MDG indicators are used to measure development outcomes and impacts in partner countries. Indicators at this level measure medium- to long-term development changes that are products of the joint efforts of partner countries, donors and other influencing factors and cannot be attributed to support provided by individual agencies. For this reason, they are not used to assess performance, but rather to provide information on the overall development context in partner countries. The information is useful as a strategic orientation (i.e. what are the results to which we wish to contribute?) and for monitoring the relevance of development strategies over time (i.e. what progress has already been made and where are further efforts needed?). Monitoring country outcomes should also strengthen the focus on development outcomes and help to align short-term activities with long-term objectives.

At level 2, development agencies measure their key contributions (i.e. intervention activities, outputs and outcomes) to the outcomes and impacts measured at level 1. Standard indicators at level 2 are agency-specific and are commonly selected with the aid of the following criteria:

- (1) they can be used regularly in ongoing projects and programmes;
- (2) they are compatible with strategic priorities;
- (3) they are easy to measure and aggregate.

The main purpose of indicators used at level 2 is to provide a snapshot of aggregate results in order to comply with external reporting requirements. Reporting on results by means of standard indicators is also an effective public relations tool as it helps to underline the positive effects of development cooperation.<sup>14</sup>

The development agencies reviewed also seek to use the indicators for management and learning purposes. For example, DFID states in its results framework that, by “*measuring results we get a much better idea of what works and what does not, so we can refine our programmes accordingly. We are also able to manage our resources to deliver these results*” (DFID 2013a, 1). Similarly, the IFC aims to use the indicators “*to drive implementation of strategy and influence operational decision-making, alongside volume targets*” (IFC 2013a).

---


13 Most of the agencies reviewed measure results as part of a three- or four-level results framework that also includes indicators of operational and organisational effectiveness (i.e. levels 3 and 4). The US Department of State/USAID and the International Finance Corporation (IFC) are exceptions to this. The US Department of State and USAID jointly monitor their contributions as part of a two-level strategic plan setting out five strategic goals and several strategic objectives (US Department of State / USAID 2014b). The IFC has developed the IFC Development Goals (IDGs) for reporting on its aggregate contributions. The IDGs are high-level targets for incremental reach. Progress towards each goal is monitored with the aid of just one standard indicator (IFC 2013b).

14 For example, the International Development Association (IDA) published a video entitled “Jobs, Clean Water and Education: The World Bank's Fund for the Poorest” in 2013 to advertise its work. The video uses information on aggregate results (e.g. “we immunised 500 million children”) to illustrate IDA’s aggregate contributions to poverty alleviation (see <http://www.youtube.com/watch?v=02NPzbnz7DQ>).

The formulation of agency-wide targets for results raises the value of standard indicators for accountability and management purposes. If targets are set, indicators can also be used to report on the effectiveness of the projects and programmes supported, i.e. the extent to which the desired results were actually achieved. In addition, by regularly collecting information on the progress made towards certain targets, development agencies can take corrective action in good time. For example, if a target is found to be off-track, management could analyse the reasons for underachievement. The findings could be used to take corrective action at project or programme level and additional funds could potentially be allocated to ensure that the target is met.

The following sections focus on the indicators used at level 2 of the results frameworks because these are the main tools for meeting accountability demands. Level 1 indicators measure the results of joint efforts and cannot be used to assess an individual agency's performance.

**Figure 4: Reporting on results in agency results frameworks**

<b>Level 1: Development progress in partner countries</b>	
Development outcomes and impacts	<u>Sample indicators:</u> <ul style="list-style-type: none"> <li>• Population living on less than \$1.25 (PPP) per day (%)</li> <li>• Under-5 mortality rate (per 1,000 live births)</li> <li>• Effective and accountable government (scale of 0-7)</li> </ul>
	
<b>Level 2: Contributions to development results</b>	
Intervention activities, outputs and outcomes	<u>Sample indicators:</u> <ul style="list-style-type: none"> <li>• Number of people with access to new or improved health services</li> <li>• Number of teachers trained</li> <li>• Number of cross-border and transnational projects supported</li> </ul>
Source: Author's research based on ADB (2013b); AusAID (2012, 20); DFID (2013a, 3); World Bank (2013, 3); and AfDB Bank Group (2013b, 3)	

## 4.2 Levels of the results chain and the inclusion of qualitative results

Standard indicators measuring contributions to development progress in partner countries (i.e. level 2 indicators) can be formulated at different levels of the results chain (see Figure 1 and Table 2). The level at which standard indicators should be defined depends on different – and in some cases conflicting – factors:

- (1) Standard indicators should be related as closely as possible to an agency's overall objective, e.g. poverty reduction and sustainable development. The latter is usually defined at impact level. Indicators that measure outputs and outcomes are a better proxy for the achievement of an agency's overall objective than indicators that measure inputs or activities.
- (2) If used for accountability purposes, standard indicators should ideally be directly attributable to the efforts of an individual agency. It should at least be possible to demonstrate that the efforts of the reporting agency were one of the main causes of observed change (or contribution<sup>15</sup>). An agency cannot be held responsible for the achievement of results beyond its control. Up to the level of outputs, it is fairly easy to attribute changes to specific interventions or development actors because inputs, activities and outputs are controlled by the stakeholders performing the intervention. Similarly, short-term outcomes – which are the changes arising directly from an intervention's outputs (such as improved access to clean water) – are fairly easy to attribute. By contrast, medium-term outcomes and long-term development impacts are influenced by many external factors, such as the economic and political environment and the activities of other development agencies. These make attribution difficult (Prennushi / Rubio / Subbarao 2002, 108–110).
- (3) The level of the results chain at which indicators are used also depends on the costs of data collection and the availability of data. Data on activities and outputs is usually available directly from project or programme records or in the form of administrative data from national or local government institutions (Prennushi / Rubio / Subbarao 2002, 113–115; Binnendijk 2000, 37). Short-term outcomes can be assessed by relatively low-cost surveys such as quick monitoring surveys, rapid appraisals or participatory methods (Binnendijk 2000, 39; Prennushi / Rubio / Subbarao 2002, 115).<sup>16</sup> Data collection and monitoring costs most if indicators are used that measure medium-term outcomes or impacts. An attribution or contribution analysis at these levels requires costly and methodologically sound evaluations.

The majority of indicators used by the development agencies reviewed in their agency-level results frameworks measure outputs and short-term outcomes. For three reasons, these levels are best suited for reporting on contributions to development in partner countries:

- (1) There is a strong and direct link with inputs supplied by the reporting agency, which allows the attribution of results or at least a contribution claim to be made.

---

15 “Contribution analysis aims to demonstrate whether or not the evaluated intervention is one of the causes of observed change”(EuropeAID 2006).

16 In practice, however, data is often provided as estimates by project or programme staff (using available data). For example, an indicator frequently used by donor agencies is the number of households or people with access to improved water sources. Data on the number of water points built or rehabilitated is taken from project or administrative records. This number is then multiplied by the average number of people or households benefiting per water source (see DFID 2013b, for example). The number may be based on estimates by project staff or available survey estimates.

- (2) Although indicators do not reflect changes at the overall objective level (i.e. impact level), results<sup>17</sup> are measured and more information is provided than by reporting on inputs or activities alone.
- (3) The cost of data collection is relatively low.

However, interventions that aim to support policy reform or build institutions and capacity in partner countries do not usually produce tangible, quantifiable outputs or short-term outcomes. To capture the contributions of these types of interventions, development agencies resort to activity-level indicators or alternatively to indicators that measure medium-term to long-term changes in partner countries (i.e. medium-term outcome or impact indicators).

Standard indicators that measure activities usually count the number of people, workshops, projects or countries supported in a specific intervention area or sector (e.g. “the number of countries where adaptation programmes have helped to reduce their vulnerability to climate change”). As with outputs, the cost of data collection is low (data is usually readily available from project or administrative records) and activities may be attributed to the efforts of a single agency. Activity-level indicators are less suited, however, for reporting on results because they do not provide information on what was achieved with the assistance provided.

