

Mid-Term Evaluation of the EFA Fast Track Initiative

Final Synthesis Report
Volume 5 – Appendices VI–VIII

APPENDIX VI
M&E FRAMEWORK AND IMPACT EVALUATION

February 2010



COMPONENTS OF THE SYNTHESIS REPORT

VOLUME 1: MAIN REPORT

Executive Summary

Main Report

VOLUME 2: ANNEXES¹ AND BIBLIOGRAPHY

Annex A Timeline

Annex B Governance and Evolution of the FTI

Annex C The Indicative Framework and UPC

Annex D Summaries of Country Studies

Annex F Monitoring and Evaluation of the FTI

Annex G The FTI and Capacity Development

Annex H The FTI and Fragile States

Annex J Interviewee Analysis

Annex K Financing Gaps

Full Bibliography

VOLUME 3: APPENDICES I–IV

Appendix I Statistical Appendix

Appendix II Trends in Financing of Basic Education

Appendix III The Catalytic Fund

Appendix IV The Education Program Development Fund

VOLUME 4: APPENDIX V

Appendix V Note on Approach and Methods

VOLUME 5: APPENDICES VI–VIII

Appendix VI M&E Framework and Impact Evaluation

Appendix VII Reforming the Catalytic Fund

Appendix VIII Findings against High Level Questions

¹ What was originally Annex E was later redesignated as Appendix VIII. Similarly Appendix VII was formerly designated Annex I. To minimise confusion, remaining annexes have kept their original letters.

CONTENTS

Acronyms and Abbreviations	iv
A. Introduction	1
B. Developing an M&E framework for the FTI	1
Existence of an FTI M&E framework	1
Proposed M&E framework	4
Next steps	12
C. Issues in impact evaluation	13
D. Options	15
E. Summary	16
References	17

Acronyms and Abbreviations

3ie	International Initiative for Impact Evaluation
AIDS	Auto-Immune Deficiency Syndrome
CAP	Consolidated Appeals Process
CRS	Creditor Reporting System
DAC	Development Assistance Committee
DQAF	Data Quality Assessment Framework
EFA	Education for All
EPDF	Education Program Development Fund
ESP	Education Sector Plan
FTI	Fast Track Initiative
GFATM	Global Fund to fight AIDS, Tuberculosis and Malaria
GMR	Global Monitoring Report
GRPP	Global and Regional Partnership Programmes
HLQ	High Level Question
IASC	Inter-Agency Standing Committee
IFS	International Financial Statistics
IMF	International Monetary Fund
IPA	Innovations for Poverty Action
JAM	Joint Assessment Mission
J-PAL	Abdul Latif Jameel Poverty Action Lab
M&E	Monitoring AND Evaluation
MTBF	Medium Term Budget Framework
MTEF	Medium Term Expenditure Framework
ODA	Official Development Assistance
OECD	Organisation for Economic Cooperation and Development
PEFA	Public Expenditure and Financial Accountability
PFM	Public Finance Management
PRS	Poverty Reduction Strategy
TOR	Terms of Reference
UIS	UNESCO Institute of Statistics
UNESCO	United Nations Educational, Scientific and Cultural Organisation
UNPD	United Nations Population Division
UPC	Universal Primary Completion

A. Introduction

1. In recent years and particularly since the Paris Declaration on Aid Effectiveness, development partners have been strongly emphasising the need to improve the performance of aid programmes defined in terms of development results. Indeed, impact is one of the five DAC criteria for evaluating development assistance.

2. In line with this orientation, the terms of reference of the FTI evaluation specified that "*a comprehensive design for FTI monitoring and evaluation, including an evaluation of impact and relevant baseline data, should be one of this evaluation's outputs*" (TOR ¶13). This being primarily a formative evaluation, it has not been expected that it would fully evaluate impact; instead it would "*help frame future monitoring and evaluation efforts*" (TOR ¶15) and "*provide guidance for the assessment of the impact of the FTI at a later stage*" (TOR ¶17)].

3. This is also in line with the GRPP Sourcebook (IEG 2007) which included among its quality standards for global programme evaluations an examination and critique of the existing monitoring and evaluation framework.

4. The issue of impact evaluation has been approached from the following angles:

- The evaluation team developed its own evaluation framework (Cambridge Education, Mokoro & OPM 2009), a detailed methodology that was circulated to the FTI partnership in January 2009. The evaluation framework document articulated a programme theory for each goal given the absence of a relevant formal FTI document.
- The challenges facing the FTI partnership on M&E were reviewed (**Section 3F of the main report, and Annex F in Volume 2**) particularly from the point of view of the M&E practices followed by the FTI partnership for management purposes.
- A study was commissioned from 3ie (White 2009) as part of this evaluation and completed in March 2009 to address methodological issues and possibilities for a final impact evaluation of the FTI.

5. This appendix responds to the relevant parts of the terms of reference on M&E and impact evaluation. Section B discusses the general issue of an FTI monitoring and evaluation framework and proposes a framework for discussion. Section C provides an overview of the impact evaluation issues. Section D presents the options available. Finally, Section E concludes.

B. Developing an M&E framework for the FTI

6. This section:

- (a) stresses the importance of an M&E framework and describes the implications that the absence of such a framework has had for the FTI partnership; and
- (b) outlines a potential structure for such a framework to serve as a basis for discussion among the partners.

Existence of an FTI M&E framework

7. This evaluation has been asked to answer, among other questions, whether the FTI had "*an adequate monitoring and evaluation framework*" that has been "*creating regular, useful inputs to decision making*" (HLQ 3h).

