

Evaluation Capacity Development through Joint Sector Reviews

Please do not distribute or copy without prior authorization of the authors.

Liesbeth Inberg and Nathalie Holvoet

Institute of Development Policy and Management, University of Antwerp, Belgium

One of the crucial reform areas agreed to in the Paris Declaration (PD) and reaffirmed in the Accra Agenda for Action (AAA) relates to monitoring and evaluation (M&E). In short, partner countries are asked to elaborate results-oriented reporting and assessment frameworks, while donors are expected to harmonise and align their policies and frameworks. Both commit to collaborate in order to strengthen country capacities and demand for results-based management. By 2010, the proportion of countries without transparent and monitorable performance assessment frameworks should be reduced by one-third. In 2007, only 7.5% of IDA-eligible countries reviewed in the 2008 PD progress report had largely developed results-oriented frameworks. Donors are reluctant to rely on M&E systems that are not yet able to satisfy the current M&E needs of accountability and learning, yet at the same time M&E systems of partner countries cannot be developed with donors persistently using their own parallel systems. At sector level, Joint Sector Reviews (JSRs), which are increasingly included as essential ingredients in Sector Wide Approaches, have the potential to take the M&E reform agenda forward. We argue that while JSRs are still overwhelmingly focused on sector substance/content issues, M&E capacity development should become one of their explicit objectives. This paper will analyse how JSRs can contribute to M&E capacity development at sector level making use of the authors' review of a sample of JSRs from Burkina Faso, Mali and Niger. In doing this, we will draw upon insights from evaluation and capacity development theory and practice.

1. Changing aid architecture, changing monitoring and evaluation

As a reaction to development aid's disappointing results, donors and partner countries committed themselves in the 2005 Paris Declaration (PD) to the principles of ownership, harmonisation, alignment, managing for results and mutual accountability in order to improve its effectiveness. These commitments are reaffirmed in the Accra Agenda for Action (AAA) signed in September 2008. In the PD, the need for sufficient capacity in partner countries is acknowledged, "the capacity to plan, manage, implement, and account for results of policies and programmes, is crucial for achieving development objectives – from analysis and dialogue through implementation, monitoring and evaluation. Capacity development (CD) is the responsibility of partner countries with donors playing a support role" (OECD/DAC, 2005: 5). The related indicator (indicator 4) is linked to the role of donors: donors commit themselves to, "align their analytic and financial support with partners' capacity development objectives and strategies, make effective use of existing capacities and harmonise support for capacity development accordingly" (OECD/DAC, 2005: 5). As far as the principle of managing for results is concerned, donors and partner countries jointly commit to strengthen country capacities and the demand for results-based management. The indicator to measure progress in the area of managing for results (indicator 11) relates to the commitment of partner countries 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" (OECD/DAC, 2005: 8). By 2010, the proportion of countries without transparent and monitorable performance assessment frameworks should be reduced by one-third (OECD/DAC, 2005: 10). In 2007, however, only 7.5% of the IDA-eligible

countries¹ reviewed in both the 2006 and 2008 PD progress reports had largely developed results-oriented frameworks in place (5% in 2005). None of the countries had a results-oriented framework which was judged to be sustainable (OECD/DAC, 2008). The progress report on the implementation of the Paris Declaration (Working Party on Aid Effectiveness, 2008) highlights the lack of country capacity for managing for development results (i.e. capacity to plan, budget, manage, implement and account for policies and programmes) as one of the major challenges in progressing on this principle. With the acknowledgement in the AAA that accounting for the achievement of development results is one of the three major challenges² to be addressed to accelerate progress on aid effectiveness (3rd High Level Forum on Aid Effectiveness, 2008), the need for progress in achieving indicator 11 has been reconfirmed.

One of the inherent problems of the M&E reform agenda is that donors and partner countries are caught in a kind of chicken-and-egg dilemma. Because the M&E systems of partner countries are not yet able to satisfy current donor M&E needs of accountability and learning, donors, who committed themselves to align to the performance assessment frameworks of partner countries, are reluctant to rely on those systems. On the other hand, M&E systems of partner countries cannot be developed with donors persistently using their own parallel systems. A possible way to break through this dilemma is to adopt a two-track approach that combines a focus on the development of the recipient's M&E system with the fulfilment of the existing M&E needs in the short and middle term. Since, in most cases, the current recipient M&E system is not yet able to satisfy present donor 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 existing recipient and donor M&E demands, is crucial to convince donors to lower their own stand-alone M&E exercises. At the same time, complementary recipient 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 the development of their own M&E system. The maturing recipient M&E system can then progressively accommodate donor M&E needs with complementary exercises (thereby evolving gradually towards a kind of meta-evaluation).

So far, there are no clear-cut recipes provided by M&E theory or longstanding M&E practices for the design of such complementary M&E exercises. At best, evaluation literature and practice inform us that the stakes involved are huge. Nevertheless and by necessity, various stakeholders on the ground seek solutions to the challenges. One such strand of ongoing M&E exercises are Joint Sector Reviews (JSRs). JSRs are increasingly included as essential in Sector Wide Approaches (SWAs) and in processes linked to sectoral and general budget support (the conclusions of JSRs often influence disbursement decisions). JSRs are gradually coming to function as the main instrument for assessing progress, resolving issues and reaching agreement on sector policy and programme targets, and are increasingly replacing individual donor missions (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 (i.e. taking stock of experiences, analysing strengths and weaknesses and formulating suggestions for improvement -see Packer, 2006). To help fill this gap, we assessed a sample of 18 JSRs in the education sectors of Burkina Faso, Mali and Niger. We reviewed the main written output of the JSRs, the Aide Mémoires, according to their fulfilment of the M&E needs of accountability and learning, and their attention to the principles of the Paris

¹ Mozambique, Tanzania and Uganda.

