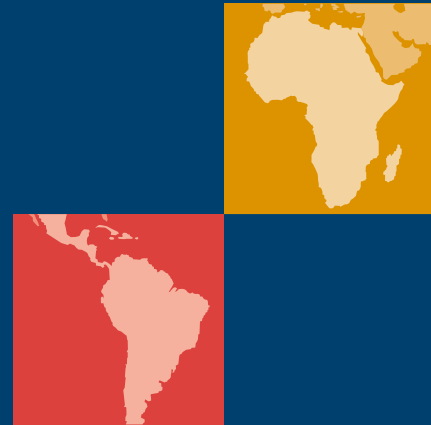


DISCUSSION PAPER / 2009.01



Monitoring and Evaluation at the sector level

**Experiences from Joint Sector Reviews
in the education sectors of
Burkina Faso, Mali and Niger.**

**Nathalie Holvoet
Liesbeth Inberg**



University
of Antwerp



INSTITUTE OF DEVELOPMENT
POLICY AND MANAGEMENT

Comments on this Discussion Paper are invited.
Please contact the authors at <nathalie.holvoet@ua.ac.be>
or at <liesbeth.inberg@ua.ac.be>

Instituut voor Ontwikkelingsbeleid en -Beheer
Institute of Development Policy and Management
Institut de Politique et de Gestion du Développement
Instituto de Política y Gestión del Desarrollo

Postal address:	Visiting address:
Prinsstraat 13	Lange Sint-Annastraat 7
B-2000 Antwerpen	B-2000 Antwerpen
Belgium	Belgium

Tel: +32 (0)3 275 57 70
Fax: +32 (0)3 275 57 71
e-mail: dev@ua.ac.be
<http://www.ua.ac.be/iob>

Monitoring and Evaluation at the sector level

Experiences from Joint Sector Reviews
in the education sectors of
Burkina Faso, Mali and Niger.

Nathalie **Holvoet** *

Liesbeth **Inberg** **

June 2009

* Nathalie Holvoet is a lecturer at the Institute of Development Policy and Management (IOB),
University of Antwerp

** Liesbeth Inberg is a researcher at the Institute of Development Policy and Management (IOB),
University of Antwerp.

TABLE OF CONTENTS

	ABSTRACT	6
	RÉSUMÉ	6
1.	INTRODUCTION	7
2.	CONCEPTUAL FRAMEWORK	10
3.	DATA, METHODOLOGY AND SETTING	15
3.1.	Data and Methodology	15
3.2.	The setting	16
4.	FINDINGS	22
4.1.	Accountability needs	22
4.2.	Learning Needs	24
4.3.	Paris Reform Agenda	25
5.	ISSUES FOR DISCUSSION	28
5.1.	Balancing ‘substance’ and ‘systemic’ issues	28
5.2.	Balancing ‘accountability’ and ‘learning’	30
5.3.	JSR as a forum for coordination and harmonisation	33
5.4.	JSR and feedback	34
6.	CONCLUSIONS	35
	REFERENCES	38

ABSTRACT

Aid policy and practice have been thoroughly shaken up over the past few years. One of the crucial reform areas relates to monitoring and evaluation (M&E). In short, aid recipients are asked to elaborate sound, results-oriented frameworks while donors are expected to harmonise and align their own policies and frameworks. Evidence has revealed that implementation on the ground is slow. This discussion paper examines the extent to which joint sector reviews (JSRs) could take the M&E reform agenda forward. JSRs are M&E exercises at the sector level which have the potential to satisfy the existing M&E needs of various stakeholders while, at the same time, also contributing to the M&E reform agenda. They are increasingly utilised on the ground, yet, so far, there do not exist any systematic stocktakings and/or analyses of them. Our own analysis of a sample of JSRs from the education sectors of Burkina Faso, Mali and Niger indicate that JSRs score highly on harmonisation, coordination, leadership and broad-based participation; but poorly on alignment. They generally prioritise accountability over learning needs; focus more on implementation than results; and largely neglect accountability and learning at the level of the sector institutional apparatus (including the M&E system). In this paper, findings from the field are contrasted with insights from evaluation theory and practice so as to provide suggestions for on the ground JSR improvements.

RÉSUMÉ

Ces dernières années, la politique d'aide a été fortement bouleversée. Un des principaux domaines de réformes se situe au niveau du "suivi et évaluation (S&E)". En bref, les pays bénéficiaires doivent élaborer des systèmes S&E performants. Dans le même temps, les donateurs sont supposés se baser et s'aligner de plus en plus sur les systèmes et les arrangements nationaux de S&E. Jusqu'à présent l'implémentation sur le terrain de ces réformes est lente et partielle. Dans cet article, nous examinons dans quelle mesure les revues sectorielles conjointes (RSC) pourraient donner un coup de pouce aux réformes dans le domaine du S&E. Les RSC sont des exercices de S&E au niveau sectoriel qui sont de nature à satisfaire les besoins de S&E des différents acteurs concernés et qui en même temps, pourraient contribuer à la réalisation des réformes énoncées par la nouvelle approche de l'aide. Bien que les RSC soient de plus en plus utilisées sur le terrain, il n'existe pas d'analyse systématique de leur qualité. Notre propre recherche des RSCs dans les secteurs d'éducation du Burkina Faso, du Mali et du Niger montre que les RSCs favorisent les besoins de rendre compte au détriment de la rétro-action, et qu'elles mettent plus l'accent sur l'implémentation des interventions que sur les résultats. Elles adhèrent parfaitement aux principes d'harmonisation, de coordination, de direction par les autorités bénéficiaires et de large participation des acteurs concernés mais elles ne s'alignent pas sur les systèmes S&E sectoriels existants qu'elles négligent largement.

1. INTRODUCTION

An important shift is taking place in aid to low-income countries. A growing number of donors is moving towards budget support where, in contrast to traditional projects, the control over the use of aid resources passes through the recipient government. The 2005 Paris Declaration (PD) and the 2008 Accra Agenda for Action (AAA) are illustrations of the growing consensus in this respect. They set out the reform agenda for donors and recipients, and propose a monitoring framework, composed of 12 indicators. One of the crucial reform areas relates to monitoring and evaluation (M&E). In short, the expectations are that “partner countries endeavour to establish results-oriented reporting and assessment frameworks that monitor progress against key dimensions of the national and sector development strategies and that these frameworks should track a manageable number of indicators for which data are cost-effectively available” (indicator 11)(OECD/DAC, 2005a: 8). From donors it is expected that “they work with partner countries to rely, as far as possible, on partner countries’ results-oriented reporting and monitoring frameworks” and that they “harmonise their monitoring and reporting requirements, and, until they can rely more extensively on partner countries’ statistical, monitoring and evaluation systems, [work] with partner countries to the maximum extent possible on joint formats for periodic reporting” (OECD/DAC, 2005a: 8). Moreover, donors and partner countries jointly commit to “work together in a participatory approach to strengthen country capacities and demand for results based management” (OECD/DAC, 2005a: 8). These commitments are reconfirmed in the AAA.

While the principle of recipient-led, donor-supported joint M&E is generally accepted, the actual implementation of the M&E reform agenda persistently stands out as one of the reform areas where progress is slow. Donors still overwhelmingly rely on their own institutional apparatus, only harmonise slowly with other donors and hardly align with the national M&E apparatus which they consider weak. Results of the 2006 Survey of the Paris Declaration highlight that harmonisation among donors is about 18% and 42% for donor missions and country analytical work respectively (OECD/DAC, 2007: 33-34). The percentage for donor missions would even be worse if joint ‘public finance management (PFM)’ missions were excluded (OECD/DAC, 2007: 33-34). The latest 2008 PD survey does not record much progress: joint activities still linger around 20% for missions and 40% for country analytical work. Donor’s alignment with country systems for PFM and procurement has only increased marginally from 40% and 39% in 2005 to 45% and 43% in 2007 (OECD/DAC, 2008: 20). An earlier survey by the OECD/DAC furthermore revealed that by the end of 2004 on average 28% of donor project portfolio was monitored and evaluated through recipient’s M&E systems (OECD/DAC, 2005c).

The recent evaluation of the implementation of the Paris Declaration (Wood et al, 2008) highlights that improvements in the use of countries systems are mainly in the areas of financial management, audit and procurement, in stronger sectors and by donors that decentralise their aid operations. While some donors have chosen to improve country systems by using them even if they are still imperfect, other donors are more reluctant to rely on country systems, even if they are known to be well developed. Reasons for this reluctance are related to policies and restrictions imposed by headquarters, remaining concerns about fiduciary risks and pressures to disburse funds faster than those systems allow.