8. Indeed, according to the GRPP Sourcebook:

A key enabling condition for effective evaluation is the early establishment of an M&E framework. Early after its launch, each GRPP should put in place an M&E framework, at least at the program level, which includes:

- (a) clear and coherent objectives and strategies
- (b) an expected results chain
- (c) measurable indicators that meet the monitoring and reporting needs of the governing body and management, and
- (d) systematic and regular processes for collecting and managing data, including baseline data. (IEG 2007 ¶2.6)

9. While some of these elements have been in place, important elements have been missing to the extent that it can be argued that the FTI has lacked an M&E framework. Taking each of the above points in turn:

- (a) The *objectives* were specified in the 2004 FTI Framework (FTI 2004) as high-level *goals* underpinned by *principles*. However, the *strategies* to be followed in order to achieve each of these goals were often unclear. For example, in the case of the data gap, at the global level, the FTI would partner "*with UNESCO to support the continued development of appropriate monitoring indicators*" according to the FTI Framework. However, it is not clear how this modest activity would lead to the achievement of results specified in the FTI Framework such as "*transparent annual monitoring of indicative framework ... targets*". At the country level, no FTI-specific activity was proposed in the FTI Framework. In many instances, the FTI was simply expected to encourage partners to engage in certain activities (e.g. according to the FTI Framework the FTI "*encourages the open sharing of information on the policies and practices of participating countries and donors alike*"). When necessary activities are vague and left to the discretion of the partners, then partners cannot be held to account if they did not implement them.
- (b) The FTI has not had a fully worked out and agreed *results chain* theory linking the available inputs and activities to the desirable outputs and outcomes to guide performance M&E. By contrast, evaluations of high-profile partnerships have relied on analyses of mechanisms of influence. For example:
 - Stern et al considered whether the assumptions linking the implementation of the Paris Declaration with its possible outputs, outcomes and impacts were plausible. Drawing on general theory and context-specific situations, a set of propositions were proposed as indicative lines of enquiry for the next phase of the evaluation. (Stern et al 2008)
 - Bennett and Fairbank presented a conceptual framework that identified the channels through which the Global Fund to Fight AIDS, Tuberculosis and Malaria (GFATM) affects health systems and health outcomes in order to provide a basis for GFATM M&E. (Bennett and Fairbank 2003)
- (c) For FTI, there have been no *indicators* of success or at least no indicators that could be plausibly linked to the role of the FTI partnership. The specification of high-level outcome indicators, such as the Indicative Framework indicators, could not be a substitute as the FTI represents only part of the overall efforts towards UPC and judging the FTI solely from the achievement of these indicators would be misplaced.

(d) Accordingly, there has been no thorough description of the *data collection processes*. For example, there was an implicit commitment to collect information on "donor performance" upon which to base regular reports. However, the data from the main source (DAC CRS) are often disputed by individual donors; yet, there has been no attempt to reconcile them with alternative sources. In practice, information on donor performance has not been collected systematically and has not been tied to follow-up on specific donor performance commitments.

10. The GRPP Sourcebook added further that in cases "*where the objectives and strategies have not been well articulated ... evaluators will need to construct a clear and agreed-upon statement of the objectives and strategies in consultation with the governing body (or oversight subcommittee or external panel) that is overseeing the evaluation*" (IEG 2007 ¶9.5).

11. It is for this purpose that, within the scope of the evaluation framework document and in consultation with the Evaluation Oversight Committee, the evaluation team developed a *logical framework* (Cambridge Education, Mokoro & OPM 2009). At a theoretical level, the framework set out a clear set of hypotheses to be tested. At a practical level, it listed relevant detailed evaluation questions, together with main sources of evidence needed to answer them – and provided a focus for the enquiries of each workstream.

12. The framework followed a standard structure that proceeded through successive levels for each one of the five streams of hypothesised channels of influence: policy and planning; finance; data and M&E; capacity; and aid effectiveness.²

- Level zero captured the context at the global and country level before the FTI partnership was established.
- Level one listed the inputs provided and activities undertaken as part of the FTI.
- Level two considered the immediate effects (short-term outputs) that would be expected on the education sector processes.
- Level three envisaged intermediate effects (medium- to long-term outputs) such as changes in sector policy, expenditure and service delivery.
- Level four set out the long-term expected outcomes on quantity, quality, access and sustainability of primary education.
- Level five was reserved for the long term "personal, institutional, economic and social" effects of expanded primary education

13. The logical sequence was couched in the context of external factors and assumptions that would need to be checked to confirm whether different stakeholders and different countries shared understanding of FTI objectives and functions. If not, then the logical sequence might have broken down with consequences for the assessment of the FTI's efficiency, effectiveness and sustainability.

14. This logical framework looked backwards trying to interpret the actions of the FTI partners. However, what the partnership needs is a forward looking logical framework on which to base a future evaluation. This is also in accordance with the terms of reference, which requested that the new framework would be "*building on the evaluation framework created for this evaluation*" (TOR ¶15).

² See this report's Appendix V (Note on Approach and Methods), which provides additional discussion and reproduces in full the extended logical framework for the mid-term evaluation.

Proposed M&E framework

15. The proposed monitoring and evaluation framework is structured like a conventional logical framework. It consists of the following elements:

- Narrative statements of goal/impact, purpose/outcomes, outputs and activities.
- Indicators and sources (implicitly also addressing data collection processes).
- Baselines, milestones and targets.
- Assumptions.

16. The *narrative statements* have been developed on the following basis:

- The *goal/impact* and *purpose/outcome* level statements are taken from the January 2009 evaluation framework of the evaluation team (Cambridge Education, Mokoro & OPM 2009).
- The *output* level statements take the *Goals* of the 2004 FTI Framework (FTI 2004) as their point of departure: (i) more efficient aid for primary education; (ii) sustained increases in aid for primary education; (iii) sound sector policies in education; (iv) adequate and sustainable domestic financing for education; (v) increased accountability for sector results; and (vi) mutual learning on what works to improve primary education outcomes and advance EFA goals. These goals have been combined into narrative statements of output as follows:
 - Goal (i) corresponds to an aid effectiveness output (Output 4)
 - Goals (ii) and (iv) correspond to a finance output (Output 2)
 - Goal (iii) corresponds to a policy and planning output (Output 1)
 - Goals (v) and (vi) correspond to a data and M&E output (Output 3).³
- The activity level statements are an attempt to pin down those activities that the partnership will directly undertake and will be monitored against. These will need to be carefully scrutinised by the partnership in view of further governance changes, changes in key elements of FTI strategy, and any decisions made on the role of the FTI Secretariat.