Indicators that measure medium-term outcomes and impacts best reflect the overall objectives of development agencies. However, the problem of attribution means that they are not very useful for reporting on the results contributed by an individual agency (Flint 2003, 43–44). The indicators at medium-term outcome and impact level that are currently used by the donor agencies reviewed for reporting on results measure country-wide changes produced by the joint efforts of partner countries, development actors and other factors. No evidence is provided from evaluations that may support an agency’s claim that its efforts contributed to the changes measured. For example, the AfDB reports on the ‘share of countries with improved transparency, accountability and corruption mitigation in the public sector’ as a measure of whether its interventions have helped countries to improve the public sector in these areas. Calculations are based on the supported countries’ Country Policy and Institutional Assessment (CPIA) scores before and after an AfDB intervention (AfDB Bank Group 2013b, 17). However, this is not sufficient to identify the AfDB’s contribution to the changes measured. Rigorous impact evaluations, which address the attribution problem by establishing the counterfactual, i.e. “*what would have occurred in the absence of an intervention*” (Leeuw / Vaessen 2009, ix) and comparing it with what happened as the result of the intervention (Leeuw / Vaessen 2009, 23), could shed more light on the extent of contributions. However, it is challenging, costly and time-consuming to provide such evidence at an aggregate level. Ideally, evaluations would have to be performed in all partner countries, taking account of all projects that may have contributed to changes in a specific standard indicator.

---

17 The OECD/DAC (2009, 42) defines a result as “the output, outcome or *impact (intended or unintended, positive and/or negative)* of a development intervention”.

<b>Table 2: Levels of the results chain (reporting on contributions to long-term development goals)</b>			
<b>Level of the results chain</b>	<b>Example indicators</b>	<b>Donor agencies that use standard indicators at this level</b>	
		<b>Total number of agencies</b>	<b>Names of agencies</b>
<b>Activities</b>	<ul style="list-style-type: none"> <li>- Number of national frameworks for climate change mitigation supported</li> <li>- Number of children in lower secondary education supported</li> </ul>	7	AusAID, DFID, EuropeAID, IDB, UNDP, USAID, World Bank
<b>Outputs</b>	<ul style="list-style-type: none"> <li>- Number of trained specialists in the healthcare sector</li> <li>- Land administration offices established or upgraded</li> </ul>	9	ADB, AfDB, AusAID, DFID, EuropeAID, IDB, UNDP, USAID, World Bank
<b>Short-term outcomes</b>	<ul style="list-style-type: none"> <li>- Number of individuals with access to improved healthcare services</li> <li>- Number of hectares under improved practices as a result of training</li> </ul>	10	ADB, AfDB, AusAID, DFID, EuropeAID, IDB, IFC, UNDP, USAID, World Bank
<b>Medium-term outcomes and impacts</b>	<ul style="list-style-type: none"> <li>- Share of countries with improved quality of public administration</li> <li>- Number of malaria-specific deaths per 1,000 persons per year</li> </ul>	8	AfDB, AusAID, DFID, EuropeAID, IDB, UNDP, USAID, World Bank
Source: Author's illustration.			

As Table 2 shows, of the donor agencies reviewed, only the ADB and the IFC do not use indicators either at activity level or at the level of medium-term outcomes or impacts. This also implies that both agencies report on interventions that focus on qualitative changes only to a very limited extent.

### 4.3 Attribution vs. contribution

One of the main questions when using standard indicators to aggregate intervention results at agency level is whether the results may be attributed directly to an individual agency's funding (e.g. "We have trained one million teachers.") or whether they are a product of a number of factors, i.e. an agency is only one of several actors contributing to the results (e.g. "Together with other development partners, we helped to provide 40 million people with access to improved water sources.") (da Costa 2009, 10–13; EuropeAID 2013a, 5–6).

I found that the majority of donor agencies reviewed,<sup>18</sup> i.e. AfDB, DFID, IDB, IFC and US Department of State/USAID, follow the attribution approach. The use of the attribution approach is a response to greater pressure by donor governments, parliaments and citizens on development agencies to account for results. For example, public pressure

18 Unfortunately, the document review did not make clear what approaches are used by AusAID and UNDP in reporting on aggregate results. UNDP states that output indicators in the strategic plan measure "only those results from schemes, services, plans, actions, etc. which are specifically supported by UNDP" (UN 2013b). However, there is no information on whether UNDP attributes these results to its support and whether full or proportional results are counted. Similarly, AusAID states that indicators "measure the contribution of Australian aid to development outcomes in [...] partner countries" (AusAID 2012, 20), without providing any further information.

in the UK has led DFID to emphasise value for money and to present aggregate results under a banner headed “*What results has DFID financed?*” (DFID 2013a, 3).

Only two of the donor agencies reviewed, i.e. the ADB and the World Bank, exclusively follow the contribution approach in reporting on aggregate results. This gives greater recognition to contributions from partner countries and other development partners. The World Bank, for instance, describes the results measured at level 2 of its corporate scorecard as “*Country Results Supported by the World Bank*” (World Bank 2013, 9). EuropeAID is also planning to follow a contribution approach for its results framework. However, it will also consider the attribution approach if results can be identified that can be linked directly to EU support (EC 2013, 10).

The practice of attribution should be viewed critically. Attribution is methodologically challenging and often disregards the contributions of development partners. As explained in section 0, attribution is still manageable at output level and becomes harder as the level of the results rises, i.e. as more factors influence results. With the exception of DFID, however, all of the agencies reviewed that follow the attribution approach only attribute results up to the level of short-term outcomes. If indicators at medium-term outcome or impact level are used, the contribution approach is followed or no claims are made about attribution or contribution at all. For example, the US Department of State and USAID explain that data on medium-term outcomes across partner countries, such as “Prevalence of Stunted Children under Five Years of Age”, serve as “*additional evidence measuring achievement of the objective*”(U.S. Department of State / USAID 2014a, 58).

It is also difficult to attribute results in projects and programmes that are supported and operated jointly by a variety of development actors. Most of the agencies reviewed use contribution factors to establish attribution if results are co-financed by a number of development actors. These are often, but not always, equal to the share of the total programme costs paid by the reporting agency. For example, the contribution factors used by the IFC to attribute results are a function of the type of intervention (i.e. advisory services or equity and loan investments) and IFC’s stake in a project. In the case of equity and loan investment, the IFC counts *pro rata* incremental reach as long as the equity provided by IFC is below 10% of total equity or if IFC loans are worth less than 20% of the total project cost. If the equity or loans exceed these percentages, however, 100% of incremental reach is measured (IFC 2013c; IFC 2013d).<sup>19</sup> For advisory services, the IFC claims 100% of the results of projects, based on an assumption that the transaction in question would not have been completed without its advisory services (IFC 2013c).

DFID also uses contribution factors to establish attribution at the medium-term outcome or impact level. For example, the “number of people achieving food security through DFID support” is calculated as follows. First, country offices use the Integrated Food Security Phase Classification (IPC)<sup>20</sup> to calculate the annual number of people becoming

---

19 The contribution rules used by the IFC for the IDGs are presented here in simplified form. See IFC (2013c) and IFC (2013d) for more detailed explanations of the contribution rules for IDG 2: Health and Education and IDG 3: Financial Services.

20 “*The IPC is an innovative tool for improving food security analysis, international status comparison, and decision-making. It is a standardised scale that integrates food security, nutrition and livelihood information into a clear statement about the food security status as well as the nature and severity of a crisis and implications for strategic response.*” (DFID 2013c)

food-secure. The number of people who have achieved food security thanks to DFID support is then calculated, based on the percentage share of DFID funding in overall funding of food security (at national or sub-national level, DFID 2013c). At a country or sub-national level, however, it is very difficult to assess DFID's share of overall funding for food security. Moreover, whether or not people move into food security is influenced by a variety of factors that are not related to the amount of funding available. These include domestic food prices, natural disasters and the economic situation in the country in question.

The contribution approach is generally perceived as being less problematic because it does not make any claims about the precise quantity of results that may be attributed to an individual agency. At the same time, however, it is also regularly criticised for producing double counts. Donor agencies following the contribution approach usually report on the full results of projects and programmes, even if they are co-financed by several actors. This means that the same results are measured and counted by many different organisations. Any attempt to aggregate outputs and outcomes across the results frameworks of different agencies would therefore lead to greatly inflated figures. To give an indication of the extent of their contributions to results, some donor agencies calculate the average share of funding they contributed to interventions that delivered the reported results. For instance, the ADB states in its 2012 Development Effectiveness Review that it paid 35% of the total cost of operations that programmed the results presented (ADB 2013a, 14).