Donor's worries regarding the use of recipient M&E systems are not ungrounded; the 2006 and 2008 PD surveys show that only 2 out of the 29 and 3 (Mozambique, Tanzania and Uganda) out of the 54 countries surveyed respectively had results-oriented frameworks that were deemed adequate (results for indicator 11)(OECD/DAC, 2007: 35; OECD/DAC, 2008: 58-59). Nevertheless, the percentage of countries that are working to improve their monitoring frameworks has increased from 42% to 56% (WP-EFF, 2008). Indicator 11, which is based on the Comprehensive Development Framework (CDF) reports (see World Bank, 2007, 2005, 2003), is composed of three sub-components: the quality of the information available, its accessibility to relevant stakeholders and the existence of a co-ordinated country-level M&E system. There are interesting variations in the scores on the different sub-indicators with improvements in the area of data generation, sharing and dissemination. However, progress in establishing coherent, co-ordinated and fully-functioning M&E institutions is lagging behind (see GTZ/BMZ 2004; World Bank 2003, 2005), even in countries where substantial improvement in sub-components such as statistical data capacities are made (World Bank, 2005). Similarly, the Evaluation of the Implementation of the Paris Declaration indicates that even though statistical reporting is improving in some countries, the next step, i.e. the actual use of statistical data as a basis for better decision-making, often remains difficult. According to the same evaluation, leadership in the country is a key factor in the use of information and the establishment of a quality system (Wood et al, 2008).

Donors and recipients are obviously caught in a kind of chicken-and-egg dilemma. Donors are not eager to align with systems that are at best only partially developed while the elaboration and maturing of recipient systems is blocked by the same donor reluctance to align. In order to move beyond this current catch-22 situation, reasonable and pragmatic approaches are needed. One possible scenario is to adopt a two-track approach that combines the set-up and/or strengthening of recipient M&E systems with the fulfilment of the existing M&E needs in the short and middle term. As in most cases the current recipient M&E system is not yet able to satisfy the existing M&E needs of accountability and learning, there is a need for complementary M&E exercises. The quality of such exercises and more particularly the degree to which they are able to cope with the recipient and donor M&E demands is crucial to convince donors to lower down their own separate stand-alone M&E exercises. At the same time, complementary exercises should be set up in such a way that they feed as much as possible into the M&E reform agenda of harmonisation, alignment, ownership and building of the recipient M&E system. The maturing recipient M&E system can then progressively accommodate the M&E needs, with complementary exercises gradually evolving towards a kind of meta-evaluation.

So far, there are no clear-cut recipes provided by M&E theory or longstanding similar M&E practices for the design of such complementary M&E exercises. At best, evaluation literature and practice inform us that the challenges at stake are huge. Nevertheless and by necessity, on the ground, various stakeholders involved seek solutions to the challenges confronted with. One such strand of ongoing M&E exercises are Joint Sector Reviews (JSRs). JSRs are increasingly included as essential ingredients in Sector Wide Approaches (SWAs), and in processes linked to sectoral and general budget support. JSRs are gradually functioning as the main instrument for assessing progress, resolving issues and reaching agreement on the sector policy, programme and targets and increasingly and progressively replace individual supervision missions of donor agencies (World Bank, 2001).

In spite of the increasing prevalence and importance of these processes on the ground, there has so far been little systematic review of JSRs that is taking stock of experiences, analyses strengths and weaknesses and formulates suggestions for improvement (see also Packer, 2006). This research aims at filling this gap. In doing so, we assess a sample of JSRs in the education sectors of Burkina Faso, Mali and Niger on a number of criteria. Additionally, we confront research findings with insights from evaluation theory and practice in order to provide some suggestions for improvement of JSRs. In short, our study follows the longstanding tradition of cross-fertilisation among M&E experiments in aid monitoring and evaluation practice on the one hand and the methodological insights from evaluation theory and practice on the other hand^[1].

The structure of this discussion paper is as follows: chapter two presents the conceptual framework, chapter three provides data, methodology and setting, chapter four summarises selected findings for JSRs in three selected countries, chapter five broadens the discussion to a number of key issues that may be distilled from the findings and finally chapter six concludes.

[1] See Cracknell B.E. (2000).

2. CONCEPTUAL FRAMEWORK

JSRs have their origin in SWAp. In a context of changing aid modalities, SWAp are used by partner countries and donors to channel aid to a specific sector. SWAp are most common in the health and education sector. Foster (2000) defines the characteristics of a SWAp as: “all significant funding for the sector supports a single sector policy and expenditure programme, under government leadership, adopting common approaches across the sector, and progressing towards relying on government procedures to disburse and account for all funds” (Foster, 2000: 9). Central to a SWAp is the collective responsibility of donors and governments for achievements in the sector. Different than in project aid, donors are not able to attribute their financial inputs to specific outputs; they have to justify their individual contributions in terms of progress against jointly agreed objectives in the sector (Cassels, 1997). Therefore, within the SWAp the need for joint assessment instruments has emerged.

Besides JSRs, joint evaluations, i.e. evaluations to which different donor agencies and/or partners participate (OECD/DAC, 2002), could be used for the assessment of performance in a sector. While the advantages of joint evaluations, such as harmonization and reduced transaction costs, are clear, in practice, joint evaluations are not yet common practice. In the OECD/DAC guidance for managing joint evaluations (2006:12) two challenges for effective joint evaluations are identified:

- “Subjects that are especially suited to joint evaluations – such as co-financed programme support at the budget or sector level, multilateral agency effectiveness and evaluating the impact of several actors – are more difficult to evaluate than traditional single agency projects.
- Processes for coordinating joint work can be complex and can increase the cost and duration of the evaluation”

Recent examples of joint evaluations are the evaluation of General Budget Support (IDD and Associates, 2006) and the evaluation of the implementation of the Paris Declaration (Wood et al, 2008).

Even though JSRs are commonplace in SWAp, so far no standard definition, handbooks or blueprints for JSR exist. A review, defined by the OECD/DAC (2002: 34) as “an assessment of the performance of an intervention, periodically or an ad hoc basis”, may be considered an instrument which is situated between monitoring^[1] and evaluation^[2]. While data provided in monitoring does not give insight in causes and attribution of change, the assessment in a review is less comprehensive and in depth than an evaluation and emphasizes operational aspects (OECD/DAC, 2002). Thus, a JSR may be described as a type of joint periodic assessment of performance in a specific sector with the aim to satisfy donor and recipient’s accountability and learning needs. ‘Performance’ is to be interpreted broadly and may include both a focus on substance at various levels (inputs, activities, output, outcome and impact) as well as on the underlying, systemic and institutional issues.

[1] OECD/DAC (2002: 27) defines monitoring as “a continuing function that uses systematic collection of data on specified indicators to provide management and the main stakeholders of an ongoing development intervention with indicators of the extent of progress and achievement of objectives and progress in the use of allocated funds”.

[2] OECD/ DAC (2002: 21) defines evaluation as “the systematic and objective assessment of a non-going or completed project, programme or policy, its design, implementation and results”.

A number of objectives of joint reviews are listed by Packer (2006: 2):

- Increase country ownership
- Give more effective support to national priorities
- Promote more efficient division of labour among aid agencies
- Improve frameworks for monitoring and evaluation
- Improve accountability to funding sources and government partners
- Improve efficiency and transparency through harmonized monitoring and evaluation systems
- Improve the aggregate effectiveness of aid.

Packer adds that there has not been any systematic comparative analysis to find out whether joint reviews on the ground actually (partly) meet these objectives.

In most cases JSRs are organized once or twice a year and take about one to two weeks. They are led and coordinated by the sector ministry, often in cooperation with the lead donor in the sector. JSRs engage a broad range of state and non-state stakeholders who are subdivided over several working groups which focus on specific topics (such as quantity and quality of outputs and outcomes, public finance management, human resources, management information system). The documentary input into JSRs diverges from country to country and includes at best a combination of various primary and secondary data sources. In most of the cases a sector performance report has been prepared by the sector ministry, including as well financial reporting from the finance department or the Ministry of Finance (and/ or sector finance departments). This is at times combined with additional (commissioned) studies on specific topics that were identified during previous JSRs and information from 'project' donors or civil society organisations active in the sector. Sometimes field missions are also planned during the JSR. Evidence from all these sources feeds into discussions in several technical working groups. Conclusions and recommendations from these discussions are mostly brought together at the Annual Review Meeting (ARM) where they are shared with all stakeholders. The Aide Mémoire (AM) is the main documentary output of the JSR.

Because there exist no clear-cut guidelines or specific criteria for JSR, which could function as the basis for a review exercise, selection of issues for analysis has mainly been driven by our research question, i.e. the extent to which JSRs satisfy existing M&E needs of accountability, learning and are in accordance with the reform agenda of the Paris Declaration. Additionally, we have also consulted guidelines that exist for evaluation exercises in general and joint exercises in particular (OECD/DAC, 1991; 1998; 2000; 2005b) and we have reviewed the 1997 WHO/UNDP guide on SWAPs as well (Cassels, 1997). This resulted in the following framework of analysis.