17. With respect to *indicators*:

- (a) The main emphasis of the proposed framework is on the indicators at the activity and output level, where indicators have been for the most part absent, a fact that has blurred issues of accountability.
- (b) At the purpose/outcome level, the indicators should be selected among the EFA goals, notably the primary completion rate. Although the list could be expanded to include some "service delivery" indicators of the current or revised Indicative Framework, it would be preferable to keep this list short.
- (c) At the goal/impact level, no indicators have been selected. The ultimate goal of universal primary education is a better quality of life. As the G8 Education Experts Group Report (an annex to the July 2009 L'Aquila G8 Declaration) states, "*education is a powerful tool, enabling communities to cope with major environmental and economic challenges and maintaining social cohesion and stability*" (G8 2009). However, the impact of education is complex and inherently difficult to measure. In particular, estimates of the impact of education (for even standard research such as growth accounting) depend on the availability of particular data series and the methodological approach in specifying the empirical

³ Note that although "capacity" appears as one of the four "gaps" there is no separate capacity development goal or output; capacity development is expected to be mainstreamed across all activities.

relationship between education and its potential effects (which tends to be endogenous). Instead of specifying and monitoring goal/impact indicators, the FTI partners could establish an initiative (see ¶42) that will compile results of research studies on the long-term impact of education on personal, institutional, economic and social aspects of life (notably health and fertility) in developing countries to inform policy.

18. In terms of setting *baselines, milestones and targets*, the evaluation team is faced with some constraints. According to the terms of reference, the framework to be proposed by this evaluation "should indicate baseline data needed for the evaluation of FTI impact" (TOR ¶15). However, the choice of *baselines* depends on the choice of indicators, which in turn must be owned by the partnership. Nevertheless, the proposed indicators have been chosen carefully to ensure that they can be monitored either through existing systems or through reasonably straightforward extensions of such systems. A different approach has been followed in defining activity-level *milestones*: these resemble the Progressive Framework (used by the FTI to describe its approach to fragile states see FTI 2007) and recognise that, as countries are at different stages, the focus of FTI activities will have to adjust to the needs of partner countries. Finally, targets and their level of ambition need to be set by the FTI partnership.

19. The *assumptions* are an adjusted version of the assumptions that were listed in the January 2009 evaluation framework. These will need to be monitored on a regular basis in order to provide confirmation that they remain valid.

20. Several important points on the proposed M&E framework need to be stressed:

- (a) The framework is being proposed as a *basis for discussion* among FTI partners. It is meant to be reviewed extensively and amended by the partners so that it better reflects what the partnership believes it should achieve. The proposed framework should not be seen as a framework that should be adopted as is. To facilitate the discussion, a number of possible decisions that need to be made by partners in the selection of indicators are suggested in footnotes but these are merely indicative and by no means exhaustive.
- (b) The framework *focuses on the monitoring of the FTI partnership* (with a particular emphasis on its activities and outputs). Although it includes a small number of key education sector indicators at the goal and purpose level, *it is not a framework for monitoring country EFA performance*.
- (c) It is essential to monitor indicators for *FTI eligible countries*, not just countries which have been endorsed by the FTI; data confined to the latter group cannot yield robust information on the value-added of the FTI.
- (d) The monitoring and evaluation of the FTI takes place at two levels:
 - The proposed framework focuses on *global-level* M&E. It assumes that the FTI Secretariat would be responsible for its coordination.
 - *Country-level* M&E is also needed to focus on aspects of the compact which are harder to observe. For example, the monitoring of the donor and country commitments to finance an education sector plan requires specific country knowledge. However, the nuances of a country-specific M&E framework need to be addressed separately and are not the object of the following matrix.

21. In summary, the M&E framework that is being proposed as a basis of discussion among FTI partners has been designed to be (i) commensurate with its span of control; (ii) aligned with its focus; and (iii) within the limits of its budget.

22. While effort has been made to include indicators that are feasible to measure, in several cases extra effort will be needed. Three different types of increased effort that might be called for are mentioned below as examples:

- (a) The FTI undertook an assessment of the quality of ESP documents (FTI 2008) and it is proposed that this initiative is continued. However, transparent criteria for the assessment need to be captured clearly in a formal document (and might also need to be extended to cover other areas of interest). Moreover, the criteria will need to be compatible with country specific conditions.⁴
- (b) There is a data quality assessment framework (DQAF) that has been adapted to education statistics by the World Bank and UIS. It is proposed that this is applied (partly or in its entirety) to assess the quality of education data collection processes. However, this will call for a programme and the allocation of resources by the FTI.
- (c) The current quality of some data may not be entirely satisfactory (e.g. public recurrent spending on education as percentage of total public recurrent discretionary spending). This will require the FTI to cooperate closely and actively with key institutions (e.g. UIS or other) to ensure that the quality of the data is improved, and may call for a formal agreement of cooperation. In fact, it highlights the fact that the active involvement of UIS, GMR and DAC in the finalisation of the M&E framework should be considered. The proposed framework specifies the likely sources of each indicator but the data collection protocol for each of these indicators (and the direct responsibility of the FTI Secretariat itself) should be clearly specified.

⁴ The work of the Public Expenditure and Financial Accountability (PEFA) initiative provides a relevant pattern for standardising such assessments.

