Capacity Development Workshop
"Country Leadership and Implementation for Results in the EFA FTI Partnership"
Cape Town, South Africa, July 16-19, 2007

QUALITY OF EDUCATION SECTOR PLANS TO ACHIEVE UNIVERSAL QUALITY PRIMARY EDUCATION

ROSEMARY BELLEW

DRAFT

This paper was prepared for an Education for All-Fast Track Initiative (EFA-FTI) workshop entitled "Country Leadership and Implementation for Results" organized by the World Bank (July 16-19, 2007). The purpose of the paper is to motivate a discussion and an exchange of experience on key issues country participants have encountered in developing a quality plan to achieve universal high quality education, and how they have addressed them for a session entitled "Elements of a good strategic plan." The paper is intended to provide a spring board for the discussions. It summarizes an assessment made by the author on the quality of existing FTI country plans, highlights strengths and weaknesses, and presents some ideas, tools and processes used by other countries that may help to country plans.

The paper is structured as follows. Section I proposes a list of elements that one might expect to find in a quality plan that aims to achieve universal primary completion for all and assesses 25 country plans in order to identify if there are systematic strengths and weaknesses across countries, and to identify areas of good practice. Section II presents a few key strategies and tools countries have used to strengthen their plans in key areas that the assessment found to be weakly addressed in the FTI country plans.

I. STRENGTHS AND WEAKNESSES IN THE QUALITY OF FTI-ENDORSED COUNTRY EDUCATION SECTOR PLANS

A main objective of the Education for All-Fast Track Initiative (EFA-FTI) is to provide support to low-income countries to develop and implement "credible" sector plans to achieve the EFA and Millennium Development Goal (MDG) of complete quality primary education for all children. Since the launch of the FTI in 2004, 28 countries have sector plans endorsed as credible by the development partners at the country level. This section explores the strengths and weaknesses of 25 FTI-endorsed country education sector that were available at the time of the review.

In order to assess the overall quality of country plans, it was first necessary to develop criteria against which the plans could be assessed. The FTI Framework document, the FTI Appraisal Guidelines and the goal of improved aid effectiveness provide some indication of elements that one would consider essential to plan quality, or credibility. Based on the FTI documents, the on-going work on aid effectiveness, and the quality assessment dimensions used by the World Bank when assessing the quality at entry of World Bank operations, I have created a list of 46 criteria grouped according to the seven dimensions below. The full list of criteria is in the annex to this paper.

1. Technical Quality
2. Leadership and Ownership
3. Institutional and Implementation Aspects
4. Monitoring and Evaluation/Accountability Arrangements
5. Financial and Economic Aspects
6. Fiduciary Safeguards

7. Environmental Aspects which concern whether an environmental impact and risk assessment was conducted and if the plan incorporates arrangements to mitigate any negative impacts of the program on the environment.

In order to get a sense of how the plans may vary in quality, and whether there are any systematic strengths or weaknesses, I have assessed the quality of the 25 country plans that were available on each of the 46 criteria. The rating scheme applied is a very simple 0,1 rating with 0 indicting the topic was not mentioned or covered in the plan, and 1 indicating that is was. Some indicators, however, such as the financial estimates and track record, which are a composite of more than one variable, I have used a .5 rating to denote partial treatment or a partial result.

For most criteria, no judgment was made as to the scope or adequacy of coverage. The exception is the quality of data. In particular, no judgment was made as to the adequacy of a particular strategy a county has chosen to pursue because the effectiveness of a particular strategy may be highly contextual. Further, it is difficult to determine if and what other strategies had been considered.

This approach to the assessment of the country plans has its weaknesses. The indicators may not be adequate to the task, and shades of grey may not be well captured. More importantly, the review is based on the plan documentation only, which may not fully reflect reality, and one peer reviewer is inadequate to achieve a reasonably unbiased and consensual assessment. However, it is useful to note that a report prepared by the FTI Secretariat on the quality of country programs shows similar results.

Given these caveats, the purpose of the review was not to assign scores to country plans, nor to rank countries, but rather to determine strengths and weaknesses with respect to aspects considered important by the FTI partnership. It is also preferable to analyze the results by indicator rather than by the total score across all indicators because the number of criteria for each indicator varies, with technical indicators being the largest, and countries vary in the depth of treatment of the criteria. Thus, a ranking by total score would not necessarily indicate a higher quality plan across all areas. However, we can point to examples of better practice within each of the indicators.

Analysis Results

Overall, the quality of the plans varies considerably, with some notable systematic strengths and weaknesses. One clear omission from country plans was the environmental considerations. Although most plans included a large school construction or rehabilitation program, often including latrines, none appeared to have developed, or plan to develop, a process or norms to ensure minimum risk to the environment. Since environmental factors were not obviously considered in any of the country plans, no analysis is presented of this indicator.

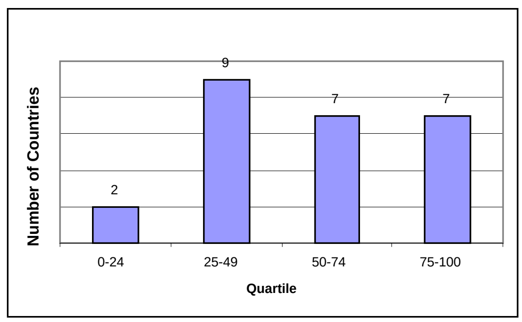
Technical Quality

The Technical/Education indicators are the largest among the seven categories of indicators. These 25 indicators cover elements such as:

- Adequacy of (minimum) data
- Analysis of priority issues, such as quality, gender equity, rural access, HIV/AIDS prevention, and meeting the needs of the poor and vulnerable children
- If there is a long-term strategic direction and medium term plan to achieve UPC,
- If the plan has clear strategies to raise completion rates, improve quality, equity, internal efficiency and other high priority issues.
- Benchmarking against the Indicative Framework
- If the plan's objectives, goals and strategies are prioritized
- Incorporates lessons learned
- Financial analysis of trade offs using a simulation model

On average, the country plans score well with respect to the technical aspects of education. Chart 1 shows the distribution of country scores by quartile where the country score is the percentage of criteria met out of a total 25 criteria. Almost 60 percent of the country plans address half to all of the technical criteria. The frequency distribution suggests that while some countries can certainly do better, many countries have built their capacity to prepare technically sound education programs as measured against the technical criteria.

Chart 1. Frequency Distribution of Technical Ratings (25 criteria)



The majority of plans is grounded in sector analysis, is benchmarked against the indicative framework, has analyzed and contains strategies to address regional, gender and other disparities, and has strategies to raise education quality and improve resource use. The majority of programs also incorporate prior lessons learned.