² The other two challenges are country ownership and building more effective and inclusive partnerships.

Declaration. Besides this, we participated in two JSR missions³ which provided us with opportunities to get better acquainted with the processes involved and interact with several of the participating stakeholders. We confronted the research findings with insights from evaluation theory and practice to provide some suggestions for JSR improvement (see Holvoet and Inberg, 2008). This paper builds on the findings of that research by analysing how JSRs can contribute to evaluation capacity development (ECD) at the sector level. In doing this, we draw upon insights from evaluation and capacity development theory and practice.

The structure of this article is as follows: section two provides some background on evaluation capacity development and joint sector reviews, section three summarizes present ECD practice in the context of JSRs, section four identifies how JSRs could stimulate ECD, and finally, section five concludes.

2. Capacity Development and Joint Sector Reviews: insights from literature

2.1 Evaluation Capacity Development (ECD)

Within development cooperation, capacity development (CD)⁴ is traditionally considered a human resource issue, with technical assistance (TA) and training as the main external interventions. However, a limited focus on human capacity building does not automatically contribute to overall capacity (Baser and Morgan, 2008) as TA, in particular, is commonly used to temporarily fill gaps in available technical capacity. Other criticisms of TA include: it is often costly and supply driven, the lack of TA coordination among donors and its use as a free good by recipient countries (Wubneh, 2003).

Increasingly the need for a broader view on capacity is acknowledged. The World Bank (2005) mentions three dimensions of capacity which should be addressed simultaneously when supporting CD: human capacity; organisational capacity and institutional capacity. *Human capacity* refers to the skills of individuals to analyze development needs, to design and implement strategies, policies and programmes, to deliver services and to monitor results. *Organisational capacity* is the capacity of a group of individuals connected by a shared purpose with clear objectives and by internal structures, processes, systems, staffing and other resources to achieve them. And finally, *institutional capacity* refers to the formal and informal norms providing the framework of goals and incentives within which organisations and people operate (World Bank, 2005).⁵ Due to wide disparities in capacities among partner countries, there is no blueprint for CD. As emphasized by Lundgren and Kennedy (2009: 84), “strategic, locally developed, carefully tailored interventions are more likely to have a significant, sustainable impact.”

ECDPM defines capacity as, “that emergent combination of individual competencies and collective capabilities that enable a human system to create value” (Baser and Morgan, 2008: 34). Five core capabilities are distinguished: the capability to carry out technical, service

³ The 9th JSR in the education sector of Burkina Faso (10-16 April 2007) and the 3rd JSR in the health sector in November 2008 (in the context of another project).

⁴ The term ‘capacity development’ is increasingly preferred to ‘capacity building’ because the notion of ‘building’ presumes a process that begins with a plain surface and involves a step-by-step creation of a new structure, which does not generally correspond to the reality (OECD/DAC, 2006b).

⁵ Institutional capacity is also referred to as an ‘enabling environment’, see for example OECD/DAC (2006b).

delivery and logistical tasks; to commit and engage; to relate and attract resources and support; to adapt and self-renew and to balance coherence and diversity. The first core capability refers to the most widespread way of looking at capacity and emphasises the functional and instrumental ways of meeting objectives. In order to achieve sustainable capacity, this capability should be supplemented and combined with the other four capabilities of which the capability to commit and engage is the most crucial. This core capability relates to the importance of ownership; actors without this capability usually lack commitment or political will (Baser and Morgan, 2008). As with all development interventions, ownership is considered a prerequisite for sustainable capacity development (Baser and Morgan, 2008; EuropeAid, 2005; OECD/DAC, 2006).⁶

EuropeAid (2005) considers an organization's outputs good proxies for capacity. This aid agency accordingly proposes to focus on outputs as the entry point for CD instead of inputs such as TA and training. EuropeAid defines capacity as, "the ability to perform tasks and produce outputs, to define and solve problems, and make informed choices" (EuropeAid, 2005: 6), and capacity development as, "the process by which people and organizations create and strengthen their capacity over time" (EuropeAid, 2005: 6). The focus on outputs is the second step in the open system approach used by EuropeAid to assess capacity. The first step is to identify a vantage point for assessment, the third is to look at the context (i.e. structural and institutional factors and organizational and individual actors), the fourth is to consider the inputs or resources, and the last is to look inside organizations and networks (EuropeAid, 2005).

Even though these general definitions of CD relate to ECD as well, a specific description of ECD has been formulated by the World Bank Operations Evaluation Department (OED)/Independent Evaluation Group (IEG). It corresponds to the broader view of capacity development, "Evaluation Capacity Development is concerned with strengthening or building M&E systems in borrower countries in order that a situation is attained where M&E is regularly conducted and utilized by the countries themselves - governments and civil society. In other words, M&E are tools and ECD is the process of strengthening country-based systems to conduct and use M&E" (OED, 2004: 1).