Proposed FTI partnership M&E framework

GOAL / IMPACT	Indicator	Baseline [200X]	Milestone 1	Milestone 2	Target [201X]
Long-term benefits from universal primary completion derived	[To be defined]				
		Source			

PURPOSE / OUTCOMES	Indicator	Baseline [200X]	Milestone 1	Milestone 2	Target [201X]
Universal primary completion / EFA goals achieved	[UPC 1] Primary completion rate by gender by FTI eligible country 1. Estimated on the basis of administrative data 2. Estimated on the basis of nationally representative household survey data for the cohort of 15-19 year olds				
	Source	1. UIS / United Nations Population Division 2. Education Policy and Data Center			
	Indicator	Baseline [200X]	Milestone 1	Milestone 2	Target [201X]
	[UPC 2] Number of out-of-school children by FTI eligible country 1. Estimated on the basis of administrative / population census data 2. Estimated on the basis of nationally representative household survey data for the cohort of 6-14 year olds ¹				
	Source	1. UIS / United Nations Population Division 2. Education Policy and Data Center			
	Indicator	Baseline [200X]	Milestone 1	Milestone 2	Target [201X]
Indicator of learning outcomes by FTI eligible country ² 1. Early grade: proportion of students who, after two years of primary schooling, demonstrate sufficient reading fluency and comprehension to "read to learn" ["School children are able to read with comprehension a country-defined number of words per minute by the end of second grade"] 2. End of cycle: proportion of students who are able to read with comprehension, according to their countries' curricular goals, by the end of primary school.	Source	To be specified The FTI partnership will promote systematically the implementation of national assessments of reading skills			

1. But such data must not be used uncritically, since they are likely systematically to underestimate out-of-school children (see commentary on data issue in [Appendix I](#)).

2. The indicators are taken from the memo *Addition of learning outcomes indicators* issued by the FTI Secretariat in July 2009.

OUTPUT 1	Indicator	Baseline [2007]	Milestone	Target [201X]	ASSUMPTIONS
[EDUCATION POLICY AND PLANNING] Sound sector policies in education developed	Quality of education sector plans in FTI eligible countries	[FTI review of 28 countries]			<input type="checkbox"/> Political stability <input type="checkbox"/> National elites support EFA <input type="checkbox"/> Governance committed to development results <input type="checkbox"/> Accommodating public service organisational and pay structure <input type="checkbox"/> Medium-term budget framework implemented
		Source			
		Adapted FTI scoring formula ¹			
	Quality of education sector plans in FTI eligible countries in terms of coverage of crosscutting issues (gender, HIV/AIDS, disability, equity)	Baseline [200X]	Milestone	Target [201X]	
		Source			
		New FTI scoring formula ²			
	Variation between education sector plan allocation and actual budget expenditure in FTI eligible countries ³ 1. Total expenditure 2. Non-salary recurrent expenditure	Baseline [200X]	Milestone	Target [201X]	
		Source			
		Joint Annual Reviews			

1. Extend and apply the methodology of the 2007 FTI ESP review to plans of all FTI eligible countries (ensuring that the methodology is clearly documented and understood). The formula scores four sections (planning; financing; data; capacity) along three dimensions (clarity, strategy, and monitoring). Consider extending the review to also cover aspects such as inclusivity/transparency in plan preparation.

2. Extend the above methodology in a straightforward way to assess cross-cutting issues.

3. The indicator could alternatively: (1) assess variation between MTEF allocation and actual budget spending (if an MTEF exists); (2) focus on primary and/or basic (rather than total) education spending.

Activity	Indicator	Low	Middle	High
1.1 Advocate UPC/EFA at the global and country level with particular emphasis on fragile states	Number of country visits by FTI staff	Visit focus: introduce FTI concepts and process	Visit focus: review plan implementation	Visit focus: support local donor groups
1.2 Advocate preparation of ESP in all FTI eligible countries	Number of FTI eligible countries with endorsed sector-wide ESP	Sub-sector plans	Sector-wide plans	Sector-wide plans embedded in PRS/MTBF
1.3 Implement education sector plan endorsement process with FTI appraisal guidelines	Percentage of appraisals completed within x months from submission of the ESP	Endorsement appraisal external to countries	Endorsement appraisal involves local stakeholders	Endorsement process is part of country processes
1.4 Advise, monitor and report on the results of the joint annual review processes with an emphasis on the use of their results for education policy decision making.	Number of FTI eligible countries whose education sector performance reports are posted on the FTI website.	No annual education sector performance report	Annual education sector performance report approved by stakeholders	Recommendations of report followed up and acted upon
1.5 Ensure that there is a capacity development strategy linked to the sector plan and FTI capacity development guidelines, report on results and update on recommendations	Number of capacity development strategies prepared (as organic part of the ESP) in past three years	No capacity development strategy	Capacity development strategy covers education sector	Capacity devt strategy links to broader PFM and public service reform
1.6 Use the EPDF to "strengthen country capacity to develop and effectively implement good education policies, robust results-oriented programmes and realistic budgets"	Volume of resources allocated to EPDF (new) Objective 1	Decisions on use of funds involve local stakeholders and are related to the ESP	Monitoring systems are used to assess EPDF funds effectiveness	Feedback from monitoring improves alignment of EPDF projects with ESP
1.7 Use the EPDF to "reinforce the role of Local Education Groups in coordinating support to the education sector, engaging and aligning a broad constituency, and advancing country leadership and accountability" ¹	Volume of resources allocated to EPDF (new) Objective 3	Ad hoc community involvement in education planning	Coordinated community involvement in education planning	Civil society involvement in and education planning and accountability

1. Use the *Country-level process guide* (or a revised FTI Framework document that will incorporate the process guide) as a basis to develop a checklist of Local Education Group activities to be monitored.