Table 1: Framework of Analysis

Broad objectives	Specific Issues
I. Accountability needs	<p><i>How and to what extent does the JSR cover accountability at the level of:</i></p> <ul style="list-style-type: none"> - substance at aggregate sectoral level <ul style="list-style-type: none"> · sector inputs (financial and human resources, quality of policy) · sectoral activities and outputs (quantity and quality of schooling, effectiveness, cost-effectiveness) · outcomes and impact (utilization, inequality in utilization, literacy) - elaboration and functioning of the institutional apparatus at the aggregate sector level
Learning needs	<p><i>How and to what extent does the JSR address learning related to:</i></p> <ul style="list-style-type: none"> - substance at the sectoral level <ul style="list-style-type: none"> · linkages among inputs – activities · linkages among activities – outputs · linkages among outputs – outcomes – impact - operational apparatus at the overall sectoral level
Reform agenda	<p><i>How and to what extent does the JSR promote the PD reform agenda?</i></p> <ul style="list-style-type: none"> - harmonization/coordination at sectoral level (including among various types of donors at sectoral level) - harmonization/coordination with national (PRSP) review processes & PAF - alignment with the existing sector M&E frameworks <ul style="list-style-type: none"> · use of information from the existing M&E system, from the routine data collection, the MIS - country leadership/ownership - broad participation of inside & outside government actors - capacity building of the M&E supply & demand side <ul style="list-style-type: none"> · line ministry · civil society, universities, M&E networks, parliament, audit offices - mutual accountability

3. DATA, METHODOLOGY AND SETTING

3.1. Data and Methodology

The selection of the sample of JSRs has been guided by a number of criteria. Firstly, the research is in the context of Belgian development cooperation, thus we selected JSRs in which Belgium participated. Secondly, as new aid modalities emphasize the incremental improvement of programmes and processes, we have selected our sample with a view to including a longitudinal perspective. Thirdly, we have tried to balance between internal and external validity^[1], with a slight preference for internal validity as there does not exist much prior research on the same topic. Therefore, we have concentrated on one sector and on similar countries. In order to increase at the same time external validity as much as possible, we have included more than one country and selected the sector where Belgium focuses on most prominently in the context of sector support. The combination of all these criteria has finally led to a selection of JSRs in the education sectors of Burkina Faso, Mali and Niger. Our research results are mainly applicable to the sample selected, but the more general findings and discussion are interesting for other sectors and other countries as well.

In order to increase the construct validity^[2] of our findings we have opted for a combination of a desk and field study. Our desk study includes a review of the Aide Mémoires, i.e. the main 'written' output of the JSRs, according to a number of specific criteria. Additionally, we have participated on the ground in one JSR mission and ARM. This provided us with an opportunity to become better acquainted with the process and to interact with several of the stakeholders that took part in the JSR.

Table 2 gives an overview of the specific JSR exercises that were included in the desk study. Field study has been done in the framework of the 9th JSR of the education sector of Burkina Faso (10-16 April 2007).

Table 2: Identifying the sample

Country	Sector programme	Specific JSRs
Burkina Faso	Plan Décennal de Développement de l'Éducation de Base (PDDEB)	2003, 2004, May/June 2005, Dec. 2005, April/May 2006, Nov. 2006, April 2007, Nov. 2007
Mali	Programme d'Investissement Sectoriel en Éducation (PISE)	2002, January 2003, June 2003, 2004, 2007
Niger	Programme Décennal de Développement de l'Éducation (PDDE)	2003, 2004, 2005, 2006, 2007

[1] Internal validity is the validity of inference about whether observed co variation between A (the presumed treatment) and B (the presumed outcome) reflects a causal relationship from A to B as those variables were manipulated or measured. External validity is the validity of inference about whether the cause-effect relationship holds over variation in persons, settings, treatment variables, and measurement variables (Shadish, W., Cook T. and Campbell D., 2002: 38).

[2] Construct validity is the validity of inference about the higher order constructs that represent sampling particulars (Shadish, W., Cook T. and Campbell D., 2002: 38).

3.2. The setting

Burkina Faso, Mali and Niger are among the poorest countries in the world, ranking 176, 173th, 174th respectively (out of 177 countries with data) on the Human Development Index (UNDP, 2007). The three countries are included in the 2006 and 2008 surveys of the Paris Declaration (PD). The table below summarises the scores on the PD-indicators.

Table 3: Results on some of the indicators of the Paris Declaration for 2006 and 2008

	Burkina Faso		Mali		Niger		Global (a)	
	2006	2008	2006	2008	2006	2008	2006	2008
Aid reported in the survey (USD m)	593	827	625	811	393	428		
Indicator 1: Operational development strategy (b)	C	B	C	C	C	C	B: 12.5% C: 57.5% D: 30.0%	B: 20% C: 65.0% D: 12.5% E: 2.5%
Indicator 2a: Reliability of PFM systems (c)	4	4	4	3,5	3,5	3,5	4.5: 2% 4.0: 17% 3.5: 37% 3.0: 17% 2.5: 20% 2.0: 7%	4.0: 29% 3.5: 32% 3.0: 22% 2.5: 7% 2: 10%
Indicator 3: Comprehensive and realistic government budget estimates (%)	68	92	60	73	99	91	42	48
Indicator 4: Co-ordinated TA (%)	3	56	15	75	15	45	48	60
Indicator 5a: Aid using country systems, PFM (%)	45	43	29	34	27	26	40	45
Indicator 5b: Aid using country systems, procurement (%)	60	54	45	35	49	37	39	43
Indicator 6: Parallel PIUs	131	102	65	60	52	47	1.817	1.601
Indicator 7: Disbursements on schedule (%)	92	92	71	68	73	78	41	46
Indicator 8: Untied bilateral aid (%)	92	92	95	93	84	84	82	87
Indicator 9: Programme based aid (%)	45	57	48	41	31	49	43	46
Indicator 10a: Co-ordinated mission (%)	17	13	7	15	21	15	18	20
Indicator 10b: Co-ordinated country analysis (%) *	39	39	26	39	35	32	42	42
Indicator 11: Monitorable PAF (d)	C	C	D	D	D	D	B: 5% C: 47.5% D: 47.5%	B: 8% C: 65% D: 30% E: 3%
Indicator 12: Mutual accountability (no countries)	No	no	no	No	no	no	12	13

* There are differences in the scores for 2006 when comparing the 2006 and 2008 reports. The percentages represented here are those from the 2008 report. The percentages mentioned in the 2006 report are 45, 30 and 40 respectively.

a Global score of countries participating in both 2006 and 2008 survey

b There are 5 scores from A to E, with A meaning that the National Development Strategy substantially achieves good practice and E meaning that the National Development Strategy reflects little action toward achieving good practice.

c These are CPIA ratings^[1] varying between 1.0 (very weak), and 6.0 (very strong), with 2.0 (weak), 3.0 (moderately weak), 4 (moderately strong) and 5.0 (strong).

d There are five scores (A, B, C, D and E) with A being very strong and E very weak.

[1] The Country Policy and Institutional Assessment (CPIA) is a diagnostic tool that is intended to capture the quality of a country's policies and institutional arrangements. The CPIA consists of 16 criteria grouped in four equally weighted clusters: Economic Management, Structural Policies, Policies for Social Inclusion and Equity, and Public Sector Management and Institutions (see Box below). For each of the 16 criteria, countries are rated on a scale of 1 (low) to 6 (high) (www.worldbank.org).

The table shows a mixed picture, with generally slightly better scores for Burkina Faso than for Mali and Niger. Burkina Faso improved the score for indicator 1 from C to B thanks to the establishment of a strategic framework for development with clear priorities and the introduction of a Medium Term Expenditure Framework (MTEF) (OECD/DAC, 2008: 30). However, findings demonstrate at the same time that in Burkina Faso and Niger the scores for the indicators in the areas of harmonisation (10a and 10b) and alignment (5a and 5b) decreased or remained the same between the 2006 and 2008 survey. Even though Mali improved its status on indicators 5a, 10a and 10b (decrease in score 5b), its scores are still behind those of Burkina Faso and the global scores.

Scores on indicator 11 highlight that none of the three countries has an adequate results orientated framework and no improvements over time are observed. CDF reports, on which the scores of indicator 11 are based, provide more detailed information on the three sub-indicators (i.e. the quality of the available information, its accessibility to relevant stakeholders and the existence of a co-ordinated country-level M&E system). According to the LEADS^[1]-scoring method, Burkina Faso has taken action on all three dimensions of an M&E system (score 'A'). Mali and Niger score similarly on 'coordinated country-level M&E' while both obtain a lower score ('E'; elements exist) for 'quality of development information' and 'stakeholders access to information'.

The JSRs under study are all related to the education sector of Burkina Faso, Mali and Niger. In what follows we present selected data on sector performance and aid as well as a short overview of the education policies in the three countries.