Five key characteristics of ECD are distinguished by the OED (2002: 34):

- ECD is concerned with the utilization of M&E by developing countries;
- ECD is closely linked with other public sector reform efforts, particularly those relating to incentives, that determine the extent to which M&E will be utilized;
- ECD is focused on the measurement of public sector performance;
- ECD is concerned with a sustainable and systemic approach to M&E conduct and utilization; and
- ECD involves a purposeful, guided process of capacity development.

The first characteristic, the utilization of (or the demand for) M&E, is frequently mentioned as a prerequisite for the institutionalization of an M&E system (World Bank, 2002; OED, 2002; Mackay, 1999 and 2007) and a necessary condition for the effective management of public expenditures for poverty reduction (Mackay, 2007). The demand side of M&E information is

⁶ The core capability "to relate and to attract resources and support" includes the capabilities to: earn credibility and legitimacy; buffer the organization or system from intrusion; earn the trust of others, such as donors and clients; and to combine political neutrality and assertive advocacy. The core capability to adapt and self-renew includes the capabilities to: improve individual and organizational learning; foster internal dialogue; reposition and reconfigure the organization; incorporate new ideas; and to map out a growth path. Finally, the core capability to balance diversity and coherence includes the capabilities to: communicate; build connections; manage diversity; and to manage paradox and tension (ECDPM, 2008).

even more important than the supply side because, “if demand for M&E is strong, then improving supply in response can be relatively straightforward, but the converse does not hold” (Mackay, 2007: 54).

The existence of incentives to use M&E data, the second characteristic of ECD, is crucial for an M&E system to be successfully institutionalized. Mackay (2007) distinguishes three types of incentives to make it work: *carrots*, which provide encouragement and rewards for M&E and utilizing the findings; *sticks*, which give punishments to ministries or individual civil servants who do not take performance and M&E seriously; and *sermons* which involve high-level statements of endorsement and advocacy with regard to the importance of M&E.⁷ In their research on the institutionalization of M&E in Uganda, Booth and Nsabagasani (2005) found that a narrow focus on outcomes and impact lead individuals to disregard data because, “feedback about outcomes and impact is, in the best circumstances, very easily dismissed as somebody else’s responsibility, or no one’s. When general incentives improve, policy makers will be much more likely to make serious use of data about things that move quickly and over which they have a reasonable degree of control and/ or which are significant in an intended chain of events that they themselves have adopted as strategy” (Booth and Nsabagasani, 2005: 28). This also relates to the third characteristic of ECD, namely that ECD should be focused on the measurement of public sector performance and on the relationship between public actions and poverty reduction (OED, 2002).

The last two characteristics of ECD endorse the need for an M&E system’s gradual development based upon existing capacity. It involves the immediate implementation of a few mechanisms instead of the instant creation of a comprehensive or very complicated system (World Bank, 2002). In this sense, ECD entails a long-term approach with a sound diagnosis of: the existing M&E demand and supply side; target setting; capacity building and follow-up, and that demands considerable time and financial resources (Feinstein and Ingram, 2003). Given that ECD is not a stroke-of-the-pen undertaking, and taking into account existing M&E accountability and learning needs, alignment to 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 to test them with (real) cash contributions (Foster, 2000).

Whereas none of the highly aid-dependent countries has at this stage a sustainable M&E system (OECD/DAC, 2008; World Bank, 2007), most countries already have a number of M&E activities and systems in place. The problem, however, is often a lack of coordination and harmonization between them (Bedi et al., 2006; Mackay, 2007). To detect the different M&E systems and the existing capacity, and to acquire a shared understanding of the strengths and weaknesses, it is by now recognized that ECD activities should preferably begin with a sound diagnosis of the M&E demand and supply side (see Mackay, 2007; Khan, 2003; Kusek and Rist, 2004; World Bank, 2001; World Bank, 2002; OED, 2002). To avoid parallel donor diagnostic ‘bombardments,’ a coordinated multi-stakeholder approach, similar to the Public Expenditure and Financial Accountability (PEFA)-initiative⁸ should, in principle, be favoured.⁹

⁷ See Mackay (2007: 63, 64) for examples of these three types of incentives.

⁸ The PEFA programme aims at strengthening recipient and donor ability to: (i) diagnose the quality of country public expenditure, procurement and financial accountability systems; and (ii) develop a practical sequence of reform and capacity-building actions. See www.pefa.org for more information.

⁹ It is worth emphasizing the obvious: a shared diagnosis need not necessarily trigger identical donor responses.

Thus far, a harmonised M&E diagnostic instrument does not exist. There are, however, some interesting independent and donor-led assessments and studies that may provide inspiration for one's eventual design. Examples include the ECD diagnostic guide and action framework (Mackay, 1999) and the highly similar readiness assessment (Kusek and Rist, 2002), the diagnostic instrument elaborated in Bedi et al. (2006), the checklist used by Booth and Lucas (2002) in their diagnosis of PRSP M&E in 21 countries and the checklist used by Holvoet and Renard (2007) in their diagnosis of PRSP M&E of eleven SSA countries to name a few. Even though these tools are mainly used for the assessment of central M&E systems, they could be useful for sector M&E systems as well. Whereas the scope of a sector diagnosis is obviously more limited, key components and guiding principles of a sector M&E system largely overlap with those of a central M&E system. An important specific issue within a sector diagnosis is the contribution of sector M&E activities to a central M&E system (Mackay, 2007).