OUTPUT 2	Indicator	Baseline [200X]	Milestone	Target [201X]	ASSUMPTIONS
[EDUCATION FINANCE] Adequate and sustainable domestic and external financing for education mobilised	Total				<input type="checkbox"/> Global economic conditions do not lead to reduced aid commitments <input type="checkbox"/> Domestic economic conditions are favourable <input type="checkbox"/> Humanitarian crises do not put a strain on the budget <input type="checkbox"/> Public financial management reforms keep pace in central government (esp. finance and planning), including both budget and fiduciary processes
	1. commitments	Source			
	2. disbursements	OECD DAC CRS			
	Indicator	Baseline [200X]	Milestone	Target [201X]	
	Share of basic education				
	1. commitments	Source			
	2. disbursements	OECD DAC CRS			
	in total education ODA to FTI eligible countries				
	Indicator	Baseline [200X]	Milestone	Target [201X]	
	Predictability of external financing: proportion of aid (to basic education) disbursed within scheduled fiscal year ^{1 2}				
	[Paris Declaration monitoring indicator 7]	Source OECD DAC CRS, Paris Declaration Survey			
	Indicator	Baseline [200X]	Milestone	Target [201X]	
Public recurrent spending on education as percentage of total public recurrent discretionary spending					
	Source				
	UIS				

1. The indicator can alternatively focus on long term (3-year) instead of short term (1-year) predictability. The Paris Declaration survey only focuses on the latter so additional information would be needed.
2. The Paris Declaration survey does not go down to (education) sector or (basic education) sub-sector level so supplementary information would be needed.

Activity	Indicator	Low	Middle	High
2.1 Engage with the Ministry of Finance to support domestic public financial management reforms that can be applied in the education sector	Quality and timeliness of in-year budget reports ¹	Weak financial planning, management (costing and budgeting), monitoring and reporting in education	Financial planning, management, monitoring and reporting processes developed in education	Consolidated public financial management system
2.2 Advocate increased levels of aid for basic education and mobilise external funds to replenish Catalytic Fund to improve the predictability of external aid flows	Commitments to Catalytic Fund	Volatile aid to education	Predictable aid to education	Predictable aid to education captured in budget documentation
2.3 Assess Catalytic Fund applications according to formula consistent with the FTI objectives	Distribution of Catalytic Fund allocations by criteria	Allocations unrelated to any formula	Introduction of formula	Allocations according to formula
2.4 Ensure that Catalytic Fund resources are provided in a timely manner to fund and support selected activities	Time lag between Catalytic Fund grant allocation and first disbursement	Non-FTI procedures delay disbursements	Disbursements as fast as alternative aid channels	Disbursements faster than alternative aid channels

1. PEFA indicator

OUTPUT 3	Indicator	Baseline [200X]	Milestone	Target [201X]	ASSUMPTIONS	
[DATA AND M&E] Reliable, timely and useful data being produced and used by national and local education stakeholders	[Indicator monitoring progress in quality of data systems in FTI eligible countries] ¹				<input type="checkbox"/> Cultural and political obstacles to evidence-based decision making overcome: sufficient demand to use data for policy <input type="checkbox"/> Staff appointed in key positions for education data management at central and local offices <input type="checkbox"/> Information and communication technology infrastructure promotes sharing of information	
		Source	FTI and UIS / World Bank			
		Indicator	Baseline [200X]	Milestone		Target [201X]
	Number of FTI eligible countries which report to the UIS their data on primary completion rate within 12 months from the end of the school year					
		Source	UIS			
		Indicator	Baseline [200X]	Milestone		Target [201X]
	Number of FTI eligible countries which conduct an annual school census that collects data from all types of primary education institutions (formal and non-formal, public and private, secular and religious)					
		Source	UIS			
		Indicator	Baseline [200X]	Milestone		Target [201X]

1. The indicator (or composite index) should be based on one or more of the dimensions of the Data Quality Assessment Framework (*A Framework for Assessing the Quality of Education Statistics*) developed jointly by the World Bank and the UIS: integrity, methodological soundness, accuracy and reliability, serviceability, accessibility.

Activity	Indicator	Low	Middle	High
3.1 Assess regularly the quality of data supporting the education sector plan and identify areas to be addressed	Number of FTI-eligible countries assessed	Review school census	Review all administrative data systems	Review all education statistics
3.2 Conduct joint programme with UIS and GMR to ensure that four key indicators (primary completion rate, net enrolment rate, adult literacy rate and gender parity) are monitored globally in a coordinated way	Number of countries for which a comparison of these four indicators (agreed between FTI, UIS and GMR to allow for comparisons across countries and over time) can be made for three consecutive years	All countries report systematically on 1 indicator	All countries report systematically on 2-3 indicators	All countries report systematically on 4 indicators
3.3 Conduct joint programmes to address gaps in data on education aid (with OECD DAC and individual major donors) and public education spending (with IMF IFS)	Number of countries for which a comparison of [core indicators of domestic and external financing to basic education] can be made for three consecutive years	Public finance and aid statistics can be monitored for 50% of eligible countries	Public finance and aid statistics can be monitored for 75% of eligible countries	Public finance and aid statistics can be monitored for all eligible countries
3.4 Conduct joint programme (with UNPD) to address gaps in out-of-school children data and population estimates	Number of FTI eligible countries where the estimation of out-of-school children is reformed	Conventional estimates of out-of-school children	Estimates take into account issue of children refugees	Estimates take into account issue of urban slum areas
3.5 Conduct joint programme to address gaps in learning outcomes data	Number of FTI eligible countries with a national systems for learning assessments / reading skills	10% of eligible countries have a national system	25% of eligible countries have a national system	50% of eligible countries have a national system
3.6 Use the EPDF to "deepen knowledge and learning around issues identified as priority by countries and in global fora, where appropriate in partnership with existing national and international institutions"	Volume of resources allocated to EPDF (new) Objective 2	Purpose of EPDF supported projects selected by international stakeholders	Purpose of EPDF supported projects selected by national stakeholders	Results of EPDF supported projects reflected in national reviews and policies
3.7 Publish reviews of policy experiences across countries to inform decision making by FTI partners	Number of publications that synthesise relevant findings on progress towards UPC/EFA	Two publications per year	Four publications per year	Six publications per year