#### 4.4 Performance assessment

Results statements based on standard indicators alone, such as “we contributed to the building or upgrading of 500,000 km of roads in 2012,” have limited informational value for accountability and management purposes. Benchmark values (i.e. baseline and/or target values) are needed in order to assess a development agency's performance in terms of its contribution to development progress in partner countries.

Where baseline values are used, performance is assessed by comparing the results achieved over two different periods of time. However, this may lead to the wrong conclusions. For example, if an agency trains two million teachers in year 1 but only one million in year 2, this is not necessarily a sign of bad performance. It could just as easily point to a shift in strategic priorities away from education. It could also be a sign of diminishing returns if, for instance, the cost of providing households with new electricity connections rises with each additional household reached. Moreover, as long as there is no information about the cost of results, as is currently the case with results reporting by development agencies, it is difficult to assess whether performance has changed from one year to the next.

Targets, which are specific indicator values that need to be attained within a specific time-frame (Binnendijk 2000, 32), are better suited for assessing performance. Agency-wide targets for results show what an agency aims to achieve. They can be used to assess the performance of the organisation as a whole by comparing the actual results achieved with the targets. Targets also increase the usefulness of standard indicators for management purposes by serving as guiding posts for gauging whether implementation is proceeding as



planned. They are also an orientation and motivation for those responsible for implementing interventions or strategies (Binnendijk 2000, 33).

Targets for standard indicators that measure aggregate results can be set using either a bottom-up or a top-down approach. Bottom-up targets are initially set for individual interventions. Agency-wide targets are then estimated or calculated as the cumulative results that all interventions are planned to deliver within a given period and which have ideally been decided jointly with development partners. Top-down targets are a tool for assisting with the implementation of strategies. They are usually chosen to reflect strategic priorities and are set at organisational level for the results to be produced by all interventions as a whole. For example, an agency seeking to increase its activities in the energy sector might set itself the following target: “We will install 20,000 km of transmission lines during the period up to 2020. This is an increase of 30% compared with the baseline period”.

Setting targets from the top down is a challenging exercise as it is difficult to estimate what results can be realistically expected across countries and interventions. There is usually a long time-lag between the time when strategic objectives are set and the point at which the first results of the strategy are observed.<sup>21</sup> Moreover, whether targets are met depends not only on the efforts of the development agency, but also on whether there is sufficient demand by partner countries for the prioritised results.

In addition to setting targets in terms of the cumulative results to be achieved within a given period, donor agencies can also set targets for achievement rates in percentage terms (i.e. results achieved / planned results). A precondition for the calculation of achievement rates is that cumulative results targets are set at intervention level and/or at agency level. Achievement rates can be calculated either initially at the level of individual interventions and then averaged at organisational level, or directly at organisational level based on the agency-wide results target.

Some agencies also use a ‘traffic-light system’ based on achievement rates, for measuring performance against targets. For instance, the AfDB rates operations that achieve 95% or more of their targets as green, operations that achieve between 60% and 94% of their targets as yellow, and operations that achieve less than 60% of their targets as red (AfDB Bank Group 2013a, 26). The purpose of a traffic-light system is to highlight areas in which performance is off-track so that timely corrective action can be taken (AfDB Bank Group 2010, 20; ADB 2012, 34).

The comparison of donor agencies (see Table 3) shows that the majority of those reviewed, i.e. AfDB, AusAID, DFID, IDB, IFC, UNDP and the US Department of State/USAID, have set cumulative results targets for standard indicators. Most of them also report on the number of results achieved in previous years (baselines). Only the ADB and the World Bank have decided against the use of cumulative results targets. The ADB instead sets itself a target expressed as a percentage of the planned outputs and outcomes it aims to achieve. The achievement rate only takes account of projects completed in the reporting period and is calculated from the bottom up, i.e. it is first calculated for each

---

21 For instance, the ADB measures the results of closed interventions at level 2 of its results framework. It states that operations assessed at this level were typically planned between six to ten years prior to the assessment year (ADB 2012, 5).

intervention and then averaged across interventions. The World Bank does not use any targets at all at level 2 of its results frameworks. The reason given by the World Bank for this is that standard indicators measure country results achieved with Bank support that are demand-driven (World Bank 2013, 9). Setting targets for standard indicators would conflict with the alignment of support with partner countries' priorities.

**Table 3: Use of baselines and targets: comparison of development agencies**

	Baseline		Target	
	Cumulative results	Achievement rate (%)	Cumulative results	Achievement rate (%)
<b>AusAID</b>	–	–	✓	–
<b>ADB</b>	–	✓	–	✓
<b>AfDB</b>	✓	✓	✓	✓
<b>DFID</b>	–	–	✓	–
<b>EuropeAID</b>	?	?	?	?
<b>IDB</b>	✓	–	✓	✓
<b>IFC</b>	–	–	✓	–
<b>WB</b>	✓	–	–	–
<b>UNDP</b>	✓	–	✓	–
<b>USAID</b>	✓	–	✓	–

Source: Author's data

## 5 Limitations and risks of using standard indicators for results reporting

By using standard indicators to report on aggregate development results, donor agencies aim both to meet the reporting requirements of external stakeholders and to generate information that is useful for management and decision-making. This chapter assesses whether the results information provided by development agencies in results frameworks forms a sufficient basis for meeting the two objectives. It revisits 'synecdoche' and the assumption of a low 'gaming' risk, both of which are preconditions for functioning performance measurement schemes, and explores the limitations, risks and adverse effects of using standard indicators. Differences in the design of indicator systems identified in chapter 4 are taken into account.

### 5.1 Limitations

Standard indicators are often criticised for providing only limited information, thus impairing their overall usability for accountability and management purposes. The assumption of synecdoche, i.e. that selected key indicators can be relied on for performance measurement and that the omission of other performance aspects does not produce adverse effects, is unlikely to be valid. There are a number of reasons for this. Many of them relate to the challenges of performance measurement in development cooperation highlighted in section 0.

First, not every result can be aggregated. Hence, measured results present an incomplete picture of an agency's total contributions to higher level development goals. In particular,

many results of development interventions are qualitative and cannot easily be standardised and captured by quantitative standard indicators at output level or at the level of short-term outcomes. For example, the indicators included at level 2 of the World Bank's Corporate Scorecard do not capture the Bank's knowledge work activities (e.g. analysis, policy advice and knowledge-sharing), even though knowledge work is one of the Bank's core activities and accounted for 31% of the Bank's budget in 2010 (IEG 2012, 65–66). Reporting on a selected set of development outputs and outcomes may be misleading, since only limited information is provided and not all aspects of development effectiveness are captured (IEG 2011, 113).

Second, the variety of objectives, approaches and circumstances inherent to development interventions is too great to be adequately reflected by standard indicators alone (Cronin / Regan 2002, 76; Vielajus et al. 2009, 61–62). In order to measure the results of a diverse set of interventions, very broadly defined key indicators are used. However, these provide only a very rough and simplistic approximation of activities, outputs and outcomes. For example, the standard indicator defined as “number of farmers trained” does not take account of the qualitative aspects of interventions and does not capture the diversity of interventions aimed at increasing farmers' capacities.<sup>22</sup>

Third, development agencies can only select a limited set of standard indicators for inclusion in their results framework if they are not to overwhelm the target audience and preserve a snapshot character. The majority of development agencies reviewed use between 20 and 40 standard indicators to report on their contributions in results frameworks.<sup>23</sup> This is a very small number considering the diversity of most development agencies' project portfolios. Weiers (2012) points out that hundreds of different standard indicators would be needed in order to report on all results and capture their specificity. The ADB (2012, vi), for instance, states that 70% of projects performed in 2011 with completion reports or extended annual reviews reported on at least one of the standard indicators in its agency results framework. This means that the results of 30% of projects were not reported on at agency level. One can also assume that, even where projects use standard indicators, only a very small share of total results are captured by the standard indicators.