When an M&E system is eventually functioning, it should perform three core processes effectively, efficiently and resourcefully, namely: collecting; interpreting and analyzing; and using processed information (Khan, 2003). Among other things, M&E information should be used for: policy planning and making through national budgets and poverty reduction strategies; supporting the design and ongoing management of activities; and for promoting accountability of government to parliament, civil society and donors (OED, 2002). Unfortunately, quite a few barriers exist to the establishment of M&E systems in developing countries: lack of genuine demand and ownership; lack of a modern culture of fact-based accountability often related to issues of ethics or corruption; lack of evaluation, accounting and auditing skills; poor quality of financial and other performance information and of accounting/auditing standards and systems; lack of evaluation feedback mechanisms into decision-making processes; and the lack of a critical mass needed to develop and support sustainable evaluation systems (Mackay, 1999).

To monitor progress and propose necessary adjustments, constant monitoring and evaluation of the M&E system is needed (World Bank, 2005). "Such evaluations provide the opportunity to review both the demand and the supply sides of the equation and to clarify the extent of actual – as distinct from the hoped-for – extent of utilization of M&E information, as well as the particular ways in which it is being used" (Mackay, 2007: 60).

2.2 Joint Sector Reviews (JSRs)

JSRs have their origin in SWAps. In the shift from project to programme aid, SWAps are used by partner countries and donors to channel aid to a specific sector. SWAps are most common in the health and education sector. Foster (2000:9) 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." Central to a SWAp is the collective responsibility of donors and governments for sector achievements. Unlike project aid, donors are not able to attribute their financial inputs to specific outputs; rather they must justify their individual contributions in terms of progress against jointly-agreed sector objectives (Cassels, 1997). Accordingly therefore, within the SWAp the need for joint assessment instruments has emerged.

Even though JSRs are commonplace in SWAps, so far no standard definition, handbooks or blueprints for JSRs exist. A review, defined by the OECD/DAC (2002) as an assessment of the performance of an intervention, periodically or on ad hoc basis, can be considered an

instrument situated between monitoring¹⁰ and evaluation.¹¹ Unlike monitoring, data provided in a review can give insights into the causes of change (World Bank, 2004, 113), but the assessment in a review is less comprehensive and thorough than an evaluation and it generally 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 of satisfying donor and recipient accountability and learning needs. ‘Performance’ is to be interpreted broadly and may include both a focus on substance at various levels (i.e. inputs, activities, output, outcome and impact, etc.) as well as on underlying, systemic and institutional issues. One of the few existing studies on JSRs highlights that the improvement of M&E frameworks and the promotion of efficiency and transparency through harmonized M&E systems are among the objectives of JSRs (see Packer, 2006). However, the same study also emphasizes that no systematic and/or comparative analysis has yet been performed on the ground to compare these objectives with actual practices.

In most cases JSRs are organized once or twice a year and take about one to two weeks. They are coordinated and led 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 spread over several working groups which focus on specific topics (i.e. quantity and quality of outputs and outcomes, public finance management, human resources, management information system, etc.). Documentary input into JSRs differs from country to country and includes, at best, a combination of various primary and secondary data sources. In most cases, a sector performance report has already been prepared by the sector ministry, including financial reporting from the finance department or the ministry of finance. 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 then feeds into several technical working groups for discussion. Conclusions and recommendations from these discussions are then usually brought together at the Annual Review Meeting (ARM) where they are shared with stakeholders. The Aide Mémoire (AM) is the main documentary output of the JSR.

In the education sectors of Burkina Faso, Mali and Niger, the first JSRs were organised in 2003, 2002 and 2003 respectively. The JSRs in Mali and Niger take place once a year, that of Burkina Faso twice a year. There were no JSRs during the preparation of the new education programme in Mali. In all three countries, one objective of the JSR is to assess the execution of the education programme in the past half year (with a focus on accountability). The JSRs of Burkina Faso and Niger also include the formulation of recommendations for programme improvement in the coming *Compte d’Affectation Spéciale du Trésor* half year (e.g. a focus on learning) among their objectives. This specific objective was also included in Mali’s first JSR, but not in later years. Notwithstanding this lapse, Mali’s Aide Mémoires of later years do include suggestions for improved performance.

3. ECD through JSRs: some evidence from the ground

¹⁰ 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”.

¹¹ OECD/ DAC (2002: 21) defines evaluation as “the systematic and objective assessment of an on-going or completed project, programme or policy, its design, implementation and results”.

Our assessment of the JSRs in the education sectors of Burkina Faso, Mali and Niger led to the conclusion that institutional and systemic issues are largely neglected during JSRs.¹² The JSRs under study disproportionately focused on sector content issues.¹³ This holds true for both accountability and learning objectives which are predominantly centred on the ‘service utilisation plan’ (i.e. the relation between activities and outputs). When attention is paid to institutional and systemic issues, these are in particular related to Public Finance Management (PFM) issues. Keeping in mind the OED’s definition, ECD is nowhere a major issue; none of the 18 JSRs really engage in diagnostic assessments of existing M&E systems or arrangements. No targets are set regarding areas for improvement, nor is there a follow-up of prior recommendations, nor even a straightforward capacity building plan. Only the Aide Mémoire of the 10th Burkina Faso mission highlights that M&E instruments do exist but that they are either not known or used at all levels. The Aide Mémoire furthermore recommends reviewing the existing tools for M&E in order to improve their relevance, efficiency and use by actors at all levels. Since this Aide Mémoire was the last available, it is not clear to what extent the recommendation has been put into practice.