OUTPUT 4	Indicator	Baseline [200X]	Milestone	Target [201X]	ASSUMPTIONS
[AID EFFECTIVENESS] Aid for primary education delivered efficiently	Alignment: Percentage of education aid that is listed in the budget and annexes [Paris Declaration monitoring indicator 3] ¹				<input type="checkbox"/> Global commitment on aid effectiveness through coordination, harmonization and alignment persists <input type="checkbox"/> Donor incentives to reform their behaviour. <input type="checkbox"/> Recipient countries demand accountability by all stakeholders
		Source			
		Paris Declaration Survey			
	Harmonisation: Use of common arrangements and procedures in education sector [Paris Declaration monitoring indicator 9]				
		Source			
		Paris Declaration Survey			
	Managing for results: Difference of amount received at Catalytic Fund grant renewal compared to original grant between strong and poor performing grants				
		Source			
		FTI			

1. The indicator can alternatively focus on how much education aid is being disbursed (1) through government PFM systems [Paris Declaration monitoring indicator 5a] and/or (2) using government procurement systems [Paris Declaration monitoring indicator 5b].

Activity	Indicator	Low	Middle	High
4.1 Encourage coordinated donor support to the education sector plan	Type of modality agreed for aid to basic education in general	Coordinated mechanism to assess need and deliver aid: JAM, CAP, IASC cluster ¹	Joint aid management and delivery: pooled funding, multi-donor trust funds	Country-led coordinating mechanism to manage aid and deliver results
4.2 Ensure CF and EPDF modalities are aligned with country processes	Type of modality agreed for grants from the Catalytic Fund	Not aligned	Partly aligned	Fully aligned
4.3 Review aid effectiveness arrangements in education sector across FTI eligible countries and act upon recommendations	Number of countries for which an annual briefing note on the implementation of the FTI compact is prepared ²	Aid provided in a fragmented manner and guided mostly by individual donor priorities	Coordination mechanisms reviewed and strengthened	Aid is coordinated, harmonised and aligned
4.4 Build capacity of Local Donor Groups ³	Number of workshops organised by FTI in country /region to inform Local Donor Groups of good practice in aid effectiveness	Poor and insufficiently consultative mechanisms in-country to discuss aid effectiveness	Stakeholders regularly provided with updated information on aid flows and aid effectiveness	Clarity on mutual commitments and time frames among partners

1. Relevant for fragile states: JAM = Joint Assessment Mission; CAP = Consolidated Appeals Process; IASC = Inter-Agency Standing Committee

2. This type of annual briefing note could become a requirement for Catalytic Fund disbursements.

3. Use the *Country-level process guide* as a basis to develop a checklist of Local Donor Group activities to be monitored.

Next steps

23. The above framework should be treated as a basis for a long overdue discussion among the partners. It is therefore necessary to make the following disclaimers, taking the following quote from the GRPP Sourcebook into account:

Although logframes are common in project evaluation, placing responsibility for the creation of a logframe on the evaluators themselves is more problematic for GRPPs. Many GRPPs have extensive authorizing environments, and the construction of a logframe should ideally be a participatory exercise among all the partners and participants in order to enhance accountability for results. (IEG 2007, ¶9.6)

24. *The framework needs to be developed and owned by the FTI.* In any case, this evaluation has recommended a substantial re-design of the FTI which should feed into this process. Therefore, the aim of the proposed M&E framework is to help initiate a process (to agree activities and outputs, indicators and targets) and should not to be seen as an off-the-shelf solution to the FTI's need for an M&E framework. This process should be seen as an opportunity to articulate a common interpretation of the objectives of the partnership and what the partnership will do for them to be achieved. It is envisaged that the search for clear indicators and agreed means of verification / data collection processes will forge joint understanding of FTI potential and limitations among stakeholders and reach agreement on what exactly the partnership attempts to accomplish, how likely it is that the selected activities will lead to the accomplishment of these targets and what the implications are in terms of resources.

25. A long-term framework for impact assessment is not a substitute for basic management monitoring of the activities that the FTI partners will commit to undertake and the inputs they will provide. In order to achieve this, *the M&E function at the Secretariat needs to be strengthened considerably.*

26. The proposed logical framework does not include *a narrative analysis of the results chain mapping the FTI channels of influence.* Various FTI documents need to be pulled together to describe in concrete terms the nature of the compact and its implications regarding the channels of influence that fall directly under the control of the FTI partnership and those that are left to the discretion of individual partners.

27. The proposed M&E framework constitutes a part of the overall recommendation of this evaluation that the FTI partnership should aim to articulate a common interpretation of the objectives of the partnership and what the partnership will do in order to achieve them. The search for clear indicators and agreed means of verification should forge a joint understanding among stakeholders of the FTI's potential and limitations. It should help in reaching agreement on what exactly the partnership attempts to accomplish, how likely it is that the selected activities will lead to the accomplishment of these targets, and what the implications are in terms of financial resources, the configuration of the Secretariat, and the sharing of responsibilities among the various bodies connected with the FTI.

28. As recommended in **Chapter 5 of the main report**, *a new constitutional document for the FTI, to replace the 2004 FTI Framework, should be developed.* Such a document would supersede not only the existing Framework document but also the country level processes guide and the governance document and would also capture the final version of the M&E framework. The M&E framework should be an intrinsic part of the design on the way forward and not just an afterthought.

C. Issues in impact evaluation

29. This section focuses on the element of the TOR that relates specifically to an eventual impact evaluation of the FTI – to "*provide guidance for the assessment of the impact of the FTI at a later stage*" (TOR ¶17)

30. Agreement on a monitoring and evaluation framework that reflects consensus on activities, outputs, purpose and goal among the partners will significantly simplify future evaluation efforts. This will be achieved primarily through the strengthening of the monitoring mechanisms that will be required.