Medium term plan for primary education. Most countries, although not all, have a medium-term plan of 3-5 years to achieve universal primary education. The few countries (4) that do not have, instead, long-term targets and interventions to achieve them but the interventions are not well prioritized and the program costing omits significant expenditure such as the cost of salaries and operating expenses.

Sector Analysis. Judged by the analysis contained in the country plans and the bibliography of materials consulted by the local donor group when appraising the country's program, over 60 percent of country strategies are underpinned by reasonable sector analysis. Most analyze internal efficiency indicators, inputs and learning outcomes. The majority of plans present an analysis of disparities in enrolment by region, gender and poverty category, although about 40 percent neglected to examine disparities and their causes.

Policies and Strategies

- Education quality. Without exception, all plans include a strategy to improve education quality from an input perspective, generally in terms of increased or better resources such as textbooks, teacher education, reforming the curriculum, etc. And, two-thirds of the plans reflect a focus on monitoring learning outcomes or intervening in a more direct way to address specific learning weaknesses.

Gender equality. The majority of countries have strategies to promote gender equality. I have made no judgment on the adequacy of these strategies but note, however, that in many cases the strategy is often non-strategic consisting either of a study and information campaign, or a long list of un-prioritized interventions which are not cost neutral, may not be effective, and may tax capacity to implement. For example, the plan to increase girl's education in one country included 12 activities to be implemented simultaneously as follows:

- Reduce the cost of school to poor families by providing stipends and other incentives
- Expand access to the second cycle in rural areas
- Provide separate latrine facilities
- Increase the number of school cantines
- Put in place a "big-sister" system
- Provide tutoring
- Provide examination preparation
- Develop local strategies for girls in particularly difficult circumstances
- Recruit more female teachers
- Train teachers in gender sensitivity
- Adopt gender quotas for teacher training
- Conduct an information campaign

Benchmarking Against the Indicative Framework - 70 percent of plans utilized the indicative framework to set targets, although a few of them did so only partially.

Weaknesses

Across all countries, three large and systematic gaps were evident in the following three key policy-related areas :

- Estimating costs and policy trade offs. A large gap is found in whether countries show evidence of having used a financial simulation model to estimate costs and policy trade-offs, particularly concerning internal efficiency. Although a financial simulation model is a powerful tool for assessing the financial impact of a country's objectives and policy trade-offs, very few countries show evidence in their plans of actually having used one. This may be a surprising observation to some especially in the Africa region where the World Bank has supported significant analytical work in countries using financial simulation models to diagnose the impact of policy choices on resource requirements. The lack of an apparent use of a financial model may be due to several factors. First, although models may have been used to assess needs and the impact of policy choices on revenues required, sustained capacity has not been build to integrate these models as on-going tools for policy development. Or, it may be that the links between the policy and financial simulations and country plan development may need to be strengthened.

- HIV/AIDS and vulnerable children. Analysis of, and strategies to address, issues related to HIV/AIDs, orphans and vulnerable children and children with disabilities. Of all criteria, these were the most neglected areas and quite universally so.

Other common weaknesses that are evident in about half the country plans include:

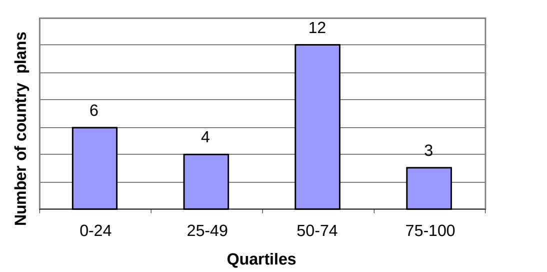
- An absence of demand-promoting strategies, although most countries have assessed disparities by income/poverty level, and have high levels of disparity.
- Prioritization of objectives or interventions. Whether there is a need to prioritize is highly context specific and difficult to assess based on program documentation only. Nevertheless, it is indicative that many assessments by the local donor groups raise reservations about country capacity and frequently characterize country programs as over-ambitious.
- Mechanical use of the Indicative Framework which results in mechanical goals and cost estimates that are not clearly linked with strategies or realistic financing plan.

Leadership/Ownership

Leadership and ownership are measured by 4 criteria -- evidence of political support for the program, the extent of stakeholder consultations, participation of civil society in plan development, and a clear description of how the donors will work together under the coordination of the country government.

Chart 2 shows the frequency distribution of the percentage of criteria attained by quartile. As for technical quality, about 60 percent of country plans fall within the top two quartiles. However, a larger number of country plans fall into the lowest quartile on this indicator. Four of the six countries also score in the bottom two quartiles on technical quality, suggesting a relationship between country leadership and technical quality, although whether this is a causal relationship and in which direction cannot be determined.

Chart 2. Frequency Distribution of Percentage Score on Leadership and Ownership (4 criteria)



By and large, however, across all country programs there are two systematic gaps:

- Donor coordination and review arrangements are covered in only half the programs, and mainly only very partially. In many instances, donor coordination is at an initial stage, and in some others it is not at as advanced as one might expect it to be given the long history of donor engagement in the country.

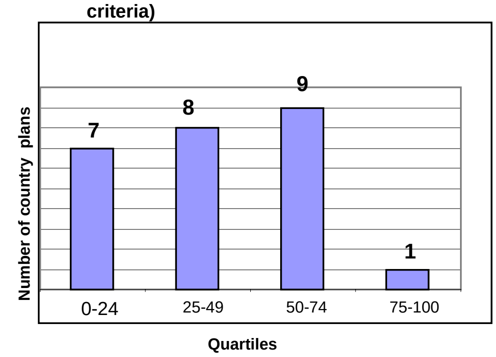
- While in the majority of countries, there is evidence of extensive consultation within the education Ministry, there is considerably less evidence of broader government consultations and the participation of civil society in the formulation of the plan. This is an important omission. Care must be taken to ensure that the education plan is not only a plan supported by upper levels of the MoE and donors but that there is broad ownership and commitment to its goals throughout the MoE, within other critical government agencies, and among the intended beneficiaries.

Institutional and Implementation Aspects

Institutional and Implementation aspects concern whether there is a short or medium term implementation plan or whether there is a plan to develop one, whether there is evidence of an analysis of implementation capacity constraints, and whether strategies are proposed to address these constraints.

Chart 3 shows the frequency distribution of the plan score as a percentage of the total possible (4). Compared to technical and leadership indicators, fewer countries fall into the top two performance quartiles (10 countries or 40 percent), while 60 percent more fall into the bottom two quartiles.

Chart 3. Frequency Distribution of Percentage Score on Institutional and Implementation Aspects (4 criteria)



Gaps

The main gap one observes across countries is the absence of an implementation plan. At this stage of the country plans, only two countries show evidence of having an implementation plan. The large majority of countries, however, recognizes this gap and plans to start developing annual plans. Nevertheless, it is a significant gap as it is difficult to determine the feasibility of achieving the objectives in the absence of a time line of critical steps or actions that need to be taken to achieve the intended results.