If anything, the attention given to ECD in the last Burkina Faso JSR does not necessarily guarantee substantial improvements over time. The JSR stocktaking, for example, made clear that recommendations formulated during one JSR are not necessarily followed up during subsequent JSRs. This is particularly true for recommendations regarding more complex and sensitive institutional and systemic issues. The Aide Mémoire of the 5th JSR in Burkina Faso is illustrative of this. It quotes the Minister of Education who describes CD as an absolute priority, but problematic as it requires above all institutional changes. The Aide Mémoire acknowledges the complexity of CD by describing its multi-dimensional, transversal and multi-partner character and recommends the elaboration of a global CD plan. However, this recommendation is not pursued in subsequent JSRs. In the 6th JSR, it was recommended to define a methodology for the elaboration of a CD plan. In the 8th JSR, it was recommended to make an internal assessment on the basis of which a CD plan should be elaborated. The internal assessment is still ongoing at the time of the 10th JSR (2007) with recommendations made for the establishment of a working group to develop a timetable for the preparation of a CD plan.

The same picture emerges in Niger, where the second JSR recommends a plan for the global training of staff, thus strongly emphasizing human capacity development. Although this specific recommendation disappeared in subsequent JSRs, the topic remained on the agenda but with a broadened focus towards institutional development. A committee was established to immediately create an action plan for the institutional strengthening of the ministry of education. The 2006 JSR highlights that despite several studies and consultations the component of institutional development is not moving forward. Subsequently, the JSR of 2007 recommends the establishment of a joint technical commission with representatives of the ministry of education and donors charged with the elaboration of the ‘institutional development’ component. Thus, despite the attention JSRs in Burkina Faso and Niger give to (the broader view upon) CD, it seems difficult to make progress.

Due to the lack of attention to ECD, maturation of the systems in these three countries is poor with little effective donor alignment. The analytical quality of the main output of the sector M&E systems, the performance report, is quite low. This is acknowledged in all three countries. In Mali, for example, the third JSR concludes that supplied documents do not allow linkages to be made between activities, objectives and results. Mali’s fifth JSR indicates that

¹² For a more extensive overview and analysis of findings, see Holvoet and Inberg (2008).

¹³ This finding was also confirmed in Rwanda’s 2008 health sector JSR.

reports do not contain an evolution and analysis of performance indicators. The Aide Mémoire stresses that this kind of monitoring and reporting does not permit a results-based management approach. Even though the quality of the reports has over time been improved in Burkina Faso and Niger, the analytical depth of JSRs remains poor and mainly includes blunt statements or recommendations.

While the focus of JSRs on substance is understandable in the short run, failing to invest in systemic issues risks boosting parallel M&E processes and undermining the M&E reform agenda. The neglect of systemic issues runs counter to the earlier recognition of the importance of institutional capacity for the successful implementation of sector-wide approaches (see Cassels, 1997), and is against the logic of the new aid modalities which endorse a shift from ‘content’ to ‘process’, thereby expecting that process conditionality, with particular attention to reforms in the field of public expenditure accounting, budget reform and public service management will lead to greater national ownership of and commitment to poverty reduction policies and programmes (see Booth, 2003).

The next section will examine how JSRs can contribute to ECD, both from a demand and a supply side perspective.

4. ECD through JSRs: some suggestions

Even though the findings show that JSRs, at present, do not explicitly focus on ECD, we believe that they have the potential to function as M&E exercises that can reconcile short and longer term objectives. To achieve this, they especially need to make room for ECD in the short run. While this may be costly, it will lead to less laborious complementary M&E exercises in the long run. Investing in maturing sector M&E systems would, over time, change the outlook of a JSR from a ‘substance’ assessment and learning tool to an approach which focuses on the quality of sector M&E systems, their main outputs as well as their actual degrees of feedback and usage. This is in line with the type of evaluations suggested in Mackay (2007) and referred to above (section 2.1). Such M&E exercises may also be termed “meta-evaluations,” which can be defined as, “a systematic review of evaluations to determine the quality of their processes and findings” (Leeuw and Cooksy, 2005: 95), or an evaluation of an evaluation. Among other things, meta-evaluations are used to identify strengths and weaknesses in evaluation practice in order to develop evaluation capacity (Leeuw and Cooksy, 2005). Two types of meta-evaluations can be distinguished: formative meta-evaluations and summative meta-evaluations. Formative meta-evaluations are conducted during the execution of an evaluation (to improve ongoing evaluation work), summary meta-evaluations are conducted after an evaluation (to assess and diagnose completed evaluations- Stufflebeam, 1974). While JSRs could develop towards a kind of summative meta-evaluation of sector M&E systems in the long run, both types of meta-evaluations may be useful throughout the ECD process.

In what follows, we distinguish between the roles which JSRs could play in capacity building on the M&E supply and demand side.