Fourth, as presently formulated, the standard indicators used by the donor agencies reviewed do not provide any information on the costs or adverse effects associated with the delivery of outputs and outcomes. Building roads, for instance, is often associated with external costs such as deforestation or loss of biodiversity (IEG 2011, xxvi). Standard indicators also do not provide an indication of whether the right people benefited from the right services (IEG 2011, 112–113).

Fifth, reporting on quantitative achievements may create an impression that more results are always better. This is not always the case, however. For instance, more road-building does not necessarily lead to lower transport costs (IEG 2011, xxvi).

---

22 For example, it may make a big difference whether a farmer attended a week-long course or a two-hour training session.

23 The IFC has defined only six indicators for measuring progress towards the IFC development goals (IFC 2013b), while UNDP uses the highest number of standard indicators in its results frameworks, i.e. 92 (UN 2013a).

Sixth, the indicators used do not provide information on the long-term effectiveness of the support provided by a development agency. This problem is also recognised by the agencies reviewed for this paper. The IDB (2013, 30), for instance, notes that the difficulty of linking outputs measured at level 2 of its corporate results framework with outcomes and impacts measured at level 1 reduces the overall usefulness of the framework.

## 5.2 Risks and adverse effects

The analysis of the limitations of standard indicators indicates that standard indicators cannot be relied upon to assess the performance of a development agency because the assumption of synecdoche is not satisfied. A high proportion of results are not measured and reported on by standard indicators. Similarly, the share of results that can only be captured by inadequate performance measures is high. Using standard indicators for accountability and management purposes despite their limitations leads to a high risk of gaming and to various unintended effects. The risk of gaming is especially high if targets are set for standard indicators as this increases pressure on development agencies to deliver results. The risk is relatively low in comparison if only baseline values are used as benchmarks in assessing performance.

Distortions may arise because the majority of standard indicators used by the development agencies reviewed focus on quantitative aspects (e.g. the number of beneficiaries reached or outputs produced) rather than on the quality of the goods and services delivered. The focus on quantities may produce an emphasis on projects that operate in more favourable environments (IFC 2013a). It may also lead to a prioritisation of activities that deliver a particularly high volume of results with the fewest resources or that reach an especially high number of beneficiaries, i.e. the low-hanging fruit. At the same time, it may sharpen the focus on larger projects that have a greater reach at the expense of smaller projects, because standard indicators that count the quantity of activities or results are influenced more by larger countries and projects (IEG 2011, 112). The Independent Evaluation Group (IEG) states, for instance, that the World Bank and the IFC could restrict their efforts to India and China (instead of working with all the member countries) if they wished to influence the results measured by standard indicators (IEG 2011, 112).<sup>24</sup> This also means that just one successful large project would be capable of concealing shortcomings in several small projects. In the same manner, one big project that failed could easily lead to the non-achievement of a target, even though the majority of projects performed well. The IFC is currently trying to solve these problems by experimenting with a system of assigning weightings to smaller projects in poorer countries (IFC 2013a). Another potential solution is to provide the public with better information about who benefits from interventions. This can be done by presenting disaggregated data for groups of beneficiaries (e.g. disaggregated by sex, location or age). The disaggregation of results

---

24 To mitigate this problem, the IEG proposes not just using reach indicators, but also reporting on the number of countries that have achieved a certain degree of progress (e.g. the number of countries in which x percent of the population lives on less than USD 2 a day or the number of countries with household electrification rates above x percent, IEG 2011, 112). However, such indicators measure country-wide, longer-term changes. It is very difficult and costly to assess whether individual actors influence changes at this level.

by countries and interventions may also help to paint a more balanced picture of who benefits from what type of intervention. The collection of disaggregated data, may, however, increase the cost of data collection and monitoring of standard indicators.

Results reporting by means of standard indicators may also conflict with the accountability demands set by aid recipients and the commitments made as part of the aid effectiveness agenda.

Indicators used by the development agencies reviewed for reporting on performance measure activities, outputs and short-term outcomes and do not sufficiently reflect the underlying longer-term objectives of development cooperation. Strong public pressure to deliver the results measured by standard indicators may thus lead to more resources being allocated to the delivery of short-term results, to the detriment of interventions focusing on long-term changes through support for policy reform, institution-building and capacity-building. At the same time, public pressure may also play a role in reducing the use of partner-country systems if short-term results can be achieved more efficiently by projects implemented by development agencies alone.

There is also concern that cumulative results targets for standard indicators set from the top down at the organisational level may conflict with efforts to align support with partner countries' priorities. If demand for support in one of a development agency's priority areas is lower than expected, either it will be impossible to achieve the target or more resources will need to be invested in order to do so. However, this may lead to an inefficient allocation of aid (or less value for money) if the additional money could have been used to greater effect elsewhere.

Moreover, the practice of some donor agencies of attributing results directly to their own funding ("we built schools") instead of measuring partner countries' results to which they have contributed ("we helped partner countries to build schools") may be regarded as being contrary to the principle of country ownership<sup>25</sup> (da Costa 2009, 10-12). The attribution of results does not acknowledge the leadership role of partner countries and their contributions to measured results. It may also conflict with commitments on harmonisation and the use of country systems, as results are harder to attribute if interventions are co-financed by several development actors and led by partner countries.

## 6 Conclusion and policy recommendations

The use of standard indicators to report on agency-wide results in results frameworks represents an attempt by development agencies to account for their results to external stakeholders and to become more results-oriented. However, the finding in this paper is that the data on aggregate results provided by the donor agencies reviewed is only of limited informational value and may not be a sufficient basis on which to hold donor agencies to account. Only a limited snapshot of activities, outputs and short-term outcomes is captured in results frameworks and there is no evidence of how much these contribute to the achievement of long-term development goals. In addition, the emphasis

---

25 Country ownership means that "partner countries exercise effective leadership over their development policies, and strategies and co-ordinate development actions" (OECD/DAC 2005/2008, 3).

on activities and short-term results is not in line with the principles of RBM, which suggests focusing on outcomes and impacts.

The analysis also raises the question of whether development agencies choose to report on short-term results so that they cannot be held truly accountable for results. Development agencies that are concerned about demonstrating results to convince an increasingly sceptical public about the benefits of aid, do not have an interest in reporting honestly on failures or the adverse effects of development interventions. Results reporting with standard indicators highlights the positive contributions of aid without providing information on the cost or adverse effects of the results. Acknowledging and learning from failures is, however, important in order to improve aid.

The paper also argues that, rather than improving aid effectiveness, the use of standard indicators to report on results may produce certain adverse effects. For example, as most standard indicators measure the quantity of results and do not account for qualitative differences, their use may lead to a focus on quantity instead of quality and to the pursuit of low-hanging fruit. In addition, the current results measurement and reporting practices of the donor agencies reviewed aggravate rather than balance the competing accountability demands of aid funders and recipients. The emphasis on accountability to funders is especially evident in the focus on short-term results (i.e. quick wins) as opposed to long-term institution-building. It is also reflected by the fact that the majority of donor agencies reviewed use the attribution approach rather than the contribution approach to show what their money has delivered.

Given the limitations and adverse effects of the use of standard indicators for reporting purposes, one of the main challenges is to ensure that accountability mechanisms drive rather than hinder aid effectiveness. Donor agencies could better balance and align the competing accountability demands and avoid some of the unintended effects of aggregate results reporting by following the recommendations set out below.

- (1) First, when reporting on agency-wide contributions, donor agencies should also report on interventions that aim to produce qualitative and long-term changes. This can help to prevent too strong a focus being placed on quantitative and short-term results. It also takes account of the realities faced by development cooperation actors, who are seeking to establish lasting institutions and capacities in partner countries. Ideally, development agencies should use indicators that reflect changes at medium-term outcome or impact level for this purpose. On the other hand, evaluations for assessing the extent of aggregate contributions to results measured at medium-term outcome and impact level are likely to prove prohibitively expensive. Evaluations need to be performed across the entire portfolio, which would direct a significant portion of resources away from programme implementation, especially if reporting is annual. Activity-level indicators should therefore be used to report on efforts that may contribute to longer-term changes at an aggregate level. This type of information could be complemented by qualitative analyses at country level.
- (2) Donor agencies are advised to report on the results they have supported (i.e. based on either a full or a proportional contribution) and not to attribute results to their own engagement. The aid effectiveness agenda calls for a general shift in results reporting from the attribution of results to individual donor agencies (“What has our aid delivered?”) towards the country level as the primary unit of analysis (i.e. “What

have countries achieved with the support of development agencies?"). The attribution of results hinders donor harmonisation and may be seen as being contrary to the principle of country ownership. Moreover, the attribution of results is methodologically challenging and is often not possible when outcome-level or impact-level results are measured that are influenced by various external factors, and when interventions are co-financed by several development partners.