The second gap relates to the identification of capacity constraints and proposed strategies to address them which is found in only 40 percent of plans. Generally, however, when capacity issues were treated in the plan, it mainly took the form of assigning specific implementation responsibilities to directions with the Ministry of Education which is an appropriate start. Lacking is any commentary or analysis on whether the human resources, the institutional environment or the incentives are present to succeed which may reflect limited capacity to analyze and assess institutional requirements. Capacity development is itself a broad topic outside the scope of this paper.

Monitoring and Evaluation/Accountability Arrangements

The monitoring and evaluation indicator includes 3 criteria that capture whether the country plan has:

- Annual targets for measuring progress on key actions and outcomes, including those of the FTI indicative framework as locally adapted

- Monitoring and evaluation arrangements that are reasonably clearly detailed

- Specific arrangements by which donors and other partners will review progress and validate monitoring and evaluation results.

Chart 4. Frequency Distribution of Percentage Score on Monitoring and Evaluation (3 criteria)

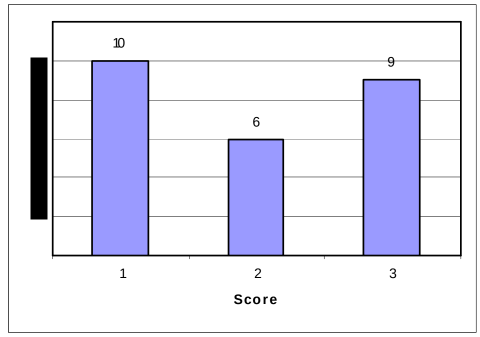


Chart 4 shows that a large number of country plans (60 percent) fall into the bottom third indicating that they have responded to none or only one M&E criteria. Most countries in the bottom third omitted monitoring, evaluation, review, and feedback mechanisms entirely while the incorporation of a process to improve accountability to beneficiaries was extremely rare.

The most common form of strategy for monitoring and evaluation was to set up or improve an EMIS to organize and analyze data gathered through annual school censuses. The most common form of review and feedback process was conducting annual reviews of progress with the local donor group.

While these two strategies may be essential, they are often not sufficient. Administrative data is often subject to error or manipulation (e.g. in situations of per capita financing, in highly centralized countries) which makes an external check on the data useful. Some information required may not be obtained from a school based census (e.g. reliable information on household income) nor may the information need require data on the entire population and can, therefore, be obtained more economically through sample surveys. Administrative data also provide little information on the exogenous factors that influence the demand for education and therefore limit guidance largely to supply side interventions. Further, impact evaluations are few.

Financial and Economic Indicators

Financial and Economic criteria relate to the existence of a financial framework that includes domestic and external resources, its credibility as assessed by the local donor group, and whether the projected financial requirements are integrated into the country's macro framework and MTEF.

Chart 5. Frequency Distribution of Percentage Score on Financial and Economic Aspects (3 criteria)

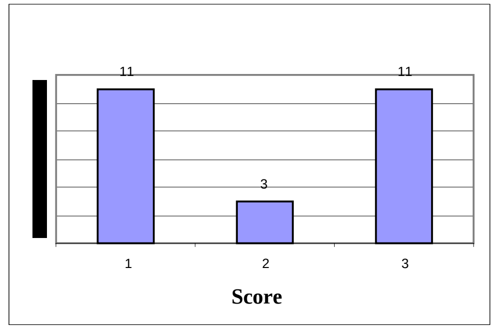


Chart 5 shows the frequency distribution of country plans according to their percentage score on these (3) indicators. Relative to the distributions found for technical quality and evidence of leadership, there are a significantly higher number of countries that score very poorly on putting together a comprehensive financial framework and linking it with an overall MTEF.

When assessing financial requirements, countries have utilized two approaches to types of tools -- an education cost simulation model or a budget model – neither of which captures the full cost. While both capture capital costs, the simulation models tend to capture better the large recurrent cost elements such as the projected wage bill under various assumptions while the budget approaches tend to capture better the detailed costs of the program and the cost of management and capacity building. In large part, countries and local donors are aware of these weaknesses and plan to work on them.

Fiduciary Safeguards

The treatment of fiduciary arrangements is perhaps the weakest aspect of the country plans. This indicator captures 5 criteria concerning whether the country plan:

- Mentions anything about the prevailing public financial management and procurement environment,
- Describes the financial management and procurement arrangements envisaged for the program, and
- Has a plan for the management of external resources.

Chart 6. Frequency Distribution of Percentage Score on Fiduciary Safeguards (5 criteria)

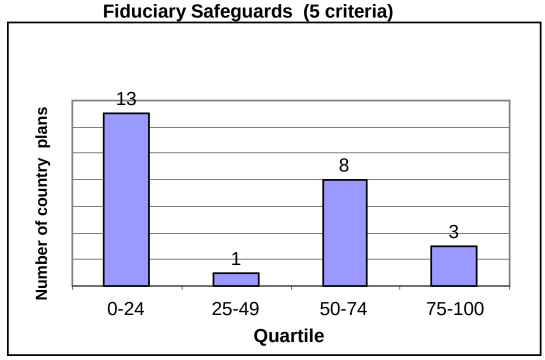


Chart 6 shows the results. The lowest scoring countries show no evidence of having a plan or having considered how the significantly higher level of resources required by the plan will be managed and whether capacity exists to manage the flows. Across all plans, procurement was a neglected area. Having in place appropriate and transparent government procurement procedures is an integral part of effective resource management. Procurement procedures and practices can strongly influence the cost of the program, the quality of inputs relative to price, and the speed at which inputs can be delivered.

Conclusions

Based on the criteria utilized in the analysis, country plans do show some areas that may be fruitful avenues for further exploration to improve country plans. However, it should be noted at the outset that achieving universal completion of quality primary education may require not only good policies and strategies, a credible and fully resourced plan, and robust financial management and procurement systems, but also strategies to build demand and momentum for change, particularly when major reforms are involved.

It is also evident that the plans represent a vast improvement compared with country plans of a decade ago and compared with a project approach to sector development even though most financing is still channeled through project financing arrangements, or compared with prior country plans. There is a very clear evolution towards sub-sector or sector wide approaches with some countries being more advanced that others.

Given the many caveats that must be placed on the type of assessment conducted by the author, and the variety of country contexts, it is difficult to draw firm conclusions. Nevertheless, the assessment points to the following areas that could be strengthened across a large number of the plans:

- Selectivity/prioritization of objectives and strategies
- Monitoring, evaluation and accountability to beneficiaries
- Civil Society engagement and demand-side policies
- Donor partnerships
- Improving public financial management and procurement in the education sector to enable an increased flow of resources through public finance channels, and establish robust mechanisms when budget management is decentralized to the district and school levels.