31. This evaluation commissioned a scoping study (White 2009) on the feasible approaches to FTI impact evaluation and their respective advantages. It contrasted two commonly used but different definitions of impact:

2.2 The tradition in evaluation has been that "impact" refers to the final level of the causal chain (or log frame), with impact differing from outcomes as the former refers to long-term effects. For example, the DAC definition of impact is "positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended". Any evaluation which refers to impact (or often outcome) indicators is thus, by definition, an impact evaluation. ...

2.3 But this definition is not shared by many working on impact evaluation Impact is defined as the difference in the indicator of interest (Y) with the intervention (Y1) and without the intervention (Y0). That is, $\text{impact} = Y1 - Y0$ (e.g. Ravallion 2005). An impact evaluation is a study which tackles the issue of attribution by identifying the counterfactual value of Y (Y0) in a rigorous manner⁵

32. The first definition focuses on monitoring the achievement of outcomes of a specific (set of) intervention(s). The second definition is not restricted to outcomes (it also covers indicators lower down the causal chain) and also places emphasis on tackling attribution.⁶

33. The scoping study focused on the potential for applying attribution analysis, in other words, on the possibility of using a particular methodology to assess the proportion of observed change which can really be attributed to the FTI or to another specific (set of) intervention(s).

34. In brief, attribution can be examined at two levels:

- (a) At the input-to-output level: whether the FTI influences policy and resource allocations at the global and country level; and
- (b) At the output-to-purpose level: whether the policy changes (supported by the FTI) have a beneficial impact on educational outcomes.

35. Evaluating the *input-to-output* relationships requires a qualitative approach starting from a mapping of the possible channels of the FTI's (direct and indirect) influence on policy and resource allocation decisions at the country level (as mentioned in ¶26 above). However, many factors outside the FTI affect these decisions. This was demonstrated in the

⁵ Note that such impact evaluations cannot easily be done in retrospect. The baseline data need to be collected at the outset of a programme in order to try to measure change over its duration.

⁶ The terms of reference of the FTI evaluation alternate between the two possible definitions of impact evaluation. For example, when talking of 'development of a manageable set of indicators for tracking progress towards sustainable impact and baseline data from case study countries on those indicators' it is a case of the first definition (with *impact* referring to monitoring of outcome trends). When asking 'how has the FTI impacted government support for other EFA objectives', it is a case of the second definition (with *impact* referring to counterfactual analysis).

case studies under this evaluation, which involved stakeholder analyses at the country and agency level to come to specific conclusions about which policies were assisted by the FTI and how. FTI activities and influences at the country level will continue to be rather heterogeneous and a small part of a complex picture. The scoping study noted the difficulty of constructing a comparison group for a counterfactual scenario that can lead to a clear interpretation of results:

- There are many *confounding factors* affecting education policy and resource allocation decisions other than the FTI.
- There is a *selection bias*, as countries are endorsed because they have sounder education policies, so these countries would likely be better performers than non-FTI countries even in the absence of the FTI.
- There are positive *spill-over effects* as non-endorsed countries may benefit from knowledge of better sector policies even if they do not join the FTI.
- *Contamination* may occur as other agencies implement parallel projects that affect the outcomes of interest. (White 2009Table 1)

36. An assessment along the lines of the mid-term evaluation, which will establish whether the activities undertaken by the FTI partnership contributed to the achievement of the outputs, is necessary. This assessment will be greatly facilitated with a clear FTI M&E framework. However, partners need to be aware that this evaluation cannot be conducted in the expectation that the results will be able to be clearly attributed to the FTI.

37. Evaluating the *output-to-purpose* relationships, in other words whether (policy, financing, data and capacity) changes relevant to FTI objectives are linked to changes in educational outcomes, also has to tackle similar problems in selecting a control group but solutions are more likely to be forthcoming. This approach will resonate with the 2004 FTI Framework goal to establish which policies work and then propagate them. Two different ways can be followed.

38. First, existing evidence can be marshalled. Since the FTI was established, a number of organisations have built up a body of evidence on what works in primary education, including, not least, some of the FTI partners themselves, such as the EFA-GMR and individual donor agencies, but also the International Initiative for Impact Evaluation (3ie), the Abdul Latif Jameel Poverty Action Lab (J-PAL) and the Innovations for Poverty Action (IPA). The FTI can prepare summaries of these reviews (particularly in the form of briefing notes for policy makers) to benchmark good practice in line with one of its working principles. This would help:

- ensure the relevance of the FTI by promoting policies with proven impact; and
- guide any original empirical work to be undertaken in areas in which there is insufficient knowledge.

39. Second, new quantitative work could be undertaken at country level. This could take two forms:

- A programme of data collection for specific indicators of the Indicative Framework for which existing systems are currently weak. This could first concentrate on intermediate outcome indicators which could be used to illuminate progress both for EFA and for the FTI and for which there are usually reasonably reliable data available although not currently monitored systematically (e.g. books per student, latrines per student) and then move on to indicators that are harder to measure (e.g. effective annual instructional hours, proportion of children coming to school without breakfast etc).

- Targeted impact studies of selected policies to understand the effects of policy changes that have been supported by the FTI (but for which knowledge gaps have been identified). These might be based on existing or newly collected data. Indeed, in the course of this evaluation the evaluation team has noted proposals to implement in-depth (and costly) studies at country level.⁷ However, the evaluation team would not recommend that such studies be pursued because – in the light of the overall agenda – it is unlikely that they would usefully (and in a cost-effective manner) address the issue of *FTI impact* that should remain the priority for the partnership. Moreover, it is unclear whether the FTI, even in a restructured form, will have a comparative advantage in directly undertaking or even commissioning such studies.

D. Options

40. The above outline has argued that a rigorous quantitative evaluation of the full impact of the FTI is not possible. However, it has presented a menu of options from which the partnership can choose in order to highlight those aspects of the FTI that demonstrate the Initiative's relevance, effectiveness and efficiency. The three options proposed below are based on an assumption that the partnership will have an evaluation budget broadly equivalent to that at its disposal for the mid-term evaluation. Rough order-of-magnitude suggestions are given for how the available budget might be divided among the three options.