The (net) enrolment rates for primary, secondary and tertiary education of Burkina Faso, Mali and Niger are lower than the average rates of countries in Sub Sahara Africa in 2006. Mali shows a slightly better score on primary education than Burkina Faso and Niger (no data for Mali on secondary and tertiary education), due to remarkable improvements in the last decade (see table 4).

Table 4: Net enrolment rate in primary education

		1991	2006
Sub Sahara Africa	Total		70
	Male		73
	Female		67
Burkina Faso	T	27	47
	M	33	52
	F	21	42
Mali	T	25	61
	M	31	67
	F	18	54
Niger	T	24	43
	M	30	50
	F	18	37

Source: <http://stats.uis.unesco.org>

[1] The LEADS 'scoring system' reads as follows: Little action (L): Due to a wide variety of circumstances, including political developments, capacity constraints and unforeseen events, action has remained at a virtual standstill; Elements exist (E): There is some basis for making progress, either through what already exists, or definite plans; Action taken (A): Progress is being made, although not yet enough, and the basis exists for even more substantive progress; Largely developed (D): Significant action taken already, although further action is needed and Sustainable (S): There are no warning signs of possible deterioration, and there is widespread expectation that the progress achieved is sustainable.

The three countries also score below average on the literacy rates (adult 15+, average is 62.1: 71.0 male, 53.7 female), with the lowest rate for Mali: 22.9 in 2006 (31.1 male, 15.6 female). The literacy rates for Burkina Faso and Niger stand at 26.0 (34.3 male, 17.9 female) and 29.8 (43.8 male, 15.8 female) respectively.

These figures demonstrate that the challenges for these countries to reach the education related Millennium Development Goals (MDG)^[1] are huge. However, all three countries have developed specific education plans^[2] aimed at attaining them by 2015. These plans are quite similar: all consist of three components: improvement of access, improvement of quality and strengthening of the institutional capacity of the Ministry of Education. In order to attain the objectives formulated in these plans, substantial investments are needed. Table 5 gives an overview of the public expenditures that are currently being targeted to the education sector. Besides striking similarities in plans and policies, the relative budgetary allocation to the education sector as well as the division over the different sub-sectors is highly similar in the three countries.

Table 5: Public expenditure on education in Burkina Faso, Mali and Niger, 2006

	Burkina Faso	Mali	Niger
Public expenditure on education			
as % of GDP	4.2	4.5	3.4
as % of total government expenditure	15.4	16.8	17.6
Distribution of public expenditure per level (%)			
Primary	66	60	64
Secondary	21	27	25
Tertiary	11	12	11

Source: <http://stats.uis.unesco.org>

Aid to the education sector is highest in Burkina Faso (i.e. 153 million USD in 2005 of which 56.9% for basic education or 39 USD per primary school-age child). The corresponding figures for Mali are 74 million USD (50.0 % and 16 USD), and 80 million USD (61.3% and 21 USD) for Niger. The share of education in total ODA is 16% in Burkina Faso, 8% in Mali and 12% in Niger (UNESCO, 2008).

[1] We refer particularly to goal 2 (achieve universal primary education), with the related target (ensure that, by 2015, children everywhere, boys and girls alike, will be able to complete a full course of primary schooling). Additionally, the target related to goal 3 (promote gender equality and empower women) focuses on education (eliminate gender disparity in primary and secondary education, preferably by 2005, and in all levels of education no later than 2015).

[2] Plan Décennal de Développement de l'Éducation de Base (PDDEB) in Burkina Faso, Programme Décennal de Développement de l'Éducation (PRODEC)/ Programme d'Investissement Sectoriel en Éducation (PISE) in Mali and Programme Décennal de Développement de l'Éducation (PDDE) in Niger.

4. FINDINGS

The first JSRs in the education sectors in Burkina Faso, Mali and Niger took place in 2003, 2002 and 2003 respectively. The JSRs in Mali and Niger are annual, the JSR in Burkina Faso is organised twice a year. There were no JSRs during the preparation of the new education programme in Mali. In all three countries an explicit JSR objective is to assess the execution of the education programme in the past (half) year (focus on accountability). In the JSRs of Burkina Faso and Niger another explicit objective is to provide recommendations for the improvement of the execution of the programme in the coming (half) year (focus on learning). This prospective objective was included as well in the first JSR in Mali, but in the subsequent JSRs, this element was left out. In spite of the absence of a clear mentioning of the ‘learning’ objective, recommendations for better performance were also formulated in the Aide Mémoires of all JSRs in Mali.

For the assessment of the JSRs the framework of analysis described in section two is used. Below, an overview of selected findings for the three countries is presented, subdivided over accountability needs, learning needs and the Paris reform agenda. Throughout the discussion we draw upon examples from the different countries. There is a bias towards the case of Burkina Faso where more JSRs took place so far and where desk study was complemented with field study.

4.1. Accountability needs

Table 6: Accountability needs in the JSR

	Burkina Faso	Mali	Niger
Substance	<ul style="list-style-type: none"> - Focus mainly on access and quality. - Since the 6th mission (Nov. 2006) more attention to financial inputs through the instalment of a working group on financial management. 	<ul style="list-style-type: none"> - Focus on access and quality with main emphasis on accountability at the activity level. Only one AM contains progress on (mainly) output indicators. - In most AMs little attention for financial inputs. 	<ul style="list-style-type: none"> - Focus on access and quality with main emphasis on accountability at the activity level. - Accountability at the financial input level is becoming increasingly important (particularly in the latest JSRs)
Institutional/ Systemic issues	<ul style="list-style-type: none"> - In the AMs especially attention for financial management and capacity building and not really for institutional and systemic issues at other levels of the causal chain. - Since the 6th and 7th JSR (2006/2007) increasing awareness of the importance of the underlying systemic and institutional issues, without much follow-up however. 	<ul style="list-style-type: none"> - Relatively little attention to institutional and systemic issues (except for the 2nd AM). - Not much attention to M&E at central and programme level. 	<ul style="list-style-type: none"> - Increasing attention to institutional development, mainly caused by findings on lagging improvements at the substance level.

Table 6 summarizes the findings regarding the coverage of accountability needs in the JSRs. Generally there is more attention during the JSRs for accountability than for learning needs, which is obvious from the emphasis put on performance stocktaking. Accountability is mainly oriented towards issues of substance and there is a predominance of ‘implementation’ over ‘results’, and more particularly a focus on the rate of execution of the planned activities (such as building of classrooms and recruitment of teachers).

The importance given to particular topics during the JSRs is mainly related to the existence of thematic working groups on the specific topic. The 5th JSR in Burkina Faso pays e.g. much attention to infrastructure and curricula, which corresponds to the existence of working groups on these subjects. From the 6th JSR onwards, a specific working group on financial management has been installed, which has put PFM issues higher on the agenda.

Progress in the set-up and functioning of the sectoral M&E system is nowhere a major issue during the JSRs. The attention for institutional and systemic issues is however gradually increasing as a result of continuous disappointing achievements at the substance level. This is particularly the case in Niger, where the implementation of the new education programme was seriously delayed due to a reorganisation/decentralisation of the Ministry of Education. Increasing attention to systemic issues does not necessarily equate to substantial improvements as recommendations which are formulated during one JSR are not necessarily pursued during the next. This is particularly true for recommendations regarding the more difficult to grasp and sensitive institutional and systemic issues. The ToR of the 9th mission of Burkina Faso acknowledge this issue and provide instructions for the formulation of recommendations: each chapter of the Aide Mémoire is allowed to have a maximum of four recommendations, which should be able to be operationalised within six months. Furthermore, recommendations related to institutional and systemic issues should be operationalised by a technical recommendation.

4.2. Learning Needs

Table 7: Learning needs in the JSR

	Burkina Faso	Mali	Niger
Substance	<ul style="list-style-type: none"> - In objectives of latest JSRs more attention to learning needs. - Analyses are weak but improving over time. 	<ul style="list-style-type: none"> - Hardly any analyses are made. - The JSR concludes that the sector performance reports (major input of JSR) contain no analyses concerning the progress of the indicators, which makes it difficult to learn. 	<ul style="list-style-type: none"> - Analysis focuses mainly on implementation issues, mostly on issues related to the service utilization plan
Institutional/systemic issues	<ul style="list-style-type: none"> - Almost each AM contains recommendations related to M&E (for the overall sector level and for the programme level), but follow up is lacking. - Increasing attention to the improvement of the quality of the reporting (implementation reports, AMs and recommendations). 	<ul style="list-style-type: none"> - Analysis and learning regarding the institutional/systemic issues is not on the agenda of the JSRs. 	<ul style="list-style-type: none"> - Recommendations are made in every JSR, but without results. Problems are mainly situated at the level of budgeting and planning.

Table 7 provides an overview of the coverage of learning needs in the JSRs. The minimal attention given to learning during JSRs is clearly due to a lack of analysis regarding the (non)-achievement of the performance targets set. When analyses are included, they are often of poor quality but improving gradually over time, particularly in Burkina Faso. However, even in the most recent Aide Mémoire of Burkina Faso, analyses are still mainly phrased in terms of statements or recommendations.