II. Tools and Processes

This section presents some tools and processes that countries have used to strengthen the first four of these gaps. Financial management and procurement aspects are not addressed in this paper. This section draws not only from the 25 FTI programs but also from examples from other countries. The processes and tools presented are not intended to be comprehensive or exhaustive but are intended to stimulate discussion among workshop participants about whether the gaps identified resonate with their situation and to share their experience about the approaches they have used in the development and implementation of their plans. Because some tools integrate civil society input into the instruments and processes for monitoring and evaluation, I present a collection of ideas tools and processes in three categories -- prioritization, methods for M&E that provide information on demand side constraints and encourage civil society participation or demand, and donor partnerships.

1. Prioritization/Selectivity

Countries with overambitious plans relative to capacity or financing, take a passive approach to prioritization. In these situations, by definition, not all objectives will be reached. The risk is that managers are likely to implement only those parts of the program that are easier to implement (e.g. school construction in urban areas or near main roads, studies), and/or have high benefits and low costs to the managers (e.g. workshops).

Other countries have utilized three main approaches to prioritization depending on (a) the degree and extent of support for the reforms being introduced, (b) implementation capacity and (c) financial capacity. Prioritization was rarely obvious in the choice of policies or interventions based on their likely impact on outcomes. Demand side strategies, particularly to raise girl's education, and strategies to improve quality often fall into this category. The example in Box 1 is an illustration of this. The point is not that all these interventions may not be useful, but that some may be more cost-effective than others.

Prioritization in an uncertain reform environment

Two approaches taken to prioritization when sensitive reforms require further consultations are what may be called "quick win" and "quick decision" approaches.

- Quick wins. The main prioritization criterion in a quick wins approach is to focus on visible results, generally increased access and more books, in order to build broad-based support for the program's goals, or raise demand for education, while planning consultations and a pathway to address more sensitive issues. There is, however, a risk that the reforms will never be adopted and implemented. History provides no shortage of examples – higher education financing in most francophone African countries, centralized approaches to school construction, performance based pay. However, for primary education, it may be the only feasible approach in the short run to build momentum in countries lacking strong political and stakeholder support for change, or where primary education is not a top priority.

- Quick Decisions. In other countries where the priority accorded to UPC was more obvious (e.g. level of public expenditure), the approach taken was to identify the most important issues on which immediate decisions could be taken while setting a path by which other issues would be worked on and addressed. A common example of this approach is the recruitment of contract teachers to resolve teacher shortages while more fundamental reforms in salary structures or professional ladders are developed.

Prioritization when implementation capacity is weak

Two approaches taken to prioritization based on capacity constraints include reducing the number of interventions or the scope of implementation by phasing in implementation geographically. These are not mutually exclusive and have been used in combination.

- Reducing interventions. Timor Leste, and the Sindh and Punjab Provinces in Pakistan provide examples of countries that have taken an approach of limiting the number of interventions to a few key ones that address the most critical gaps in access or quality while focusing efforts on building capacities and systems for financial management, procurement, monitoring and policy analysis/evaluation. In a second stage, or medium term plan, the program is gradually scaled up and expands in scope once reasonably reliable systems are in place to both manage a significantly increased volume of resources and provide reliable indications of results. This approach was highly successful in Sindh Province of Pakistan and has been replicated in other provinces.

- Phasing in Implementation. Some countries have started in regions with either stronger management capacity, or where the needs were greatest, in order to test out the strategy, learn lessons and incorporate those lessons into the program design before gradually scaling up to other regions. This learning-by doing approach has obvious advantages serving both to strengthen capacity and improve the program design which should result in greater efficiency in implementation when scaled up to other regions. Often, however, experience shows that geographically phased implementation is not always politically or ethically acceptable, particularly when needs are vast.

Tools to Aid in Prioritization

- Financial Simulation Models – A simulation model is a useful, indeed necessary tool especially at the stage of plan preparation. It aids in identifying and assessing critical policy choices (e.g. student:teacher ratios, class size) that have a large impact on recurrent and capital costs and financial sustainability over the long run under various assumptions of resource mobilization.

- Budget models and MTEFs are tools for prioritization and resource allocation in the medium-term, generally over a three-year period. The MTEF sets the overall financial envelope determined by the Ministry of Finance, while the budget model allows estimates to be made within this framework, and allocates them across the sub-sectors and program components.

In order to facilitate the linkage between the budget and long term policy impacts, Rwanda, for example has integrated a budget model within a finance simulation model in order to accomplish both with one tool. The tool is more cumbersome but captures other costs and is directly linked with the country's budget structure for easy integration which both eliminates the need for parallel exercises and maintains the long-term financial simulation model up to date when resources, costs or policies change.

- Impact evaluations have been increasingly conducted for programs in Latin America and Asia, but they have rarely been used for education policy making in Africa.[1] Impact evaluations seek to answer the question: Are we doing the right things? Impact evaluations can be used to assess whether a particular intervention is having the desired impact, and to compare the relative effectiveness of different interventions. When combined with cost data, it provides robust information on the relative cost-effectiveness of different interventions.

Impact evaluations seek to establish causality between the intervention and the measured outcome. In order to place confidence in the causality link, a control group that has not been exposed to the intervention, that will allow the attribution of any change in the treatment group to the intervention, is necessary. The control group must be identical in observable and unobservable dimensions as much as possible and not subject to any spillover effects from the intervention itself to net out potential omitted variable or selectivity bias. Randomized selection of treatment and control groups with pre (baseline)- and post-treatment data collection provide the greatest assurances of capturing the effect of the program only.

Two examples of how impact evaluations have been used in the education sector come from Kenya where a series of evaluations assessed the impact and costs of three different interventions that aimed to increase primary school attendance. The most cost-effective intervention was the administration of a twice yearly deworming treatment which raised attendance at an average cost of $3.50 per additional year of schooling. Providing school meals to pre-schoolers cost an average of $36 per additional year of schooling, and the provision of free uniforms cost $99 (Duflo et al 2007).

In another study (Glewwe et al. 2000), the authors show that the impact of flip charts in schools on learning achievement is significantly overestimated in traditional retrospective studies that analyze the effect of educational inputs on learning outcomes. In retrospective production-function type studies, flip charts were estimated to increase student achievement by 21 percent of a standard deviation, compared with only 5 percent when flip-chart students are compared with non-flip chart students. The authors attribute this large difference to omitted variable bias where observed inputs such as textbooks or more qualified teachers are correlated with other unobserved variables (e.g. when parents who provide better home environments are also more successful at mobilizing more school resources) that affect learning outcomes.