- (3) Donor agencies should refrain from setting cumulative results targets for standard indicators at the organisational level. Such targets may conflict with efforts to align support with partner countries' development priorities and strategies, especially if they are set from the top down. Lessons learned from complexity theory also make a strong case against setting agency-wide targets from the top down. When problems are complex, as is often the case in development cooperation, decisions should ideally be taken at the lowest possible organisational level. This is because the characteristics of a problem become more apparent the closer one gets to the reality of lived experience (Eyben 2008, 24). As an alternative to cumulative results targets, I suggest using achievement rates (i.e. results achieved / planned results). These should first be calculated for each individual intervention and then averaged across interventions to avoid the unintended consequences of agency-wide targets. For example, if averaged across interventions, achievement rates do not have an adverse effect on alignment and prevent a situation from arising in which the achievement of targets depends on just a few large projects or programmes.
- (4) Donor agencies should invest in harmonising definitions of standard indicators and reporting methodologies and make these definitions publicly available. This would decrease the cost of coordination among development partners when implementing joint projects and programmes. It would also reduce the overall cost of data collection and monitoring and lower the burden placed on partner countries' statistical systems. In addition, it would facilitate the comparison of results information presented by different donor agencies.
- (5) Fifth, given that the information on aggregate results reported in results frameworks is only of limited value for accountability purposes, I suggest exploring complementary or alternative options for meeting accountability demands. Reporting on aggregate results should always be complemented by greater transparency on individual aid interventions, rigorous impact evaluations and in-depth qualitative analyses, and by the systematic incorporation of beneficiary feedback into monitoring and evaluation efforts.

Greater aid transparency at intervention level, and reporting on the results of individual interventions (for example, in an online database or as part of the International Aid Transparency Initiative Standard) would allow the public in donor and recipient countries to hold development actors accountable for achieving project or programme objectives. Although development actors already undertook to do so at the 4th High Level Forum on Aid Effectiveness in Busan in 2011, barely any progress has been made in implementing this commitment to date.

Conducting rigorous impact evaluations and qualitative analyses for selected interventions and country programmes can help actors to better understand the complex processes leading to development impacts. Moreover, rigorous impact evaluations allow longer-term impacts to be attributed to an intervention, facilitating

a model presentation of the contribution of one or more donor organisations to longer-term development results.

Beneficiary feedback mechanisms can help to strengthen accountability to aid recipients. They provide additional information on whether projects have delivered the intended results, on the quality of the outputs and services delivered, and on whether funds have reached the intended beneficiaries (Natsios 2010, 40; Darendeliler / Gurkan 2012, 3). Pressure to openly account for the findings of beneficiary feedback may make donor agencies more responsive to recipients' needs, and incentivise donor agencies to provide better and more targeted services (Darendeliler / Gurkan 2012, 3). In order to provide a snapshot of agency-wide performance to external stakeholders, it is important to find ways of aggregating the findings of evaluations and beneficiary feedback at agency level.



## Bibliography

- ADB (*Asian Development Bank*) (2012): Review of the ADB results framework; online: [www.adb.org/sites/default/files/review-of-adb-results-framework-2012.pdf](http://www.adb.org/sites/default/files/review-of-adb-results-framework-2012.pdf) (accessed 25 Nov. 2013)
- (2013a): Development effectiveness review 2012, Manila
- (2013b): Results framework 2013-2016; online: <http://www.adb.org/sites/default/files/results-framework.pdf> (accessed 7 Oct. 2013)
- AfDB Bank Group (*African Development Bank Group*) (2010): Bank group results measurement framework; online: <http://www.afdb.org/fileadmin/uploads/afdb/Documents/Publications/Bank%20Group%20Results%20Measurement%20Framework%20-%20REV%20%5B1%5D.pdf> (accessed 7 Oct. 2013)
- (2013a): Annual development effectiveness review 2013 : towards sustainable growth for Africa : annual review, Tunisia
- (2013b): The one bank results measurement framework 2013-2016; online: [http://www.afdb.org/fileadmin/uploads/afdb/Documents/Project-and-Operations/The\\_One\\_Bank\\_Result\\_Measurement\\_Framework\\_2013-2016.pdf](http://www.afdb.org/fileadmin/uploads/afdb/Documents/Project-and-Operations/The_One_Bank_Result_Measurement_Framework_2013-2016.pdf) (accessed 16 July 2014)
- AusAID (*Australian Agency for International Development*) (2012): Helping the world's poor through effective aid : Australia's comprehensive aid policy framework to 2015–16, Canberra; online: <http://www.ausaid.gov.au/about/documents/capf.pdf> (accessed 7 Oct. 2013)
- Bach, S. / L. Bordogna (2011): Varieties of new public management or alternative models? The reform of public service employment relations in industrialized democracies, in: *The International Journal of Human Resource Management* 22 (11), 2281-2294
- Barder, O. (2012): What are the results agenda?; online: <http://www.owen.org/blog/5228> (accessed 28 April)
- Bester, A. (2012): Results-based management in the United Nations development system : progress and challenges : a report prepared for the United Nations Department of Economic and Social Affairs, for the quadrennial comprehensive policy review; online: [http://www.un.org/esa/coordination/pdf/rbm\\_report\\_10\\_july.pdf](http://www.un.org/esa/coordination/pdf/rbm_report_10_july.pdf) (accessed 14 July 2014)
- Bevan, G. / C. Hood (2006): What's measured is what matters : targets and gaming in the English public health care system, in: *Public Administration* 84 (3), 517-538; online: <http://dx.doi.org/10.1111/j.1467-9299.2006.00600.x>
- Binnendijk, A. (2000): Results based management in the development co-operation agencies : a review of experiences: DAC Working Party on Aid Evaluation; online: <http://www.oecd.org/development/evaluation/1886527.pdf> (accessed 7 Oct. 2013)
- Bogart, W. T. (1995): Accountability and nonprofit organizations: an economic perspective, in: *Nonprofit Management and Leadership* 6 (2), 157–170
- BP (2011): Busan Partnership for Effective Development Co-operation : Fourth High Level Forum on Aid Effectiveness, Busan, Republic of Korea; online: [http://effectivecooperation.org/files/OUTCOME\\_DOCUMENT\\_-\\_FINAL\\_EN2.pdf](http://effectivecooperation.org/files/OUTCOME_DOCUMENT_-_FINAL_EN2.pdf) (accessed 4 Aug. 2014)
- Chapman, J. (2002): System failure : why governments must learn to think differently, London: Henty Banks
- Cook, T. J. et al. (1995): Performance measurement : lessons learned for development management, in: *World Development* 23 (8), 1303–1315
- Cronin, D. / J. O. Regan (2002): Accountability in development aid : meeting responsibilities, measuring performance : a research report for Cómhlamh: Cómhlamh; online: [http://www.dochas.ie/Shared/Files/4/Comhlamh\\_Research.pdf](http://www.dochas.ie/Shared/Files/4/Comhlamh_Research.pdf) (accessed 7 Oct. 2013)
- da Costa, P. (2009): Study on communicating development results: OECD DAC Development Co-operation Directorate & DevCom Network; online: <http://www.oecd.org/dev/devcom/44836250.pdf> (accessed 7 Oct. 2013)