4.1 Supply side

As suggested by EuropeAid (2005), the outputs of an M&E system are good entry points for ECB. An M&E system at sector level should in principle deliver information from the

management information system (MIS), including data on the performance of individual delivery centres, provide aggregate assessments of sector performance and report on achievements in policy, finance, budgetary, institutional and system development (Cassels, 1997). Normally, one expects all this information to be included in the sector performance reports. These reports are among the main outputs of the sector M&E system, while they also function concurrently as the most important inputs into JSRs. So far, the analytical quality of sector performance reports is poor in the education sectors of the three countries studied, and this obviously influences the overall quality of the JSRs. Even though the low quality is acknowledged and improved in Burkina Faso and Niger, more structural attention to the quality of these reports is essential for the success of JSRs as joint assessment instruments within SWAp. After all, if donors are not able to satisfy their information needs through JSRs, they will be inclined to maintain or to return to their own M&E activities.

While monitoring information is an essential component of performance reports, in itself it is insufficient to assess performance as it usually covers only certain programme dimensions. A more balanced interpretation of performance is provided through evaluations. However, as evaluations are more time consuming and expensive, monitoring is useful to identify problems including possible bottlenecks that need more investigation through evaluation (OED, 2002). An M&E system ideally provides or coordinates these evaluations in order to make more comprehensive and in-depth information available. JSRs generally function as a kind of forum that gathers and triangulates information from various data sources. In doing this, they obviously could also contribute to the identification of remaining information gaps and issues on which more thorough evaluations are useful or necessary.

Within the new aid architecture and in particular its growing attention to results, there is an increasing demand for assessing the impact of programmes. A common approach to assessing overall impact is the performance of an evaluation synthesis. This is, “a study that synthesizes evidence from previous evaluations and other sources in order to come to an overall conclusion about effectiveness” (Leeuw and Cooksy, 2005: 96). A first step in an evaluation synthesis is the performance of a meta-evaluation with the aim of selecting those individual reports which meet quality standards and whose contents are defensible. Another reason to perform a meta-evaluation, besides identifying strengths and weaknesses in evaluation practice and contributing to decisions of specific evaluation findings in evaluation syntheses, is to inform stakeholders about whether and how to use evaluation findings (Leeuw and Cooksy, 2005). This is also relevant in JSRs where evaluations are used to make decisions.

A meta-evaluation normally judges the quality of evaluations or evaluation systems against the professional standards and principles of the evaluation field. The harmonious conduct and the potential impact of a meta-evaluation are enhanced when stakeholders reach a clear understanding of the criteria, principles and/or standards to be applied in the meta-evaluation. Several standards exist including those of the Program Evaluation Standards of the American Joint Committee on Standards for Educational Evaluation (AJCSEE) and the DAC Evaluation Quality Standards.¹⁴ The standards of the AJCSEE are often used by organizations as a basis for the establishment of their own standards.¹⁵ They consist of 30 standards grouped into four categories: utility; feasibility; priority and accuracy.¹⁶ Different standards are intended to

¹⁴ The DAC Evaluation Quality Standards consist of 32 standards divided into 10 categories: rationale, purpose & objective of the evaluation; evaluation scope; context; evaluation methodology; information sources; independence; evaluation ethics; quality assurance; relevance of the evaluation results and completeness (OECD/DAC, 2006a).

¹⁵ Examples include the Swiss Evaluation Society and the African Evaluation Society.

¹⁶ Stufflebeam uses these standards in his definition of meta-evaluation, “meta-evaluation is the process of delineating, obtaining, and applying descriptive information and judgmental information about the utility,

ensure the quality of different M&E dimensions: utility standards focus on the satisfaction of the information needs of intended users; feasibility standards foster evaluations that are realistic, prudent, diplomatic and frugal; priority standards promote evaluations that are conducted legally, ethically, and with due regard to the welfare of those involved in the evaluation, as well as those affected by its results; and finally accuracy standards are intended to ensure that an evaluation will reveal and convey technically adequate information about the features that determine the value or merit of the programme being evaluated (AJCSEE, 1994).

The use of meta-evaluations in JSRs might also be useful as they could provide insight into the specific bottlenecks that block the supply of sound performance and evaluation reports which are the main outputs of the sector M&E system. The identification and analysis of these factors feeds into the identification of the needed CD activities. Optimally, the identification and planning of specific ECD activities should be included in an overall CD plan for the sector, in which all three levels of capacity should be addressed (OECD/DAC, 2006b), and in which clear targets should be set with a corresponding budget (Feinstein and Ingram, 2003). Such a CD plan will improve the coordination of ECD activities which are crucial for ECD efforts. Lundgren and Kennedy (2009) highlight that coherence and efficacy are particularly improved when partner countries and donors actively shape a joint approach (Lundgren and Kennedy, 2009). Since donors will probably finance the implementation of a CD plan, they could use consensual conditionalities,¹⁷ based on jointly-formulated targets to stimulate the implementation of the plan (use of sticks, see 2.1). As JSRs score strongly on coordination and harmonization, they could be an ideal forum for the coordination, harmonization and alignment of different donor ECD activities and implementation monitoring. Where sector M&E systems are not successfully functioning and/or are still in development, JSRs could aid the harmonization of parallel M&E systems.