Although randomization is not always possible when, for example, programs are targeted or participation is voluntary, or when an ethical dilemma is posed by denying investments to other groups, there are various methods for designing comparison groups, data requirements and statistical methods that can minimize selectivity or omitted variable

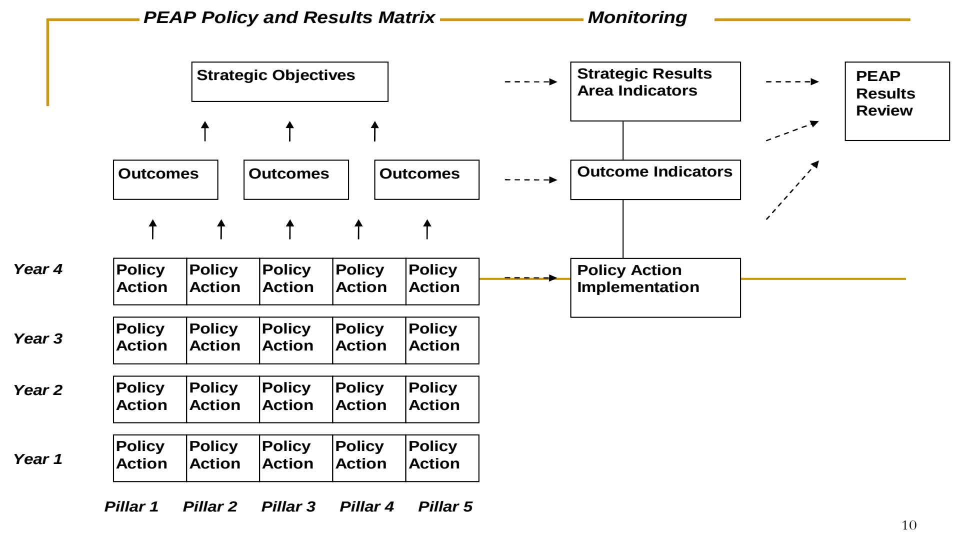
[1] Several examples are an evaluation of a voucher program in Colombia, El Salvador's EDUCO program, Nicaragua's school autonomy reform.

bias depending on the objective of the program. For example, the two groups can receive different improvements or the same improvements at different dates as long as the two groups do not get the same intervention at the same time. A treatment of the various statistical methods and methods of impact evaluation (randomized/experimental, pipeline comparison, single difference, double difference, etc) and their impact on the sample, comparison group, and data and statistical methods is outside the scope of this paper. Interested readers may wish to consult the references for more information. The point I would like to make here is that there are many ways in which impact evaluations can be conducted. By knowing who is eligible for the program, where the program will be implemented and when the intervention will take place, it is always possible to design a valid control group and methodology that adapts to the specificities of the program.

Impact evaluations demand, however, a substantial amount of information, time and resources. It is necessary, therefore, to sufficiently narrow the interventions to be evaluated by selecting those that are of particular strategic importance, and/or those that have a large impact on costs.

- Results Framework and Policy Matrices . One of the commitments that signatories to the Paris Declaration on Aid Effectiveness have made concerns making efforts to improve results frameworks. In response, increasingly, countries are using logical frameworks to develop results frameworks and policy matrices that are used to monitor key outcomes, outputs and actions in the medium-term that are SMART -- specific, measurable, appropriate, relevant, and time-bound -- to gauge successful implementation. In principle, results frameworks are characterized by fewer and better indicators to keep the monitoring effort manageable, with a focus on policy and institutional changes that contribute to desired outcomes, and flexibility with a process that allows for revision on an annual basis through a (budget) review cycle building incrementally on progress and results achieved. The logic is that a set of policy actions can positively influence outcomes. Success or failure is assessed against both the big picture and critical actions assumed to affect big picture trends including further policy and plan development. The schematic depiction below shows the logic of the framework developed by the Ugandan Government for its Poverty Reduction Program.

The Logical Framework of PEAP Results Review:



Source: Presentation by Peter M. Ssentongo, Office of the Prime Minister, Uganda. "Translating Poverty Reduction Strategy Objectives into Results" for The Workshop on the Paris Declaration November 16-17, Entebbe.

The table below shows an example of a Results Framework adapted from a project in the Punjab, Pakistan. The Pillars of the program are enhancing fiscal sustainability, equitable access to education, improving quality and relevance, public sector governance and public sector management. Outcomes pertain both to the big picture as well as to critical actions associated with achieving the outcomes over a three-year period.

Results Framework and Policy Matrix: Punjab Province, Pakistan

<table>
<tr><td>Pillar 2: Increasing equitable access and improving quality and relevance</td><td colspan="3">Policy Actions and Benchmarks</td></tr>
<tr><td>Outcomes</td><td>Year 1</td><td>Year 2</td><td>Year3</td></tr>
<tr><td>1. Improved access to all levels of education and reduced gender disparity
Monitorable Indicators:
Overall
<table>
<tr><td></td><td>Baseline</td><td>2009</td></tr>
<tr><td>Primary NER (M,F)</td><td>58 (60;55)</td><td>66 (68;64)</td></tr>
<tr><td>Primary GER (M;F)</td><td>95 (100;89)</td><td>100%+</td></tr>
<tr><td>Rural Female NER (elem;secondary)</td><td>14/7</td><td>20/10</td></tr>
<tr><td>Primary Completion rate</td><td></td><td></td></tr>
</table></td>
<td>Provide missing facilities to primary schools adopting a full school approach

Initiate recruitment of 15,000 teachers based on enrolment loads of existing schools and new vacancies in upgraded middle and secondary schools

College sector improvement plan prepared and under implementation</td>
<td>Infrastructure program with full schools approach

Recruitment of additional teachers as per enrolment load and new vacancies</td>
<td>Continue implementing full schools approach</td></tr>
</table>

2. Multi-Purpose Surveys for Monitoring, Evaluation and Accountability

A large majority of countries regularly collect and process data on a range of crucial input and output indicators such as enrolment, staff numbers, and number of graduates etc., through data gathered in annual school censuses and maintained in an Education Management Information System (EMIS). Increasingly, countries are also implementing assessment systems to measure and track progress in learning achievement.

Exclusive reliance on administrative data, however, presents two main weaknesses. First, although administrative data is essential, it is often difficult to determine the degree of reliability and consistency of the data. Survey data, for example, often shows vastly different enrolment ratios than those computed with school census data. Second, informed policy requires an understanding of both the supply and demand-side constraints, that is an understanding of the priorities and problems of families whose children do not attend or complete primary school as well as of users' perceptions of the quality of services being provided.