- Darendeliler, Z. / A. Gurkan* (2012): Monitoring through beneficiary feedback : a stocktaking of monitoring by non-state actors in bank-funded projects; online: <http://www.worldfishcenter.org/sites/default/files/Stocktaking-Monitoring-through-Beneficiary-Feedback.pdf> (accessed 20 May 2014)
- de Renzio, P. / S. Mulley* (2006): Promoting mutual accountability in aid relationships, Briefing Paper, London: Overseas Development Institute
- de Renzio, P.* (2014): The impossible geometries of foreign aid. mimeo
- Delorme, P. / O. Chatelain* (2011): Policy steering : the role and use of performance measurement indicators: Aid Delivery Methods Programme; online: [http://www.dochas.ie/Shared/Files/4/Guide\\_on\\_Performance\\_Measurement.pdf](http://www.dochas.ie/Shared/Files/4/Guide_on_Performance_Measurement.pdf) (accessed 14 Jan. 2013)
- DFID (Department for International Development)* (2011): UK aid : changing lives, delivering results; online: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/67584/BAR-MAR-summary-document-web.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/67584/BAR-MAR-summary-document-web.pdf) (accessed 10 Dec. 2013)
- (2013a): DFID's results framework, London; online: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/175715/DFID-external-results.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/175715/DFID-external-results.pdf) (accessed 19 August 2014)
  - (2013b): Indicator methodology note : number of people with sustainable access to clean drinking water sources through DFID support, in: DFID (ed.), DFID's results framework; online: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/332571/clean-drinking-water.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/332571/clean-drinking-water.pdf)
  - (2013c): Indicator methodology notes : number of people achieving food security through DFID support, in: DFID (ed.), DFID's results framework online: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/361381/food-security5.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/361381/food-security5.pdf) (accessed 21 Oct. 2014)
- Doucouliafos, H. / M. Paldam* (2009): The aid effectiveness literature : the sad results of 40 years of research, in: *Journal of Economic Surveys* 23 (3), 433–461
- Easterly, W.* (2006): Multilateral development banks : promoting effectiveness and fighting corruption : “Accountability for multilateral development banks”, Committee on Foreign Relations, United States Senate; online: [http://williamaeasterly.files.wordpress.com/2010/09/senateforeignrelationscommittee\\_oraltestimony\\_032806.pdf](http://williamaeasterly.files.wordpress.com/2010/09/senateforeignrelationscommittee_oraltestimony_032806.pdf) (accessed 30 June 2014)
- (2007): The white man's burden : why the West's efforts to aid the rest have done so much ill and so little good, New York: Oxford University Press
- Easterly, W. / T. Pfutze* (2008): Where does the money go? Best and worst practices in foreign aid, in: *Journal of Economic Perspectives* 22 (2), 29–52
- Ebrahim, A.* (2003): Accountability in practice : mechanisms for NGOs, in: *World Development* 31 (5), 813–829
- EC (European Commission)* (2011): Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions : increasing the impact of EU development policy : an agenda for change, Brussels
- (2013): Commission staff working document : paving the way for an EU development and cooperation results framework; online: <http://register.consilium.europa.eu/doc/srv?l=EN&t=PDF&gc=true&sc=false&f=ST%2017709%202013%20INIT&r=http%3A%2F%2Fregister.consilium.europa.eu%2Fpd%2Fen%2F13%2Fst17%2Fst17709.en13.pdf> (accessed 16 May 2014)
- Edwards, M. / D. Hulme* (1996): Too close for comfort? : The impact of official aid on nongovernmental organizations, in: *World Development* 24 (6), 961–973
- EuropeAID* (2013a): Results study; online: [http://capacity4dev.ec.europa.eu/sites/default/files/file/12/12/2013\\_-\\_1742/results\\_study\\_volume\\_i\\_11\\_12\\_13.pdf](http://capacity4dev.ec.europa.eu/sites/default/files/file/12/12/2013_-_1742/results_study_volume_i_11_12_13.pdf) (accessed 16 May 2014)
- (2013b): Sector indicator guidance for programming, s. 1
- Eyben, R.* (2005): Donors' learning difficulties : results, relationships and responsibilities, in: *IDS Bulletin* 36 (3), 98–107

- (2008): Power, mutual accountability and responsibility in the practice of international aid : a relational approach; online: <https://www.ids.ac.uk/files/Wp305.pdf> (accessed 19 Aug. 2014)
- (2013): Uncovering the politics of 'evidence' and 'results'. A framing paper for development practitioners; online: <http://bigpushforward.net/wp-content/uploads/2011/01/The-politics-of-evidence-11-April-20133.pdf> (accessed 27 March 2014)
- Flint, M.* (2003): Easier said than done : a review of results-based management in multilateral development institutions; online: [http://gametlibrary.worldbank.org/FILES/920\\_Results%20Based%20Management%20in%20Dev%20Partners.pdf](http://gametlibrary.worldbank.org/FILES/920_Results%20Based%20Management%20in%20Dev%20Partners.pdf) (accessed 7 Oct. 2013)
- Fry, R. E.* (1995): Accountability in organizational life : problem or opportunity for nonprofits?, in: *Nonprofit Management and Leadership* 6 (2), 181-195
- Guijt, I.* (2010): Exploding the myth of incompatibility between accountability and learning, in: Jan Ubels / N.-A. Acquaye-Baddoo / a. A. Fowler (eds.), *Capacity Development in Practice*, London: Earthscan
- Halligan, J.* (2010): NPM in Anglo-Saxon countries, in: T. Christensen / P. Lægrend (eds.), *The Ashgate research companion to new public management*, Farnham and Burlington: Ashgate Publishing Company
- Health and Development Networks* (n.d.): What is the Monterrey Consensus?; online: [http://siteresources.worldbank.org/KFDLP/Resources/461197-1122319506554/What\\_is\\_the\\_Monterrey\\_Consensus.pdf](http://siteresources.worldbank.org/KFDLP/Resources/461197-1122319506554/What_is_the_Monterrey_Consensus.pdf) (accessed 30 June 2014)
- Holzapfel, S.* (2014): The role of indicators in development cooperation : an overview study with a special focus on the use of key and standard indicators, Bonn: DIE (Studies 81); online: [http://www.die-gdi.de/uploads/media/Studies\\_81.pdf](http://www.die-gdi.de/uploads/media/Studies_81.pdf)
- Hudson, A. / L. Jonsson* (2009): Donor domestic accountabilities : UK case study, London: Overseas Development Institute
- IDA (International Development Association)* (2002): Measuring outputs and outcomes in IDA countries; online: [http://www.worldbank.org/ida/papers/IDA13\\_Replenishment/measuring%20outputs.pdf](http://www.worldbank.org/ida/papers/IDA13_Replenishment/measuring%20outputs.pdf) (accessed 30 June 2014)
- IDB (Inter-American Development Bank)* (2013): Mid-term evaluation of IDB-9 commitments; online: <http://idbdocs.iadb.org/wsdocs/getdocument.aspx?docnum=37653359> (accessed 9 Dec. 2013)
- IEG (Independent Evaluation Group)* (2011): IEG annual report 2011 : results and performance of the World Bank Group volume I : main report, Washinton, D.C.: The World Bank Group
- (2012): Results and performance of the World Bank Group 2012 : volume I, Washington, D.C.: The World Bank Group
- IFC (International Finance Corporation)* (2013a): IFC's development goals (IDGs); online: [http://www.ifc.org/wps/wcm/connect/c7a466004c3e205e8939cbf12db12449/IFC\\_DOTS\\_factsheet\\_DevelopmentGoals\\_8-3.pdf?MOD=AJPERES](http://www.ifc.org/wps/wcm/connect/c7a466004c3e205e8939cbf12db12449/IFC_DOTS_factsheet_DevelopmentGoals_8-3.pdf?MOD=AJPERES)
- (2013b): IFC development goals online: [http://www.ifc.org/wps/wcm/connect/Topics\\_Ext\\_Content/IFC\\_External\\_Corporate\\_Site/IDG\\_Home/IFCDevelopmentGoals/](http://www.ifc.org/wps/wcm/connect/Topics_Ext_Content/IFC_External_Corporate_Site/IDG_Home/IFCDevelopmentGoals/) (accessed 29 July 2013)
- (2013c): IFC development goals (IDGs) overview : IDG2 : health and education; online: [http://www.ifc.org/wps/wcm/connect/d21460804bc9ade78649e71be6561834/IDG2+H\\_E+flyer.pdf?MOD=AJPERES](http://www.ifc.org/wps/wcm/connect/d21460804bc9ade78649e71be6561834/IDG2+H_E+flyer.pdf?MOD=AJPERES) (accessed 9 Dec. 2013)
- (2013d): IFC development goals (IDGs) overview : IDG3 : financial services; online: <http://www.ifc.org/wps/wcm/connect/1d2a21004bcb4f538fbaef1be6561834/idg+3+flyer+financial.pdf?MOD=AJPERES> (accessed 9 Dec. 2013)
- Jackson, P.* (2012): Value for money and international development : deconstructing myths to promote a more constructive discussion; online: <http://www.oecd.org/dac/effectiveness/49652541.pdf> (accessed 22 July 2014)
- Kabba, O. et al.* (2002): Statement of heads of multilateral development banks : better measuring, monitoring, and managing for development results, International Roundtable on Better Measuring,