Generally, the analytical depth of the JSR largely depends upon the analytical quality of the sectoral performance report, the major documentary input provided by the ministry of education. In all three countries the low analytical quality of the performance reports is acknowledged and over time the quality of the reports has been improved in Burkina Faso and Niger. Interestingly, the AMs of Mali and Niger recognise the value added of field missions for the improvement of the analytical depth of performance reports.

Similar to the coverage of accountability needs, analysis and learning is also predominantly centred on the 'service utilisation plan' (relationship between activities and outputs). When there is attention for progress in the elaboration and functioning of the underlying institutional apparatus, it is mainly related to Public Finance Management (PFM) issues.

4.3. Paris Reform Agenda

Table 8: Paris reform agenda and the JSR

	Burkina Faso	Mali	Niger
Coordination and Harmonisation at sector level	- Coordination and harmonisation is taking place through the 'Cadre Partenarial'.	- Especially in the 2nd AM attention to the non-harmonisation of procedures of the different development partners	- Attention to harmonisation, results in this area are made.
Harmonisation with other national review processes	- Indicators are harmonised with those of the PRSP. - Overlap of actors that participate at various processes.	- Harmonisation with the central PRSP process is an explicit objective (see first AM).	- The education programme is integrated in the PRSP and in conformity with the MDGs.
Alignment	- There is alignment on paper in the 'Cadre Partenarial'. The move towards alignment and budget support is however weakened in the final version of the 'cadre partenarial'. The effective degree of alignment is not clear.	- The degree of alignment is not clear from the documents.	- The M&E unit plays a role in alignment of information, including information from projects and programmes.
Leadership/ownership	- Joint coordination and joint presidency of Secretary General of the ministry of education and the leading donor.	- Ownership/leadership is not clear from the documents.	- The Ministry of basic education coordinates the education programme and the JSR. The Minister presides the JSRs.
Broad participation of actors	- The first AM mentions explicitly the presence of the ministry of education, donors in the education sector, social partners (labour union, teacher-parents associations), NGOs, private sector, representatives of other ministries (without further detail) and representatives of the deconcentrated structures.	- Representatives of the national ministry of education, development partners, deconcentrated and decentralised structures and civil society. Not clear whether the ministry of finance or the central M&E unit are present.	- Many donors are present, also non-budget support donors. Representatives of the Ministry of Finance and the PRSP secretariat are present.
M&E capacity building of demand & supply side	- Weak capacity is acknowledged, but no specific capacity building in M&E.	- Only attention in 2nd AM to capacity building in M&E for the decentralised 'education' structures at local level.	- No specific attention to capacity building at the level of M&E.
Mutual accountability	- Donor effectiveness not really an issue on the agenda.	- Little and in the last two AMs even no attention to donor effectiveness	- AM increasingly critical on the fulfilment by donors of agreements made in the 'Cadre Partenarial'.

Table 8 summarises the findings related to the Paris reform agenda. When it comes to the implementation of the Paris reform agenda, there is clearly more progress in coordination and harmonisation than alignment. There are efforts to harmonise sector review processes with central PRSP review processes in order to avoid the continuous existence of parallel processes. In practice, this is achieved through the use of similar indicators and the participation of key-actors of the central review processes (mostly from the ministry of finance) in sectoral review processes. The harmonisation and coordination among various types of donors at the sectoral level is a topic of discussion during JSRs. Whereas coordination is certainly not yet optimal, progress is being made. A first important step is obviously the participation in the JSRs of non-budget support donors who have projects in the education sector.

In line with the above-mentioned limited discussion regarding the systemic and institutional issues during the JSRs and the lack of specific performance targets and systematic follow-up in this area, the actual degree of alignment with the sectoral M&E system is low or at least not always clear from the JSR. Similarly, capacity building of the M&E supply and demand side does not seem to figure prominently on the agenda of the JSR. From these findings it is obvious that the use of the JSR as an instrument to feed into the elaboration and the maturing of the own national sector M&E system remains largely underexploited. There is no real systematic diagnosis of the existing M&E system or arrangements, little effective target setting regarding areas for improvement, no follow-up of recommendations, no straightforward capacity building plan and (consequently) little effective donor alignment. Reversely, leadership and broad-based participation seem to score relatively well: in all cases the ministry of education takes the lead in the JSR, alone or jointly with the sector lead donor and various groups of stakeholders, both from within and outside government, participate in the JSR. Participation is enhanced through the organisation of discussions in different working groups which are composed of various actors.

While JSRs are mentioned in the evaluation of the implementation of the Paris Declaration (Wood et al, 2008) as one of the mechanisms for mutual assessment of progress, donor's effectiveness in terms of providing financial and human inputs is not on the agenda of the JSRs studied, except in the case of Niger where donors' fulfilment of the agreements made in the 'Cadre Partenariat' is increasingly critically assessed. Reports of consultants are much more critical on the partnership between donors and government. For instance, a consultant of the Belgian Technical Cooperation wrote in his report on the 5th JSR in Burkina Faso: 'L'absence d'un vrai "partenariat" entre MEBA (Ministry of Health) et PTF (donors) était assez visible: les PTF posaient des questions critiques et le MEBA répondait; la présence de volonté de chercher des solutions ensemble n'était pas très apparente.' In his review of JSRs in the education sectors of Ethiopia, India and Rwanda, Packer (2006) similarly highlighted that the degree to which donor's accountability was tackled during the JSRs was not clear.

5. ISSUES FOR DISCUSSION

Confronting the research findings with insights from evaluation theory and practice raises a number of issues for broader discussion and from which we might learn to improve actual JSR practices on the ground. We have structured the discussion according to four major items:

- the balance between 'substance' and 'institutional/ systemic issues'
- the balance between 'accountability' and 'learning'
- the JSR as a forum for harmonisation and cooperation
- the JSR and feedback

5.1. Balancing 'substance' and 'systemic' issues

Our review highlights that so far JSRs overwhelmingly focus on 'substance' issues while they largely neglect the deeper-rooted institutional and systemic issues. The focus on performance 'stock' measurement is understandable given the existing short-term M&E needs. The disregard of systemic issues, however, runs counter to the logic of the new aid modalities. These endorse a shift from 'content' to 'process', hereby expecting that process conditionality, with in particular attention to reforms in the area of public expenditure accounting, budget reform and public service management, will lead to greater national ownership and commitment for poverty reduction policies and programmes (see Booth, 2003). The negligence of systemic issues neither matches the growing understanding of the importance of institutional capacity for a successful implementation of sector-wide approaches. Cassels (1997), for instance, has stressed the importance of government capacity to lead the process of sectoral development and to facilitate the introduction of common management arrangements (see Cassels, 1997).

While the need to focus on systemic issues is progressively recognised in the JSRs of particularly Niger and Burkina Faso, none of these explicitly integrates the elaboration, strengthening, follow-up of the underlying apparatus for policy-making, implementation, monitoring and evaluation among its objectives and targets. This is certainly not unique to the countries or sectors under study. While weaknesses in particularly monitoring and evaluation have been identified by Dfid as the most fundamental threat to the sustainability of the sector-wide approach (Grindle, page cited in World Bank, 2001: 27), the development of M&E systems generally remains weak in the development of sector programmes (Jones, cited in World Bank, 2001 : 27).

It entails a long-term approach^[1] with sound diagnosis of M&E demand and supply side, target setting, capacity building and follow-up demanding considerable time and financial resources (Feinstein and Ingram, 2003). Given that M&E system and capacity building is not a-stroke-of-the-pen undertaking, and taking into account the existing M&E accountability and learning needs, alignment with M&E systems will necessarily be incremental. As Foster (2000: 30) puts it, ‘the presentational risks from too rapid a reliance on government systems can be high: lack of capacity may lead to a dip in disbursements, or exposure of major fraud or mismanagement may lead to a reversal of progress’. Nevertheless, there is an important role for ‘path-finder’ donors willing to develop procedures and systems and then test them by daring to put money through them’ (Foster, 2000: 31). Refraining from investing in ‘systemic’ issues risks to boost parallel M&E processes and to undermine the M&E reform agenda.