Several types of multi-purpose surveys offer opportunities to provide a check on administrative data and information on the quality of the services being delivered. When institutionalized and used as a monitoring tool, these surveys can provide a mechanism to strengthen the accountability of service providers to beneficiaries.

Three types of surveys that have become popular over the past several years that aim to accomplish one or more of these objectives are Quality of Service Delivery Surveys (QSDS), Citizen's Report Card Surveys (CRCS), and Public Expenditure Tracking Surveys (PETS). All of these surveys share a common goal of providing feedback to policy makers on the performance of service providers. They have been used in the Education and Health sectors, among others, as:

- A *Diagnostic tool* to provide quantitative (inputs) and qualitative (perceptions although quantified) information about service delivery and comparative information by geographic area, and sub-groups to identify critical variations and excluded groups.
- An *Accountability tool* to help identify where institutions have not fulfilled their obligations
- A *Benchmarking tool* to track variations in service quality over time.

Quality of Service Delivery Survey (QSDS)

QSDS are micro-surveys of service providers, such as schools, that provide both a check on inputs and results, and capture various other dimensions of performance that are not assessed by administrative data. QSDSs can be comprehensive or targeted towards a specific question. A good example of the latter are studies that focus on measuring teacher absenteeism (Chaudhury et al 2003, Reinikka and Smith 2003).

Citizen Report Cards (CRCs)

Citizen Report Cards are participatory client feedback surveys that provide a quantitative measure of user perceptions of the quality, efficiency and adequacy of public services such as health, education, and public security (police) services. CRCs were first developed by an NGO – the Public Affairs Center -- in Bangalore, India. Like a student's school report card, the CRC aggregates scores across a variety of indicators in easily understood ranks. They are used in situations where demand side data such as user perceptions on quality and satisfaction with public services is absent. CRCs are used to provide information on where and how basic services are failing people, on key constraints people face in accessing public services, and on the reach and benefit of reform measures on the basis of perceptions of households. By systematically gathering and disseminating public feedback, citizens can collectively inform agencies about their performance and produce pressure for change.

CRCs typically involve the participation of stakeholders – providers and users – in the survey design through the use of focus groups where providers can provide input on their mandates and areas where feedback would be useful, and initial impressions can be sounded out among the users for further probing.

CRCs typically involve not only assessing citizen's perception of service quality but also solicit their advice on suggestions for improvement. Some of the actual applications include:

- In India – for state comparisons in access, use and satisfaction
- In Uganda -- to supplement national QSDS surveys
- In South Africa -- to compare the cost, quality and performance of different municipalities

In the education sector, the CRC approach has been used in Zanzibar (2 districts, 1,000 households) and Ethiopia (10 districts, 3,900 households) to:

- Verify access of households to school,
- Costs incurred by parents
- Citizens' perceptions of quality,
- Citizen's level of satisfaction with the services, and
- Major suggestions for improvement.

In Zanzibar, the CRC results showed high levels of enrolment and satisfaction with teacher behavior, but a high level of dissatisfaction with the availability and quality of learning materials. It also revealed that most parents pay fees, or make other contributions, of which 61 percent was reported to have been demanded by authorities. In Ethiopia, the sources of parental dissatisfaction with rural schools centered on poor management of attendance (47%), compulsory school uniforms (25%) and poor conduct of teachers (16%).

Some of the benefits that countries have cited in using CRCs are their simplicity, low cost and replicability which provides an efficient way to benchmark service quality and can help to tackle bureaucratic inertia by its effect on stimulating collective action for reform.

But experience also shows that the impact of CRCs has varied depending on the degree of participation, the ability of citizens to lobby for change, and the presence of a political champion. For CRCs to be effective, they require:

- Technical competence to undertake the survey. The survey work is generally outsourced to a market research agency, policy institute or NGO with adequate statistical and survey analysis skills. In South Africa, for example the Human Science Research Council is undertaking the surveys. In Ethiopia, they are conduced by the Poverty Action Network of Civil Society. A World Bank Social Development Note (2004) provides a summary of the process involved.
- Advocacy to bring the results into the public domain, often NGO advocacy
- Steps aimed at institutionalizing the practice

However, whether or not all these elements are in place, client feedback surveys can be useful in guiding the policy priorities and informing strategies that address demand side constraints.

Public Expenditure Tracking Surveys

PETs are a category of quantitative exercises that enable the monitoring of how public funds are spent by tracing the resource flow from its origin to final destination and determining the location and scale of any anomalies. They can also be used to track inputs delivered centrally such as textbooks or teachers. PETS in the education sector have been conducted in numerous countries including Ghana, Papua New Guinea, Peru, Rwanda, Tanzania, Zambia, Cambodia and Uganda.

The first and most well-known PETs was conduced in Uganda. The idea was initiated by a puzzling situation – the Government of Uganda had been steadily increasing expenditure on education since the late 1980s, but primary enrolments stagnated. To obtain a better understanding of why there appeared to be so little impact of the increased expenditure, it was first necessary to determine whether the resources were in fact reaching their intended destination. The results showed that between 1991-1995, 87 percent of the annual per student grant allocated to district officials did not reach the schools. Fewer than half the schools received any funding at all. The Government responded by publicizing the amounts transferred to districts, requiring officials to post the information at the district offices and schools, and training communities in how to use the information to hold local authorities accountable. In 1999, the school survey was replicated. The results showed a dramatic improvement with schools receiving 90 percent of the capitation grant.

Results from a PETS in Zambia also documented leakage on a massive scale in non-wage expenditure. The PETs found that funds allocated to schools at the discretion of local authorities had less chance of reaching them than a formula based funding mechanism.

Like all surveys, however, the information is only useful if the results are widely known and acted upon as was the case in Uganda. Tanzania provides a counter example where three PETs were undertaken – in 1999, 2001 and 2003. The first survey suggested that 57 percent of funds allocated for other charges (non-wage) were diverted. It also gave important insights into governance and accountability issues at the district and service center levels. Sundet (2004) reports that while the results of the first two surveys had a positive impact on making the advertisement of transfers from the central treasury to local councils routine, they have not served to motivate and sustain a debate on improving governance and accountability.

As illustrated, PET surveys can be highly useful tools to monitor if resources allocated are reaching their destination, as well as provide important insights into financial management systems and structures of accountability. PETS, however, are complicated to design and administer, and the larger the number of budget categories or inputs to track, the more demanding the exercise becomes. In the interest of keeping the exercise manageable, countries would be best advised to limit the scope to one or two critical flows or research questions. Further, a PETS that aims to investigate financial leakage may not yield useful results when the PFM rules and practices are designed in such a way that resources cannot be tracked.

One important aspect that all three survey methods – QSDS, CRC, and PETS -- have in common is that they are undertaken by third parties, not the MoE. This is necessary to eliminate MoE stakeholders as a potential source of bias. It is also helps to build local capacity to conduct surveys, a capacity that would be difficult for MoEs to develop in-house.