- Monitoring, and Managing for Development Results, Washington, D.C.; online: <http://www.mfdr.org/documents/MDbProgressReportonMfDRAnnexA.pdf>
- Kaplan, R. S. / D. P. Norton (1992): The balanced scorecard - measures that drive performance, in: *Harvard Business Review*, 70–79
- (2007): Using the balanced scorecard as a strategic management system, in: *Harvard Business Review* 85, 1–14
- Klingebiel, S. (2012): Accountability and the effectiveness of development cooperation: German Development Institute/Deutsches Institut für Entwicklungspolitik (DIE); online: [http://www.die-gdi.de/uploads/media/Accountability\\_Klingebiel.pdf](http://www.die-gdi.de/uploads/media/Accountability_Klingebiel.pdf) (accessed 29 April 2014)
- Koeberle, S. / Z. Stavreski (2006): Budget support : concept and issues, in: S. Koeberle / Z. Stavreski / J. Walliser (eds.), Budget support as more effective aid? Recent experiences and emerging lessons, Washington, D.C.: World Bank
- Lancaster, C. (1999): Aid effectiveness in Africa : the unfinished agenda, in: *Journal of African Economics* 8 (4), 487–503
- Leeuw, F. / J. Vaessen (2009): Impact evaluations and development NONIE guidance on impact evaluation, Washington, D.C.: The Network of Networks on Impact Evaluation
- Lensink, R. / H. White (2011): Are there negative returns to aid?, in: *Journal of Development Studies* 37 (6), 42–65
- Meier, W. (2003): Results-based management : towards a common understanding among development cooperation agencies, Discussion Paper prepared for the Canadian International Development Agency, Performance Review Branch, for consideration by the DAC Working Party on Aid Effectiveness and Harmonisation; online: <http://www.managingfordevelopmentresults.org/documents/Results-BasedManagementDiscussionPaper.pdf>
- MfDR (*Managing for Development Results*) (n.d.): About managing for development results; online: <http://www.mfdr.org/1About.html> (accessed 30 June 2014)
- Ministry of Finance, Nepal (2010): Joint evaluation of the implementation of the Paris Declaration, phase II; online: <http://www.oecd.org/countries/nepal/47671748.pdf> (accessed 14 July 2014)
- Natsios, A. (2010): The clash of the counter-bureaucracy and development: Center for Global Development; online: [http://www.cgdev.org/files/1424271\\_file\\_Natsios\\_Counterbureaucracy.pdf](http://www.cgdev.org/files/1424271_file_Natsios_Counterbureaucracy.pdf) (accessed 5 June 2013)
- Nuscheler, F. (2008): Die umstrittene Wirksamkeit der Entwicklungszusammenarbeit, Duisburg: Universität Duisburg. Institut für Entwicklung und Frieden (INEF-Report 93/2008)
- Obser, A. (2007): Multilateral Organisations Performance Assessment: Opportunities and Limitations for Harmonisation among Development Agencies, Bonn: German Development Institute / Deutsches Institut für Entwicklungspolitik (DIE) (Discussion Paper 19/2007)
- ODI (*Overseas Development Institute*) (2011): A results take-over of aid effectiveness? How to balance multiple or competing calls for more accountability; online: <http://www.odi.org.uk/events/2715-results-take-over-aid-effectiveness-balance-multiple-competing-calls-more-accountability> (accessed 25 March 2014)
- OECD (*Organisation for Economic Co-operation and Development*) (2011a): 2011 survey on monitoring the Paris Declaration : glossary; online: <http://www.oecd.org/site/dacsmpl1/glossary.htm> (accessed 26 Sept. 2013)
- (2011b): Aid effectiveness 2005–10 : progress in implementing the Paris Declaration; online: <http://effectivecooperation.org/files/resources/2011%20Report%20on%20Monitoring%20the%20Paris%20Declaration%20ENGLISH.pdf> (accessed 28 June 2013)
- /DAC (*Development Assistance Committee*)(2005/2008): The Paris Declaration on Aid Effectiveness and the Accra Agenda for Action; online: <http://www.oecd.org/development/effectiveness/34428351.pdf> (accessed 10 Feb. 2013)

- (2008a): Emerging good practice in managing for development results : Sourcebook, 3rd edition; online: <http://www.mfdr.org/Sourcebook/3rdEdition/SourceBook3FINAL.pdf> (accessed 10 July 2013)
  - (2008b): Reporting directives for the creditor reporting system : corrigendum on programme-based approaches; online: <http://www.oecd.org/dac/stats/44479916.pdf> (accessed 23 July 2014)
  - (2009): Glossary of key terms in evaluation and results based management : Glossar entwicklungspolitischer Schlüsselbegriffe aus den Bereichen Evaluierung und ergebnisorientiertes Management; online: <http://www.oecd.org/development/evaluation/dcdndep/43184177.pdf> (accessed 28 Apr. 2013)
- Office of Management and Budget* (1993): Government performance results act of 1993; online: <http://www.whitehouse.gov/omb/mgmt-gpra/gplaw2m#h4> (accessed 30 June 2014)
- OHCHR/ CESR (Office of the United Nations High Commissioner for Human Rights/Center for Economic and Social Rights)*(2013): Who will be accountable? Human rights and the post-2015 development agenda, New York and Geneva: United Nations
- Pidd, M.* (2005): Perversity in public service performance measurement, in: *International Journal of Productivity and Performance Management* 54 (5/6), 482–493
- Posner, P. L.* (2006): Accountability institutions and the policy process : the United States experience, in: *OECD Journal on Budgeting* 5 (3), 71-96
- Prennushi, G. / G. Rubio / K. Subbarao* (2002): Monitoring and evaluation, in: J. Klugman (ed.), A sourcebook for poverty reduction strategies : volume 1 : core techniques and cross-cutting issues, Washington, D.C.: The World Bank
- Radelet, S.* (2006): A primer on foreign aid: Center for Global Development; online: [http://www.cgdev.org/files/8846\\_file\\_WP92.pdf](http://www.cgdev.org/files/8846_file_WP92.pdf) (accessed 22 May 2014)
- Ramalingam, B.* (2013): Aid on the edge of chaos : rethinking international cooperation in a complex world, Oxford, United Kingdom: Oxford University Press
- Roberts, D. / N. Khattri* (2012): Designing a results framework for achieving results : a how-to guide, Washington, D.C.: Independent Evaluation Group, World Bank
- Saltmarsh, D. / M. Ireland / J. A. McGregor* (2003): The performance framework : a systems approach to understanding performance management, in: *Public administration and development* 23, 445–456
- Savedoff, W. D.* (2011): Incentive proliferation? Making sense of a new wave of development programs: Center for Global Development; online: [http://www.cgdev.org/sites/default/files/1425405\\_file\\_Savedoff\\_Incentive\\_Proliferation\\_FINAL\\_0.pdf](http://www.cgdev.org/sites/default/files/1425405_file_Savedoff_Incentive_Proliferation_FINAL_0.pdf) (accessed 5 June 2013)
- Schacter, M.* (1999): Means... ends... indicators : performance measurement in the public sector, Policy Brief No. 3, Ottawa, Ontario, Canada: Institute on Governance
- (2001): Sector wide approaches, accountability and CIDA : issues and recommendations: Institute on Governance; online: <http://bdmta.org.ua/ta/files/library/200911041716220.Sector%20Wide%20Approaches%20%28CIDA%29.pdf> (accessed 25 April 2013)
- Schedler, A.* (1999): Conceptualizing accountability, in: A. Schedler / L. J. Diamond / M. F. Plattner (eds.), The self-restraining state : power and accountability in new democracies, Boulder and London: Lynne Rienner Publishers, 13–28
- Shamash, L. / S. Burall / B. Whitty* (2013): Resetting the aid relationship; online: <http://www.involve.org.uk/wp-content/uploads/2013/05/Resetting-the-Aid-Relationship1.pdf> (accessed 18 June 2014)
- Shand, D.* (2006): Managing fiduciary issues in budget support operations, in: S. Koeberle / Z. Stavreski / J. Walliser (eds.), Budget support as more effective aid? Recent experiences and emerging lessons, Washington, D.C.: World Bank
- Sjöstedt, M.* (2013): Aid effectiveness and the Paris Declaration : a mismatch between ownership and results-based management?, in: *Public administration and development* 33 (2), 143–155