Besides their contribution to the development of sector evaluation systems, JSRs could also strengthen human evaluation capacities. JSRs could, in fact, potentially function as a kind of learning-by-doing process for participants. They offer ample opportunities to improve general knowledge on ‘monitoring’ and ‘evaluations.’ Where field missions are included, participants could be exposed to different M&E tools and methods such as participatory rural appraisals, beneficiary assessments or they might improve their knowledge regarding issues such as sampling and validity. National universities, research institutes and national M&E platforms and organizations, which are sometimes present in the JSR, could also play an important role in ECD by providing expertise.

4.2 Demand side

Besides being a mechanism for coordination and harmonisation, the JSRs in Burkina Faso, Mali and Niger score quite strongly on ownership. In all cases the ministries of education take the lead in the JSR, alone or jointly with the lead sector donor. Moreover, the JSRs commonly involve a broad range of actors, both from within and outside government. Through their participation in the ARM at the end of a JSR, participants might get a shared understanding about the results as well as an increased sense of engagement and ownership.¹⁸ This is

feasibility, propriety, and accuracy of an evaluation in order to guide the evaluation and to publicly report its strengths and weaknesses” (2000: 96).

¹⁷ Consensual conditionality can be defined as, “a conditionality for which there exists a genuine measure of ownership on the recipient side” (Molenaers and Renard, 2008: 12). This type of conditionality is more likely to be effective than the ‘hard’ conditionalities which were applied by the IMF and World Bank in the context of Structural Adjustment Programmes which were often not based on a country-owned policy or plan (see Molenaers and Renard, 2007).

¹⁸ This is related to Patton’s notion of the ‘process use’ of evaluations (see Patton, 1998).

potentially conducive to ECD as ownership and demand for M&E has been identified as crucial for successful ECD. ECD is more likely to be sustainable if the sector ministry puts ECD on the agenda of a JSR and formulates demands for M&E information. Where sector ministries are not acting as M&E demand side actors (after all, lack of demand and ownership are among the major bottlenecks, see above), there are still others which might pressurize for M&E. Stakeholders such as civil society organisations or parliamentarians, who operate outside the executive, are also potentially important M&E demand side actors. If the M&E capacity of national non-government actors improves, it may help them to hold their own governments more accountable and may eventually contribute towards ‘downward accountability.’

Donors could contribute as well by emphasizing the merits of evaluation systems (use of sermons, see 2.1), “but until countries accept the strengths of such arguments; or themselves conclude that evaluations have much to offer, an evaluation system is unlikely to be sustainable” (Mackay, 1999: 4). Spotlighting the outputs of evaluation systems early in the ECB process (e.g. through an influential high-quality evaluation) could help convince concerned individuals of their merits (Lundgren and Kennedy, 2009).

5. Conclusion

While the Paris Declaration (PD) sets out a major reform agenda regarding M&E, actual implementation in the field lags behind. Various surveys have demonstrated that partner countries did not succeed in establishing results-oriented frameworks. Donors, from their side, are not inclined to use recipient arrangements which currently do not meet existing accountability and learning needs. This reluctance, however, concurrently blocks the set-up and development of recipient systems. To move beyond this stasis, donors and partner countries need to seek a middle way between the satisfaction of existing short run M&E needs and the construction of recipient M&E frameworks in the longer run. Joint Sector Reviews (JSRs), which are essential components of SWAPs, might fill the gap.

This article analyses whether JSRs could contribute to Evaluation Capacity Development (ECD), defined by the OED/IEG (2004: 1) as, “the process of strengthening country-based systems to conduct and use M&E.” Potential ECD entry-points are explored on the basis of previous stocktakings and analyses (by the authors) of actual JSRs in the education sectors of Burkina Faso, Mali and Niger. Findings of this research demonstrate that the development of sector M&E systems are not a declared or major issue in any of the JSRs. JSRs do not pay systematic attention to systemic or institutional issues, but rather focus on content issues. However, in order to contribute to the PD reform agenda, JSRs should preferably evolve from an assessment of sector ‘content’ towards a kind of meta-evaluation in which the quality of the sector M&E system and its main outputs are assessed. In the mean time, JSRs could contribute to ECD in several ways. First of all, one of the potential entry points for ECD are sector performance reports which are among the most important outputs of the sector M&E system. As the same reports provide essential inputs to JSRs, the latter may function as a good forum for the improvement of these reports. Furthermore, JSRs score quite positively on coordination, harmonization and ownership. Thus they could prove useful as a forum for the coordination and harmonization of ECD initiatives, which are essential for effective ECD. As ownership is a prerequisite for ECD, the existing high levels of ownership by the government and non-government actors of JSRs might increase the demand for M&E information and for related ECD activities. Additionally, JSRs could play a role in the selection of topics for more in-depth evaluation and in decisions related to whether and how to use evaluation findings.

Finally, JSRs could promote the development of human evaluation capacity, in particular through participant learning-by-doing on field missions and attendance/participation at Annual Review Meetings (ARMs).

In summary, JSRs have the potential to contribute to ECD and to function eventually as a meta-evaluation exercise. While this may demand considerable (additional) short run investment, it will lead to less laborious complementary M&E exercises in the long run. If this is not done, JSRs are in danger of becoming institutionalized as parallel M&E exercises that persistently replace the recipient's M&E systems, thereby undermining the M&E reform agenda and the new aid architecture.

References

3rd High Level Forum on Aid Effectiveness. 2008. *Accra Agenda for Action*.

American Joint Committee on Standards for Educational Evaluation. 1994. *Program Evaluation Standards*. Sage Publications.