41. First, a second/final evaluation of the FTI will be undertaken in five years' time along the lines of the mid-term evaluation. Contribution analysis will be used to demonstrate whether or not the FTI was one of the causes of observed changes (at the national and global levels) from the point of view of achieving the expected outputs, outcomes and impacts. This evaluation will have a much better and more efficient starting point than the mid-term evaluation if an M&E framework is adopted and if FTI monitoring systems are strengthened and adjusted to collect information on all the selected indicators. [*Dedicate 50% of impact evaluation budget to this activity*]

42. Second, indicative meta-analyses and reviews of core education policy questions could be undertaken on a range of issues that the FTI has supported or is interested to support: conditional cash transfers, abolition of fees, regulation of private schools, school proximity, decentralisation of education decisions, early childhood education programmes, special needs education programmes, teacher training programmes, monitoring of attendance, school grants, school feeding programmes, provision of instruction materials and textbooks etc.⁸ There is a great deal of knowledge about the effectiveness of these programmes. What is needed is a substantial job to synthesise and present them in ways that are appropriate for informed policy discussions to take place. The FTI could therefore generate a series of policy briefings drawing on studies already undertaken. It is noteworthy that, compared to health services research, the tradition of systematic reviews is in its infancy in the international education field (e.g. there is no equivalent to the internationally recognised Cochrane collaboration⁹ in the health field). The proposal here therefore is to draw up a list of topics and produce a series of low cost policy briefings based on existing

⁷ The commissioned paper from 3ie includes outline TOR and approach for such studies, together with indicative costs and a suggestion as to which countries might present least difficulty in terms of availability of data (e.g. from household surveys), see White 2009, Annex E.

⁸ This activity could also be used to compile the results of studies on the impact of education on personal, institutional, economic and social aspects of life (notably health and fertility) in developing countries, as discussed in ¶17.

⁹ See www.cochrane.org/.

research. Once a body of reviews has been completed (e.g. within the next 18 months), it would be appropriate for the FTI (in conjunction with GMR and UIS) to participate in a series of workshops on sensible approaches to impact evaluation for the education sector. [*Dedicate 15% of impact evaluation budget to this activity*]

43. Third, strengthen the monitoring of particular indicators which are central to the FTI partnership yet for which there are substantial gaps. These might include indicators of assumptions of what is (or might be) good educational policy practice (such as the effective number of instructional hours), indicators of adherence to the FTI compact (such as basic education aid commitments and disbursements) or indicators of desirable educational outcomes (such as learning outcomes). The model would be to support existing institutions collecting data, such as the UIS, DAC or other. Their involvement in the design of the M&E framework should be considered. [*Dedicate 35% of impact evaluation budget to this activity*]

E. Summary

44. This appendix has argued that a useful impact evaluation (and indeed an effective partnership) requires a clear programme logic with well-specified indicators. This element has been missing and has resulted in lack of clarity about what the partnership is meant to be doing and what resources it is supposed to be expending accordingly. A detailed M&E framework has been proposed as a basis for discussion among FTI partners. It calls for a sharper focus of the monitoring activities to be undertaken directly by the FTI.

45. With respect to an eventual impact evaluation that identifies the contribution of the FTI to improved policies, finance, data and capacity at the country level, the same challenges that faced the mid-term evaluation are likely to continue. In particular, although the focus of the partnership may sharpen, it is likely that the activities of the FTI will continue to be a small part of a larger picture. In this context, it is questionable whether impact evaluation case studies of particular (small and perhaps non-replicable) packages of FTI inputs would be as good a use of evaluation resources as analyses of specific policies that the FTI does or might promote, or improvements in the information base.

References

- Bennett and Fairbank 2003 *The System-Wide Effects of the Global Fund to Fight AIDS, Tuberculosis and Malaria: a conceptual framework*. Bethesda Maryland: Partners for Health ReformPlus. Sara Bennett and Alan Fairbank, 2003.
- Cambridge Education, Mokoro & OPM 2009 *Mid-Term Evaluation of the EFA Fast Track Initiative: The Evaluation Framework-Evaluation Team Guidelines on Process and Methodology*. Cambridge Education, Mokoro and OPM, 2009.
- FTI 2004 *Accelerating progress towards quality universal primary education: Framework*. FTI, 2004.
- FTI 2007 *FTI Support to fragile situations: Progressive Framework*. FTI, May 2007.
- FTI 2008 *Strengthening Quality Assurance Mechanisms within the Fast Track Initiative: A review of the Twenty-eight Education Sector Plans*. FTI, 2008.
- G8 2009 *Sharing Responsibilities to Advance Education For All*. G8 Education Experts Report 2009. L'Aquila, July 2009.
- IEG 2007 *Sourcebook for Evaluating Global and Regional Partnership Programs: Indicative Principles and Standards*. Washington: World Bank. Independent Evaluation Group, 2007.
- Ravallion 2005 *Evaluating Anti-Poverty Programs*. Development Research Group, World Bank. Martin Ravallion, 2005.
- Stern et al 2008 *The Paris Declaration, Aid Effectiveness and Development Effectiveness*; DaRa, Madrid. Ministry of Foreign Affairs of Denmark, Copenhagen. Elliot D. Stern with contributions by Laura Altinger, Osvaldo Feinstein, Marta Marañón, Nils-Sjard Schulz and Nicolai Steen Nielsen, November 2008.
- White 2009 White, Howard (2009) *Working Paper 3b: Impact Evaluation Scoping Study*, Prepared for the Mid-Term Evaluation of the Education For All Fast Track Initiative, draft, 11 March 2009, International Initiative for Impact Evaluation (3ie).
- WB & UNESCO 2003 *A Framework for Assessing the Quality of Education Statistics*. World Bank Development Data Group and UNESCO Institute for Statistics. January 2003.