JSRs that invest in the set-up of sector M&E systems and capacity-building would, over time, also benefit from higher-quality sector performance reports. These are among the main documentary inputs into the JSR and their quality largely determines the overall quality of the JSR. Investing in maturing sector M&E systems would gradually shift the essence of a JSR away from ‘substance’ assessment and learning to more systemic evaluation exercises and formative and summative meta-evaluations which focus on the quality of the sector’s M&E systems and their main outputs. In systemic evaluative exercises the functioning of the whole system for supplying reliable evaluative information is assessed (see Schartz and Mayne, 2005). A meta-evaluation is a method to analyse and assess the quality of evaluations (Bustelo, 2003) and/or evaluation systems. Two types of meta-evaluation can be distinguished. Formative meta-evaluation is pro-active and particularly helpful in improving ongoing evaluation work (Stufflebeam, 1974). It can be used to identify strengths and weaknesses in evaluation practice (including M&E systems) in order to develop evaluation capacity (Leeuw en Cooksy, 2004). On the other hand summative meta-evaluations are more retroactive and mainly serve accountability (Stufflebeam, 1974). Reasons for doing a summative meta-evaluation include informing stakeholders’ decisions about whether and how to use evaluation findings and helping evaluators’ and researchers’ to select evaluation findings for evaluation syntheses (Leeuw en Cooksy, 2004). In order to ensure objectivity and quality, meta-evaluations are usually performed on the basis of evaluation standards^[2].

A practice that might be particularly valuable for M&E institutional capacity building is ‘twinning’, whereby linkages are created between institutions with similar functions but with different levels of development. Contrary to more traditional technical assistance which is more based on individual capacity building, twinning focuses more on the development of organizations and institutions. Activities generally used in twinning arrangements are formal off-job training courses, in-house training courses, on-job training and study visits. In order to really fulfil the aim of institutional capacity building, partners should be carefully chosen and expectations on cooperation should be clear for all partners involved (Jones and Blunt, 1999).

[1] The ‘Developmental Approach’ of Schwartz and Mayne (2005, p. 321) might be interesting in this respect. It starts with the building blocks of standards, capacity building and education; continues with formative approaches and then friendly summative approaches and periodic systems checks, supplemented if necessary with less friendly summative approaches.

[2] An example of evaluation standards is the Programme Evaluation Standards of the Joint Committee on Standards for Educational Evaluation (1994), which groups thirty evaluation standards into four dimensions: utility, feasibility, propriety and accuracy. Stufflebeam (2001: 185) uses these four dimensions in his definition of meta-evaluation; “the process of delineating, obtaining, and applying descriptive information and judgemental information about the utility, feasibility, propriety and accuracy of an evaluation in order to guide the evaluation and to publicly report its strengths and weaknesses”.

As JSRs involve a broad range of actors, both from within and outside government, M&E capacity building should not only be confined to the traditional M&E supply side actors. When M&E capacity building becomes one of the explicit objectives of JSRs, M&E demand side actors outside government may also benefit. This will eventually contribute towards 'downward accountability' whereby national non-government actors hold their own government accountable.

5.2. Balancing 'accountability' and 'learning'

While it is not always explicitly stated in these terms, JSRs strive to satisfy an odd mixture of accountability and learning needs, mostly with an inclination towards accountability. There are contributions to evaluation literature indicating that it is preferable to unite both functions. Perrin (2002), for instance, suggests that a limited focus on accountability and performance targets is useless in providing future directions as it leads to manipulation of data. This has also been evidenced off in the case of JSRs where governments have tried to keep negative comments out of the reports, hereby undermining the confidence of donors in the process and reinforcing the tendency to stick to existing parallel M&E channels (see Brown et al., 2001). Instead, Perrin proposes an alternative framework for accountability in which programmes should not be held accountable for exactly executing what was planned, but for demonstrating that they are focused on outcomes and that they take a learning approach. Lehtonen (2005) indicates that there might be some ways to reconcile both functions, but he nevertheless concludes that one primary function and main client should be identified for each evaluation exercise.

Generally, standard evaluation theory and practice suggest that exercises that blur the distinction between accountability and learning run the risk of falling short in both areas, because more often than not different functions necessitate the involvement of different actors, a different focus and the use of different methodologies. Although there is some overlap in audience, the primary audience for accountability are media and civil society organisations in donor and recipient countries, parliament and national audit offices in donor countries, senior management within the international agency and the partner government. On the other hand, the primary audience for learning are government departments, development agencies and senior management, policy units and operational departments within the international agency in donor countries and in-country staff within the international agency, operational staff in other development agencies, counterparts involved in implementing programmes and policy makers in government and other in-country organisations in recipient countries (OECD/DAC, 2001). A different audience needs a different approach, with more emphasis on rigour, independence, replicability and efficiency in the case of accountability and more focus on the process in the case of learning, whereby space is needed for discussion of experience and lessons learned (OECD/DAC, 2001).

Additionally, accountability exercises demand careful scientific preparation. In order to be able to extrapolate findings from a limited sample, for instance, one needs to increase as much as possible the external validity of the assessment. Concretely, this implies that samples need to be representative. Failing to respect this weakens the credibility of the assessment and gives the implementing agency under scrutiny every excuse to downplay or flatly ignore 'negative' findings. On the other hand, 'learning' necessitates zooming into and analyzing unexpected results, success stories, failures or specific experiments, for instance in the way services are delivered. Here external validity is obviously not a concern, on the contrary; internal validity

should be optimized, so as to make credible conclusions regarding causal linkages of the underlying programme theory. In order to identify issues under study and to ensure that feedback for future policy-making and management is optimal, one obviously needs to involve those responsible for policy and programme design and implementation. So far, the analytical quality of the JSRs is particularly low, which is obviously related to the above-mentioned lack of sector evaluative capacity. That the usefulness of these ‘learning’ exercises potentially extends beyond the particular programme, sector, or even country setting, even though the costs of such ‘evaluative exercises’ are substantial, justifies heavy donor (financial) investment in the building of recipient evaluative capacity and systems. The shift towards more joint ‘evaluative’ exercises in the context of sector and general budget support is particularly promising in this respect as it might lower the public goods problem that also exists among donor agencies when it comes to impact evaluation (Centre for Global Development, 2006).

Keeping in mind the tensions that might exist among ‘accountability’ and ‘learning’, it might be worth splitting JSRs into two separate exercises or adopt a kind of hybrid multi-partner approach^[1] where not all stakeholders participate similarly in all dimensions and phases of the joint review exercise. If one wants to do a reality check on government’s performance, for instance, then obviously there is a need to respect as much as possible the principle of autonomy and independence which implies the involvement of independent actors on the demand side, such as the auditor general’s office, independent policy research institutes and independent CSOs. One could even argue that donors (or at least some departments in donor agencies) are themselves not independent enough to be involved in these accountability exercises^[2]. The involvement of national independent actors in ‘accountability’ exercises may at the same time also function as a kind of learning-by-doing approach to build national evaluative capacity.

Those in favour of one exercise to cover both accountability and learning needs consider the use of one system as a way to tackle the limitations of the current move towards narrowly confined performance measurement (PM) systems to satisfy accountability needs. Criticism is related to different issues such as the organisational and systemic side-effects including organisational paralysis, ‘tunnel’ vision, ‘goal displacement’, the construct validity of performance indicators, (see Ebrahim, 2005; Lehtonen, 2005; Perrin, 1998), and the difficulties in identifying the appropriate level of performance measurement. There is a vivid discussion among academics and aid practitioners regarding the appropriate identification of indicators for accountability^[3] (see Adam and Gunning, 2002; Adam et al., 2004; Foster, 2000). Most of the critics consider PM valuable as a kind of ‘early warning system’, which indicates whether output has indeed resulted in the intended outcome/impact (Booth and Lucas, 2002) but PM indicators should be distinguished from indicators which are used for aid allocation (Adam and Gunning, 2002), because PM does not explain the causes behind the observed outcomes and outputs (Lehtonen, 2005). PM is by preference integrated within a broader evaluation framework that also

[1] OECD/DAC (2005b) identifies ‘hybrid multi-partner’ as one of the three categories of joint evaluations. It regroups a wide range of more complex ways of joint working whereby not all stakeholders are similarly involved in all phases or dimensions of joint M&E exercises.

[2] In a revealing interview, an auditor general confided that in his view donors did not really support him in his goal to undertake independent performance auditing, because they feared that what might come out would not please them.

[3] Some argue that upward accountability which influences aid disbursements should only be based on real impact indicators as a focus on inputs and outputs is ineffective and counter to the ownership principle. Others rather defend the idea of accountability at the output level as this reflects the level that is under control of the government and less dependent upon exogenous factors.

takes on board learning needs and includes an analysis of the reasons for achievement or non-achievement of the stocks. Notwithstanding the value of the arguments, it remains important to acknowledge the inherent tensions when trying to reconcile both objectives of ‘accountability’ and ‘learning’

5.3. JSR as a forum for coordination and harmonisation

The 1998 Review of OECD/DAC Principles for Evaluation of Development Assistance pointed out that the principles in least compliance were collaboration among donors and the participation of recipients. Our review of AMs and our own JSR participation suggest that JSRs have the potential to turn the tide. They gather a broad range of stakeholders, including various government representatives, both from the sector and the finance ministry, national non-government actors, and budget and non-budget support donors.