Baser H., Morgan P. 2008. Capacity, Change and Performance. Study Report. *Discussion Paper No 59B. ECDPM*.

Bedi T., Coudouel A., Cox M., Goldstein M. and Thornton N. 2006. *Beyond the Numbers. Understanding the Institutions for Poverty Reduction Strategies*. World Bank: Washington D.C.

Booth D., Lucas H. 2002. Good Practice in the Development of PRSP Indicators and Monitoring Systems. *ODI Working Paper No. 172*. Overseas Development Institute (ODI): London.

Booth D. 2003. Introduction and Overview. *Development Policy Review* 21 (2): 131-159.

Booth D., Nsabagasani X. 2005. Poverty Monitoring Systems: An Analysis of Institutional Arrangements in Uganda. *ODI Working Paper No. 246*. Overseas Development Institute (ODI): London.

Cassels A. 1997. *A guide to sector-wide approaches for health development*. WHO: Geneva.

EuropeAid 2005. Institutional Assessment and Capacity Development. Why, what and how? *Tools and Methods Series. Reference Document No. 1*. European Communities: Luxembourg.

Feinstein O., Ingram G. 2003. *Lessons learnt from World Bank experiences in joint evaluation* (internal 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.

Holvoet N., Inberg L. 2008. Joint Sector Reviews: M&E experiments in an era of changing aid modalities. Experiences from JSRs in the education sectors of Burkina Faso, Mali and

- Niger. Unpublished paper. Institute of Development Policy and Management: University of Antwerp.
- Holvoet N., Renard R. 2007. Monitoring and Evaluation under the PRSP: Solid Rock or Quicksand? *Evaluation and Program Planning* 30: 66-81.
- Khan K. 2003. *Strengthening of Monitoring and Evaluation System*. Pakistan Poverty Alleviation Fund: Islamabad.
- Kusek J.Z., Rist R.C. 2002. 'Building Results-Based Monitoring and Evaluation Systems: Assessing Developing Countries Readiness.' *Zeitschrift für Evaluation* 1: 151-8.
- Kusek J.Z., Rist R.C. 2004. *Ten Steps to a Results-Based Monitoring and Evaluation System. A Handbook for Development Practitioners*. World Bank: Washington D.C.
- Leeuw F. L., Cooksy L. J. 2005. Evaluating the performance of development agencies: The role of metaevaluations in G. K. Pitman, O. N. Feinstein & G. K. Ingram (Eds.), *World Bank series on evaluation and development: Vol. 7. Evaluating development effectiveness* (pp. 95-108). New Brunswick, NJ: Transaction Publishing.
- Lundgren H., Kennedy M. 2009. Supporting Partner Country Ownership and Capacity in Development Evaluation. The OECD DAC Evaluation Network. In UNICEF, *Country-led Monitoring and Evaluation Systems. Better evidence, better policies, better development results*.
- Mackay K. 1999. A Diagnostic Guide and Action Framework. *Evaluation Capacity Building - ECD Working Series No. 6*. World Bank: Washington D.C.
- Mackay K. 2007. *How to Build M&E Systems to Support Better Government*. The World Bank: Washington D.C.
- Molenaers, N., Renard, R. 2008. Policy Dialogue under the New Aid Approach: Which Role for Medium-sized Donors? Theoretical Reflections and Views from the Field. *Discussion Paper* 2008.05. Institute of Development Policy and Management: University of Antwerp.
- OECD/DAC. 2002. *Glossary of Key Terms in Evaluation and Results Based Management*. OECD/DAC: Paris.
- OECD/DAC. 2005. *Paris Declaration on Aid Effectiveness*. OECD/DAC: Paris.
- OECD/DAC. 2006a. *DAC Evaluation Quality Standardse*. OECD/DAC: Paris.
- OECD/DAC. 2006b. *The Challenge of Capacity Development, Working Towards Good Practice*. OECD/DAC: Paris
- OECD/DAC. 2008. *2008 Survey on Monitoring the Paris Declaration. Effective Aid by 2010 ? What it will take?* OECD/DAC: Paris.
- OED. 2002. *Annual Report on Evaluation Capacity Development*. World Bank: Washington, D.C.

- OED. 2004. *Evaluation Capacity Development: OED Self-Evaluation*. World Bank: Washington, D.C.
- 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.
- Stufflebeam D.L. 1974. Meta-evaluation. *Occasional Paper Series* No. 3. Western Michigan University Evaluation Center: Kalamazoo MI.
- Working Party on Aid Effectiveness. 2008. *Aid Effectiveness. A Progress Report on Implementing the Paris Declaration*.
- World Bank. 2001. *Education and Health in Sub-Sahara Africa. A Review of Sector-Wide Approaches*. World Bank: Washington, D.C.
- World Bank. 2002. *A Source Book for Poverty Reduction Strategies*. World Bank: Washington, D.C.
- World Bank. 2005. *Capacity Building in Africa. An OED Evaluation of World Bank Support*. World Bank: Washington D.C.
- World Bank. *Compte d'Affectation Spéciale du Trésor 2007. Results-Based National Development Strategies. Assessment and Challenges Ahead*. World Bank: Washington D.C.
- Wubnek M. 2003. *Building Capacity in Africa: The Impact of Institutional, Policy and Resource Factors*. African Development Bank.