2. Donor Partnerships

Annual Reviews

Annual reviews are the main process by which countries and donors come together to review progress towards the achievement of the education sector plan's goal. Several studies over the past few years on the strengths and weakness of donor partnerships at the country level point to the following key challenges in the annual review process (ODI 2000, FTI, OECD 2006):

- Annual reviews are not sufficiently organized, and preparatory documents are not available upstream
- Reviews tend to focus on implementation progress and lose sight of the medium term perspective.
- The tendency of many reviews is to emphasize that they are joint reviews which may restrain critical views
- Donors sometimes see their role as an opportunity to have their interests represented and exercise a lack of discipline and prioritization in recommendations.

Countries have used a range of strategies to address these issues ranging from the simpler actions to more complex agreements – from improving the availability of information, to more formal agreements stipulated in MoUs or Joint Finance agreements.

Availability of information. To improve the availability of information, a number of countries have invested in posting all relevant documents and data on a public MoE web site (Cambodia). This helps to improve transparency for all stakeholders, the quality of reviews, and reduces demands on MoE for information.

Strategies to Make (annual) reviews more effective.

- Introduce bi-annual reviews one focused on policy, strategy and results, the other on financing rather than one annual review. In fact, this is generally now the norm across countries.

- Shift from reporting on progress in implementing work programs to seeking technical value added of donors on specific topics, policies, strategies or constraints to avoid the development of an atmosphere that places MoE in a hot seat, reporting and fielding questions only with little value added of the meeting to the MoE.

- Utilize third party validations of MoE monitoring and evaluation reports so that donors are more fully prepared and there is some independent check on the monitoring and results information being provided. It can also help to balance the partnership when donors are perceived to exercise an undue influence on policy choices, provided that the external evaluation is not beholden to either the Government or the donors. Third party evaluations of the partnership between the Government and donors have also been utilized to provide outside insights on perceptions, strengths and weaknesses.

- Donors coordinate themselves internally to provide coherent information on their financing and a consensual validation of MoE reports.

MOUs, Partnership Declarations. Since the development of the country plans, a number of countries have continued to develop the donor partnership through MoUs such as Honduras and Yemen. MoUs attempt to establish principles and procedures to which the country and donor partners are committed to endeavor to achieve. Although MoUs share some common features, they are highly context specific in content depending on the stage of the partnership, and issues that the partners are working to address. Common features include mutual commitment to the modalities of financial and TA support, progress reviews, monitoring indicators, and MoE coordination of the process. However, the nature and specificity of the commitments varies considerably.

The table below shows excerpts from three MoUs in the education sector which illustrate the variety. In Yemen, the purpose of the MoU is to guide the transition from projects to joint donor-support of a common program for basic education utilizing Government procedures. In Honduras, the purpose of the MoU is to have a common understanding among parties on general principles and procedures for support to, and implementation of, the country's EFA plan. These different purposes reflect the stage of development of the country's strategic plan for EFA-UPC, the stage of development towards donor harmonization, and specific country issues. Thus, although models of MoUs exist, based on the Paris Agreement indicators, their contextualization is highly dependent on country circumstances.

Honduras	Yemen	Uganda
• The EFA plan will be managed as a single program • Each partner may utilize its own mechanism to provide support for the EFA Plan • Partners will endeavor to use a common framework of expected results and indicators for the EFA plan • Partners will endeavor to exchange information on various aspects of the EFA plan • Honduras will develop and implement a M&E strategy • Honduras and partners will actively encourage consensus building • Over time, cooperating partners should move to consolidate some functions of their PCUs within the ST-EFA	• Make available to other partners all information on all relevant interventions and work to ensure there is no overlap • Contribute to the development and implementation of a unified monitoring systems for BEDS and PRSP objectives and use the system as the main source of education information • Establish a joint annual review cycle • Provide schools with workable budgets • Promote the optimal use of technical advisory services by sharing information and terms of reference	• Funding agencies will endeavor to provide three year commitments which will be integrated into the MTEF • GoU to be more proactive in identifying where it wants TA support and managing the process • TA should be primarily for institutional capacity building. Specific sectoral TA would be expected to draw on GOUs own resources and the budget support (channeled through (a consolidated TA Fund) as its first funding source. • Reviews will monitor critical undertakings (financial commitment of Gov, public expenditure management and quality) and process undertakings which focuses on strategic areas where progress is required

A main benefit of MoUs is that they use international agreements to advantage (such as the Paris Declaration) thereby facilitating the process for all countries.

Experience shows that MoUs can raise certain sensibilities that need to be managed. Three key principles can help to smooth the way:

- Avoid exclusion – Earlier experience with "like minded" donors forming a group for pooled funding often led to a feeling of exclusion among the donors unable to participate in the pooled fund, although they may have provided financing through other channels. All donors can still work within a common results framework. This implies respect for donor financing procedures although respect should not be used as an excuse for the donor not to harmonize nor for a donor to expand the country's results framework.

- Maintain Commitment but Informality -- MoUs are moral rather than legal agreements. Donor individual requirements will always take precedence over the conditions set forth in a MoU. Understanding that at the outset, can help to manage expectations.

- Be clear about the value added. In countries with few partners, MoUs may not add value. In some other situations, the initiative for MoUs has mainly come from donors, so MoUs may often appear one-sided in that they do not commit the donors to the levels, timing or channels by which they would commit funds. Therefore, there may not be much interest on the part of government (Ethiopia education).

III. Discussion Questions

- Reflecting on your own education sector plan, what would you say are the strengths of the plan and areas for improvement?

- What tools and processes have you applied to:
    - Prioritize
    - Improve information on demand-side constraints and the quality of services being provided at the school level
    - Engage civil society in plan preparation and implementation
    - Coordinate donors

- What constraints did you encounter and how did you address them?