For budget support donors the rationale for participation in JSRs is obvious: as they shift from isolated projects to sector-wide support, it is no longer possible to attribute results directly to their inputs. For ‘project’ donors, the participation in JSRs is interesting as they gain insights into the sector they work in. Moreover, through their participation in the JSRs they get familiarised with this new M&E instrument, which is an important first step in the transition from individual M&E to joint M&E. For budget support donors the participation of project donors and non-government actors is particularly interesting as they bring in information from the ‘ground’. Characteristic for budget support donors is that they often no longer operate at the local level, which cuts them off from local knowledge and realities. Whereas some JSRs also include field visits, in many cases JSRs take entirely place in the capital as visits to the field are often logistically complex, time-consuming and costly (see Packer, 2007). Information brought in by project donors and non-government actors may include e.g. participatory service delivery surveys, tracking studies, in-depth evaluations of pilot interventions, information on specific interventions in particular areas or target groups^[1]. This does not only allow triangulation of information, it also functions as a ‘reality check’ on the information that is provided through conventional management information system (MIS) and household surveys (Packer, 2007; OECD/DAC, 2000). The use of information from non-government actors in the JSR process may also stimulate their own data collection activities and its quality (Ginsberg and Pane, 2005: 230).

Improving the soundness of the JSRs is essential as a means to convince all involved to gradually break down their stand-alone M&E systems or at least to design their M&E exercises in a way which enriches the overall JSR. Continuous parallel stand-alone missions (including from budget support donors) add to the cost of ‘joint’ M&E processes, which are particularly high in the initial stages (see e.g. Martinez 2006). To decrease the costs of joint review processes, some (see e.g. Cassels, 1997) have suggested working with smaller missions, and/or experimenting with donor representation using other donors (‘silent partnerships’). Whereas all actors need not collectively participate in all different phases and dimensions of the JSRs, it is not certain that transaction costs for the recipient will be lessened when fewer actors participate in the ‘joint’ processes. Given their own accountability needs to their constituencies, it is likely that ‘silent partners’ may somehow be allured to perform a kind of additional reality check.

[1] It might be interesting to perform a kind of meta-evaluation to check the quality of the different evaluative information (see Stufflebeam, 1974; Schwartz and Mayne, 2005).

5.4. JSR and feedback

Our own review shows that there has been little systematic follow-up on JSR recommendations. This lends support to some other studies (Martinez, 2006) which have raised similar concerns and indicated a failure to translate JSR conclusions and recommendations into policy-making, budgeting and implementation plans. This finding is indicative of a deficiency in the underlying institutional apparatus to address these systemic issues in the JSR.

The organisational set-up of the JSR has, however, a number of attractive features to stimulate dissemination and influence of its findings. First, most JSRs are linked to an (A)RM where the conclusions and recommendations of the different working groups are brought together, discussed and adopted. As all stakeholders at the JSR participate in this forum, it is often labelled as painfully laborious and repetitive, involving too many people. Nevertheless, it potentially has a pay-off in terms of the dissemination of information and the influence^[1] the M&E exercise can have. A joint review implies debate and dialogue which might in particular increase the process use of the evaluation^[2]. It may create a shared understanding and might lead to changes in individual and collective thinking and behaviour which might eventually induce changes in outcomes. It should be noted however, that quality of the input into the JSR, generally sector performance reports, should be satisfactory as to provide an incentive for debate (Lehtonen, 2005). What is less clear is the degree of dissemination of the findings beyond those that directly participate at the JSR, towards the wider audience each of those participants represents.

Second, JSRs are not generally stand-alone, one-shot events but part of a continuous M&E process. They are based upon data collection that takes place throughout the year, including information from MIS, non-government actors, additional commissioned studies on specific topics, etc.. At the same time the JSR feeds into this continuous process by indicating issues for follow-up and capacity-building, by indicating on which specific topics additional analyses or data collection is needed. The quality of the JSR depends to a certain extent upon the quality of the continuous M&E process and a high quality JSR also improves the continuous M&E process. A continuous M&E process also requires regular dialogue which in turn necessitates the presence of strong sector and M&E working groups (see also Packer, 2006; Cassels, 1997).

Third, in many cases the JSR adopts an 'expanded' approach to M&E. An evaluability assessment^[3] of many sector programmes would indeed indicate that evaluative exercises should not yet be adopted as there is often no clear-cut sector programme with sound indicators. However, the role of the review exercise is in many cases also to feed into the elaboration or at least improvement of the sector policy, strategies, plans and indicators. The insights from evaluability assessments are nevertheless interesting because they could help set an agenda of issues that, if not covered before the JSR starts, could at least be on its agenda.

[1] We deliberately use the notion of 'influence' instead of 'use'. See on this Henry and Mark (2003), Kirkhart (2000).

[2] See Lehtonen (2005), Patton (1998) on 'process use'.

[3] An evaluability Assessment is a systematic process that helps identify whether program evaluation is justified, feasible, and likely to provide useful information (see Wholey, 1994).

6. CONCLUSIONS

The 2005 Paris Declaration (PD) confronted both donors and recipients with an ambitious reform agenda. Recent follow-up surveys have revealed that M&E is one of the areas where implementation is lagging seriously behind. In sharp contrast to their own declarations, donors hardly harmonise and align with recipient systems or arrangements that are often non-existent. This very reluctance however blocks at the same time the set-up and maturing of recipient systems. To move beyond the current chicken-and-egg dilemma they are caught in, donors and recipients need to seek a middle way between the satisfaction of existing M&E needs in the short run and the building of recipient M&E frameworks in the longer run.

This paper focuses on Joint Sector Reviews as potentially one interesting type of M&E exercises that might fill the gap. While JSRs are increasingly used on the ground and considered by practitioners as vital components of sector programmes, they have so far remained under-researched. Based on a combination of desk and field study, this paper analyses the extent to which JSRs in the education sectors of Burkina Faso, Niger and Mali effectively reconcile objectives of accountability and learning, while at the same time feed into the M&E reform agenda. Our stocktaking leads to a set of nuanced conclusions. The JSRs under study prioritise accountability over learning needs and mainly focus on implementation. When analysis is included, it is mostly of low quality and mainly dependent upon the analytical quality of the sector performance report, the main documentary input of the leading ministry. The lenient quality of these reports hints at the same time at the deficient quality of the underlying recipient M&E system or arrangements. In spite of these observed weaknesses, JSRs predominantly focus on sector 'content' issues to the detriment of institutional and systemic issues. While some JSRs underscore the importance of the underlying apparatus, none really engages in diagnostic assessment of the existing M&E system or arrangements. No targets are set regarding areas for improvement, there is no follow-up of recommendations nor is there a straightforward capacity building plan. Unsurprisingly, there is little maturing of the system and little effective donor alignment.

On the positive side, JSRs score strongly on coordination, harmonisation, leadership and broad-based participation, M&E principles that traditionally prove difficult to realise. JSRs are led by the sector ministry (alone or with the lead donor), bring together a broad variety of actors, make considerable efforts to coordinate and harmonise at the sector level and between the sector and the central level. While cross-reading among various data sources has so far remained underutilised, JSRs create a forum for information triangulation. This may improve validity of the exercises and enhance trust, which is particularly vital for those donors who are conventionally reluctant to align. Furthermore, JSRs also have a number of attractive organisational features to stimulate feedback: they are linked to an (annual) review meeting where conclusions and recommendations are discussed and negotiated, and they are generally integrated into a continuous process of follow-up and negotiation through sector working groups.

In summation, JSRs have the potential to function as M&E exercises that reconcile short and longer-term objectives. To realize this, they particularly need to address, in the short run, longer-term systemic and institutional issues. While this may involve considerable addi-

tional investment, it should lead to less laborious and complementary M&E exercises in the long run. Failure to do this risks institutionalising JSRs as parallel M&E exercises that hinder the further elaboration and maturing of the recipient sector M&E system. Instead they should be ‘chameleon-like’, modifying their outlook to accommodate the changing status of the recipient’s M&E systems.