- Smith, P.* (1995): On the unintended consequences of publishing performance data in the public sector, in: *International Journal of Public Administration* 18 (2-3), 277–310
- UN (United Nations)* (2002): Monterrey Consensus on Financing for Development; online: <http://www.un.org/esa/ffd/monterrey/MonterreyConsensus.pdf> (accessed 5 March 2013)
- (2013a): Annex 2 : integrated results and resources framework in, UNDP strategic plan, 2014-2017, New York
- (2013b): UNDP strategic plan 2014-2017, New York
- UNDG (United Nations Development Group)*(2011): Results-based management handbook : harmonizing RBM concepts and approaches for improved development results at country level; online: [http://www.ilo.org/public/english/bureau/program/dwcp/download/undg\\_rbm1011.pdf](http://www.ilo.org/public/english/bureau/program/dwcp/download/undg_rbm1011.pdf) (accessed 7 Oct. 2013)
- UNDP (United Nations Development Programme)* (2009): Handbook on planning, monitoring and evaluating for development results, New York
- US Department of State / USAID (United States Agency for International Development)* (2014a): FY 2015 annual performance plan; online: <http://www.usaid.gov/sites/default/files/documents/1870/FINAL%20-%20FY%202015%20Annual%20Performance%20Plan%202014-04-01.pdf> (accessed 16 July 2014)
- (2014b): Strategic plan FY 2014-2017; online: <http://www.state.gov/s/d/rm/rls/dosstrat/2014/pdf/index.htm> (accessed 27 May 2014)
- Vähämäki, J. / M. Schmidt / J. Molander* (2011): Review : results based management in development cooperation: Riksbankens Jubileumsfond; online: [http://www.rj.se/Documents/Rapporter/2011/RBM\\_Review\\_Feb\\_2012.pdf](http://www.rj.se/Documents/Rapporter/2011/RBM_Review_Feb_2012.pdf) (accessed 19 Aug. 2014)
- Vielajus, M. et al.* (2009): The challenge of accountability for development agencies within their own countries and before their peers : a comparative study : draft document for an AFD-DevCom workshop entitled “transparent, reliable, understandable and interactive : challenges in donors' domestic accountability”, 15-16 Dec. 2009; online: <http://www.oecd.org/dev/devcom/44251093.pdf>
- Weiers, G.* (2012): Does aid work? The challenges of operationalizing the results agenda; online: <http://naiforum.org/2012/03/the-challenges-of-operationalizing-the-results-agenda/> (accessed 11 Nov. 2013)
- Wood, B. et al.* (2011): The evaluation of the Paris Declaration : phase 2 : final report, Copenhagen: Danish Institute for International Studies
- World Bank* (2013): World Bank corporate scorecard : integrated results and performance framework; online: <http://databank.worldbank.org/data/download/World%20Bank%20Corporate%20Scorecard.pdf> (accessed 10 June 2013)

## **Annex**





Table A: List of interviewees

<b>Name</b>	<b>Job title/Unit</b>	<b>Organisation</b>	<b>Date of interview</b>
Mehdi Hussein	Senior Statistics Advisor	DFID	22 May 2013
Franco Conzato	Deputy Head, Quality of Delivery Systems Unit	Directorate-General for Development and Cooperation (DG DevCo) – EuropeAid	29 May 2013
Milena Reinfeld	Statistics Advisor, Quality of Delivery Systems Unit	DG DevCo – EuropeAid	29 May 2013
GisuMohadjer	Manager, Results Unit	World Bank	4 June 2013
Lisandro Martin	Results Specialist Operations Officer	World Bank	4 June 2013
Vincent Fruchart	Consultant	Institutional Systems	7 June 2013
Johannes Jütting	Manager	Paris 21 Secretariat	10 July 2013
Hans Lundgren	Team Leader, Results and Evaluation, Development Co-operation Directorate	OECD	10 July 2013
Debbie Chappat	Policy Analyst, Development Co-operation Directorate	OECD	20 July 2013
Patricia Meduna	Office of Strategic Planning and Development Effectiveness	IDB	22 July 2013
Ana Maria Torres	Results Measurement Specialist, Development Impact Unit, Investment Services Unit	IFC	22 July 2013
Ugo Amoretti	Private Sector Development Specialist, Development Effectiveness Unit	IFC	22 July 2013
Lita N. Echiverri	International Trade Specialist	USAID	22 July 2013
Georg Weiers	Principle Results Specialist	AfDB	12 November 2013
Amy Lewis	Office of Strategic Planning and Development Effectiveness	IDB	14 November 2013
Bernard Woods	Principle Results Measurement Specialist, Strategy and Policy Department, Results Management Unit	ADB	22 November 2013



Publications of the German Development Institute /  
Deutsches Institut für Entwicklungspolitik (DIE)

Studies

- 85      *Zelli, Fariborz / Daniela Erler / Sina Frank / Jonas Hein / Hannes Hotz / Anna-Maria Santa Cruz Melgarejo*: Reducing emissions from deforestation and forest degradation (REDD) in Peru: a challenge to social inclusion and multi-level governance, 178 pp., Bonn 2014, ISBN 978-3-88985-651-7
- 84      *Fues, Thomas / Jiang Ye (eds.)*: The United Nations Post-2015 Agenda for Global Development: perspectives from China and Europe, 373 pp., Bonn 2014, ISBN 978-3-88985-649-4
- 83      *Horstmann, Britta / Günther Schulz-Heiss*: Providing international adaptation finance for vulnerable communities: a study on potentials and limits of social investment funds, 106 pp., Bonn 2014, ISBN 978-3-88985-647-0
- 82      *Knörich, Jan / Axel Berger*: Friends or foes? Interactions between Indonesia's international investment agreements and national investment law, 158 pp., Bonn 2014, ISBN 978-3-88985-589-3

[Price: EUR 10.00; publications may be ordered from the DIE or through bookshops.]

Discussion Papers

- 30/2014    *Brandt, Hartmut / Michael Brüntrup*: Erfahrungen und Perspektiven nachhaltiger Intensivierung der Landwirtschaft in subsaharischen Ländern, 80 pp., Bonn 2014, ISBN 978-3-88985-656-2
- 29/2014    *Altenburg, Tilman*: From combustion engines to electric vehicles: a study of technological path creation and disruption in Germany, 44 pp., Bonn 2014, ISBN 978-3-88985-654-8
- 28/2014    *Lütkenhorst, Wilfried / Tilman Altenburg / Anna Pegels / Georgeta Vidican*: Green industrial policy: managing transformation under uncertainty, 51 pp., Bonn 2014, ISBN 978-3-88985-653-1
- 27/2014    *Leiderer, Stefan*: Who gets the schools? Political targeting of economic and social infrastructure provision in Zambia, 107 pp., Bonn 2014, ISBN 978-3-88985-652-4
- 26/2014    *Bhasin, Shikha*: Enhancing international technology cooperation for climate change mitigation: lessons from an electromobility case study, 40 pp., Bonn 2014, ISBN 978-3-88985-650-0
- 25/2014    *Keijzer, Niels / Heiner Janus*: Linking results-based aid and capacity development support: conceptual and practical challenges, 23 pp., Bonn 2014, ISBN 978-3-88985-648-7

[Price: EUR 6.00; publications may be ordered from the DIE or through bookshops.]

For a complete list of DIE publications:  
[www.die-gdi.de](http://www.die-gdi.de)