ANNEXES

ANNEX TABLE 1. COUNTRY PLANS INCLUDED IN THE REVIEW

Albania	Guinea	Mongolia
Burkina Faso	Honduras	Mozambique
Cambodia	Kenya	Niger
Cameroon	Kyrgyzstan	Rwanda
Djibouti	Lesotho	Senegal
The Gambia	Madagascar	Tajikistan
Ghana	Mali	Timor Leste
	Mauritania	Vietnam
	Moldova	Yemen

ANNEX TABLE 3: CRITERIA APPLIED IN THE ASSESSMENT OF COUNTRY PLANS

Technical Aspects	Financial and Economic	Institutional & Implementation	Leadership/ Ownership	Monitoring & Evaluation/Accountability Framework	Fiduciary aspects	Environmental
Has a long term strategic direction for the entire sector	Has a medium-term financial framework based on projected domestic and external resources	Has an implementation plan	Evidence of country ownership & political support for the reforms included in the program	Annual targets for measuring progress on key policies and outcomes	Financial Management – draws on relevant analyses of the country's and sector's public financial management	Quality of environmental impact and risk assessment
Has a medium term plan for the whole sector	The financial plan is embedded in the country's macro framework and MTEF	Plans to have annual implementation plans	Arrangements for consultations and participation in formulating the strategy	Monitoring and evaluation arrangements are clear	Consideration of financial management arrangements and capacity	Arrangements to ensure the program will have no negative impacts on the environment
Has a long term direction for UPC	Adequacy and realism of the financing plan as determined by donor appraisal	An assessment of capacity to implement the program	Civil society participation	Arrangements by which donors will review progress and validate monitoring and evaluation findings	Extent to which the program draws on analysis in the Country Procurement Assessment Report or other relevant documents.	
Has Medium term plan (3-5 years) to achieve UPC		Strategies are proposed to address implementation capacity issues	Describes how donors will work together		Adequacy and quality of procurement arrangements and capacity	
The UPC plan is costed					Funding modalities considered	
Used a simulation model to assess financial impact of policies and trade-offs						
Strategy and interventions are prioritized						
Adequacy of country and sector knowledge underpinning the strategy						
Adequacy of basic data						

ANNEX TABLE 3: CRITERIA APPLIED IN THE ASSESSMENT OF COUNTRY PLANS (CONTINUED)

Analysis of high priority issues						
Regional disparities						
Gender disparities						
Poverty disparities						
HIV						
Orphans						
Children with disabilities						
The medium-term plan has clear strategies:						
- Primary completion						
- Learning outcomes						
- Education quality (inputs)						
- Gender equality						
- Internal Efficiency						
- HIV/AIDs						
- Demands at other levels of education						
- Poverty targeting						
Plan includes benchmarks for Indicative framework indicators						
Lessons are incorporated from past experience						

FTI Country Plans – are available on the FTI web site at:
http://www.fasttrackinitiative.org/content.asp?contentid=529

Impact Evaluation

World Bank Impact Evaluation Web Site
http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/EXTISPMA/0,,menuPK:384336~pagePK:149018~piPK:149093~theSitePK:384329,00.html

Duflo, E. et al. 2007. "Randomized Evaluations of Interventions in Social Science Delivery." Washington, DC: World Bank Institute.

Glewwe, P. et al. 2000. Retrospective Vs. Prospective Analyses of School Inputs: The Case of Flip Charts in Kenya. Cambridge, MA: National Bureau of Economic Research, Working Paper 8018.

Ravallion, Martin. The Mystery of the Vanishing Benefits: An Introduction to Impact Evaluation. 2001. World Bank Economic Review, Vol. 15, No 1, 115:140.

Citizen Report Cards

Asian Development Bank. 2005. *Improving Local Governance and Service Delivery: Citizen Report Card Tool Kit.* Manila.

Bekele, Eshetu. 2006. *The Case of Citizen Report Cards in Ethiopia*. Paper presented at the HDCA 2006 Annual Meeting.

Paul, Samuel. 1998. *Making Voice Work: The Report Card on Bangalore's Public Service.* Bangalore: Public Affairs Center.
http://www-wds.worldbank.org/external/default/main?pagePK=64193027&piPK=64187937&theSitePK=523679&menuPK=64187510&searchMenuPK=64187283&theSitePK=523679&entityID=000009265_3980625102855&searchMenuPK=64187283&theSitePK=523679

Paul, Samuel. 2004. *Holding the State to Account though Citizen Report Cards in India.* Washington, DC: World Bank, Working Paper no. 30824.
http://www-wds.worldbank.org/external/default/main?pagePK=64193027&piPK=64187937&theSitePK=523679&menuPK=64187510&searchMenuPK=64187283&theSitePK=523679&entityID=000090341_20041227134108&searchMenuPK=64187283&theSitePK=523679

World Bank 2004. *Citizen Report Card Surveys.* Washington, DC: Social Development Notes Participation & Civic Engagement. February

World Bank. 2004. *Citizen Report Card Surveys – A Note on Concept and Methodology*. Social Development Note No 91, February. Washington, DC:

QSDS and Public Expenditure Tracking Surveys

Ablo, Emmanuel and Ritva Reinikka. 1998. *Do Budgets Matter? Evidence form public spending on education and health in Uganda*, Washington, DC: World Bank, Policy Research Worlking Paper no. 1926.
http://www-wds.worldbank.org/external/default/main?pagePK=64193027&piPK=64187937&theSitePK=523679&menuPK=64187510&searchMenuPK=64187283&theSitePK=523679&entityID=000009265_3980702115711&searchMenuPK=64187283&theSitePK=523679

Chaudhury, N. et al. 2003. Teacher and health care provider absenteeism : a multi-country study. Washington, DC: World Bank.

Chaudhury, N. et al. 2005. Missing in Action:Teacher and Health Worker Absence in Developing Countries. Washington, DC: World Bank
http://econ.worldbank.org/external/default/main?theSitePK=477916&contentMDK=20562060&menuPK=546432&pagePK=64168182&piPK=64168060
The above site provides studies from numerous countries.

Dehn, J. Ritva Reinikka and Jakob Svensson. *Survey Tools for Assessing Performance in Service Delivery*. World Bank.
www1.worldbank.org/publicsector/pe/PETS1.pdf

PETS in the Education Sector – Uganda, Tanzania and Zambia – The Anti-Corruption Research Center.
http://www.u4.no/themes/pets/petseducationsector.cfm

Reinikka, Ritva and Jakob Svensson. 2002. *Assessing Frontline Service Delivery,* Washington, DC: World Bank, Development Reserach Group.
www.worldbank.org/wbi/governance/assessing/pdf/reinikka.pdf

Sundet, Geir. 2004. Public Expenditure and Service Delivery Monitoring in Tanzania: Some International Best Practices and a Discussion of Present and Planned Tanzanian Initiatives. Dar es Salaam: HakiElimu Working Paper No. 04.7.
www.hakielimu.org/WP/WPSeries7_2004.pdf

Donor Coordination

Links to various reports
http://www.aidharmonization.org/

Links to signed MoUs
http://www.aidharmonization.org/ah-cla/secondary-pages/cla-actionCountryMOU

Foster, M., A. Brown, A. Norton, F. Naschold. 2000. *The Status of Sector Wide Approaches*, Overseas Development Institute for Ireland Aid, February.

OECD 2006. *The Challenge of Capacity Development Working Towards Good Practice*. Paris

Workshop on the Paris Declaration: Implications and Implementation, Bamako Mali March 27-29, 2006.