REFERENCES

- 3rd High Level Forum on Aid Effectiveness (2008) *Accra Agenda for Action. 3rd High Level Forum on Aid Effectiveness*: Accra.
- Adam, C., Gunning, J.W. (2002) "Redesigning the aid contract: Donors' Use of Performance Indicators in Uganda". *World Development* 30 (12): 2045-2056.
- Adam, C., Chambas, G., Guillaumont, P., Jeanneney, S.G., Gunning, J.W. (2004). "Performance-based Conditionality: A European Perspective". *World Development* 32 (6): 1059-1070.
- American Joint Committee on Standards for Educational Evaluation (1994) *Program Evaluation Standards*. Sage Publications.
- Booth, D. and Lucas, H. (2002) "Good Practice in the Development of PRSP Indicators and Monitoring Systems". *ODI Working Paper* 172. ODI: London.
- Booth, D. (2003) "Introduction and Overview". *Development Policy Review* 21 (2): 131-159.
- Brown, A., Foster, M., Norton, A., Naschold, F. (2001) "The Status of Sector Wide Approaches". *ODI Working Paper* 142. ODI: London.
- Bustelo, M. (2003) "Evaluation of Gender Mainstreaming: Ideas from a Meta-Evaluation Study". *Evaluation* 2003 (9): 383. Sage Publications
- Cassels, A. (1997) *A guide to sector-wide approaches for health development*. WHO: Geneva
- Center for Global Development (2006) "When will we ever learn? Improving Lives through Impact Evaluation". *Report of the Evaluation Gap Working Group*. Center for Global Development: Washington, D.C.
- Cracknell, B.E. (2000) *Evaluating Development Aid: Issues, Problems and Solutions*. Sage Publications: New Delhi, London.
- Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ)/German Federal Ministry for Economic Cooperation and Development (BMZ) (2004) *National Monitoring of Strategies for Sustainable Poverty Reduction/ PRSPs (vol. 1)*. GTZ, Governance and Democracy Division, Mainstreaming Poverty Reduction Project: Eshborn.
- Ebrahim, A. (2005) "Accountability Myopia: Losing Sight of Organizational Learning". *Nonprofit and Voluntary Sector Quarterly* 2005; 34: 56.
- Feinstein, O., Ingram, G. (2003) *Lessons learnt from World Bank experiences in joint evaluation* (room document submitted to the OECD/DAC Evaluation Network). OECD/DAC: Paris.
- Foster, M. (2000) "New Approaches to Development Co-operation: what can we learn from experience with implementing Sector Wide Approaches?" *ODI Working Paper* 140. ODI: London.
- Ginsberg, A.L., Pane, N. (2005) "Decentralisation does not Mean Poor Data Quality : A Case Study from the U.S. Department of Education." In: *Quality Matters: Seeking Confidence in Evaluation, Performance Auditing and Performance Reporting*, Schwartz R, Mayne J (eds). Transaction Publishers: New Jersey.
- Grindle, J. (1999) "Donor Harmonization in Sector Programmes: Issues and Challenges", *Key-note Paper for PROAGRI High Level Meeting*, Dublin, March 29.
- Henry, G.T., Mark, M.M. (2003) "Beyond Use : Understanding Evaluation's Influence on Attitudes and Actions". *American Journal of Evaluation* 24 (3) : 293-341.

IDD and Associates (2006) Evaluation of General Budget Support: Synthesis Report. University of Birmingham.

Joint Committee on Standards for Educational Evaluation (1994) *Programme Evaluation Standards*

Jones, M.L., Blunt, P. (1999) " 'Twinning' as a method of sustainable institutional capacity building", *Public Administration and Development* 19 (4): 381-402.

Jones, S.P. (1997) "Sector Investment Programs in Africa : Issues and Experience", *World Bank Technical Paper* No. 374. World Bank: Washington, D.C.

Kirkhart, K. (2000) "Reconceptualising Evaluation Use. An integrated theory of influence". In: *New directions for Evaluation* 88. Jossey-Bass: San Francisco; 5-23.

Leeuw, F., Cooksy, L. (2004) "Evaluating the Performance of Development Agencies: The Role of Meta-Evaluations". In: G.Keith Pitman et al (eds), *Evaluating development effectiveness, World Bank Series on Evaluation and Development*, Vol 7, Transaction Publishers, Rutgers, 2004: 95-108

Lehtonen, M. (2005) "OECD Environmental Performance Review. Accountability (f)or Learning". *Evaluation* 11 (2): 169-188.

Martinez, J. (2006) "Implementing a sector wide approach in health : the case of Mozambique", *HSLP Institute Technical Paper*. HSLP Institute: London.

Ministère de l'Economie et des Finances, Ministère de l'Enseignement de Base et de l'Alphabétisation (2007) *Cadre Partenarial entre le gouvernement du Burkina Faso et ses Partenaires Technique et Financiers (P.T.F) du Secteur de L'Education de Base.*

Ministère de l'Education de Base et de l'Alphabétisation de Niger. 2003, 2004, 2005, 2006, 2007. *Aide-Mémoires des Missions Conjointes du Programme Décennal de Développement de l'Education* (PDDE)(2003, 2004, 2005, 2006, 2007).

Ministère de l'Education National de Mali. 2002, 2003, 2004, 2007. *Aide-mémoires des Missions Conjointes de Suivi du PISE* (2002, January 2003, June 2003, 2004, 2007).

Ministère de l'Enseignement de Base et de l'Alphabétisation de Burkina Faso. 2003, 2004, 2005, 2006, 2007. *Aide-mémoires des Missions Conjointes de Suivi du PDDEB* (2003, 2004, May/ June 2005, December 2005, April/May 2006, November 2006, April 2007, November 2007).

OECD/DAC (1991) *DAC principles for evaluation of development assistance*. OECD/DAC: Paris.

OECD/DAC (1998) *Review of the DAC principles for evaluation of development assistance*. OECD/DAC: Paris.

OECD/DAC (2000) *Effective practice in conducting a joint multi-donor evaluation*. Paris: OECD/DAC.

OECD/DAC (2001) *Evaluation Feedback for Effective Learning and Accountability*. Paris: OECD/DAC.

OECD/DAC (2002) *Glossary of Key Terms in Evaluation and Results Based Management*. OECD/DAC: Paris.

OECD/DAC (2005a) *Paris Declaration on Aid Effectiveness*. OECD/DAC: Paris.

OECD/DAC (2005b) *Joint evaluations: recent experiences, lessons learned and options for the future*. OECD/ DAC: Paris.

- OECD/DAC (2005c) *Survey on Harmonisation and Alignment: Measuring Aid Harmonisation and Alignment in 14 Partner Countries*. OECD/DAC: Paris.
- OECD/DAC (2006) *Guidance for Managing Joint Evaluations. DAC Evaluation Series*. OECD/DAC: Paris.
- OECD/DAC (2007) *2006 Survey on Monitoring The Paris Declaration*. OECD/DAC: Paris.
- OECD/DAC (2008) *2008 Survey on Monitoring The Paris Declaration. Effective Aid by 2010 ? What it will take*. OECD/DAC: Paris.
- Packer, S. (2006) *Joint monitoring review missions in the education sector*. UNESCO : Paris.
- Patton, M.Q. (1998) "Discovering Process Use". *Evaluation* 4 (2) : 225-233.
- Perrin, B. (1998) "Effective Use and Misuse of Performance Measurement". In: *American Journal of Evaluation*, vol 19, no 3, 367-379 (1998)
- Perrin, B. (2002) *Towards a New View of Accountability*. Paper presented to the European Evaluation Society annual conference, Seville, Spain.
- Schwartz, R., Mayne, J. (2005) "Assuring the quality of evaluative information : theory and practice". *Evaluation and Program Planning* 28 (1) : 1-14.
- Shadish, W., Cook, T., Campbell, D. (2002) *Experimental and Quasi-experimental Design for Generalized Causal Inference*. Houghton Mifflin Company: Boston, New York
- Stufflebeam, D.L. (1974) "Meta-evaluation", *Occasional Paper Series* No. 3. Western Michigan University Evaluation Center : Kalamazoo MI.
- Stufflebeam, D.L. (2001) "The Metaevaluation Imperative". *American Journal of Evaluation* 2001; 22; 183.
- UNDP (2007) *Human Development Report 2007/2008. Fighting climate change: Human solidarity in a divided world*. UNDP: New York.
- UNESCO (2008) *EFA Global Monitoring Report, Education for All by 2015. Will we make it?* Oxford University Press: Oxford.
- Wholey, J.S. (1994) "Evaluability Assessment". In: Wholey, J.S., Hatry, H.P., Newcomer KE (eds.), *Handbook of practical program evaluation*, San Francisco : Jossey-Bass; 33-61.
- Wood, B., Kabell, D., Muwanga, N., Sagasti, F. (2008) *Synthesis Report on the First Phase of the Evaluation of the Implementation of the Paris Declaration*. Ministry of Foreign Affairs of Denmark: Copenhagen
- Working Party on Aid Effectiveness (2008) *Aid Effectiveness. A Progress Report on Implementing the Paris Declaration*
- World Bank (2001) *Education and Health in Sub-Saharan Africa. A Review of Sector-Wide Approaches*. World Bank : Washington, D.C.
- World Bank (2003) *Toward Country-led Development. A Multi-Partner Evaluation of the Comprehensive Development Framework. Synthesis Report*. World Bank : Washington, D.C.
- World Bank (2005) *Enabling Capacity to Achieve Results – 2005 Comprehensive Development Framework (CDF) Progress Report*. World Bank, Operations Policy and Country Services : Washington, D.C.
- World Bank (2007) *Results-Based National Development Strategies. Assessment and Challenges Ahead*. World Bank : Washington, D.C.



University
of Antwerp



INSTITUTE OF DEVELOPMENT
POLICY AND MANAGEMENT