EVALUATION DEPARTMENT

Norad

Evaluation of Norwegian Multilateral Support to Basic Education: **Ethiopia Case Study**

Evaluation of Norwegian Multilateral Support to Basic Education: Ethiopia Case Study

Development Portfolio Management Group

October 2015

USC University of
Southern California

Acknowledgements

Table of Contents

Tables

Figures

Acronyms and Abbreviations

ABE	Alternative Basic Education
ABEC	Alternative Basic Education Center
BE	Basic Education
DPMG	Development Portfolio Management Group
ECD	Early Childhood Development
ECE	Early Childhood Education
EMIS	Education Information Management System
ESDP	Education Sector Development Plan
FTI	Fast-Track Initiative
GDP	Gross Domestic Product
GEQIP	General Education Quality Improvement Program
GER	Gross Enrolment Rate
GP	Global Partnership
GPE	Global Partnership for Education
HIV/AIDs	Human Immunodeficiency Virus Infection/Acquired Immune Deficiency Syndrome
IDA	International Development Association
LEG	Local Education Group
M&E	Monitoring & Evaluation
MDGs	Millennium Development Goals
MOE	Ministry of Education
NER	Net Enrollment Ratio
NLA	National Learning Assessment
PEFA	Public Expenditure and Financial Accountability
PFM	Public Financial Management
PTAs	Parent Teacher Associations
SIP	School Improvement Plan
ToC	Theory of Change
UN	United Nations
UNICEF	United Nations Children's Fund
USAID	United States Agency for International Development
WASH	Water, Sanitation and Hygiene

Executive Summary

Introduction

1. This is one of the four country case studies that has been conducted by the Development Portfolio Management Group (DPMG) as input in its evaluation of Norway's support to basic education through the United Nations Children's Fund (UNICEF) and the Global Partnership for Education (GPE). The overall evaluation covers the period 2009 to 2013 and so does this case study, except that the UNICEF Country Programme examined here started in 2007. The GPE, during that time in Ethiopia, was merged with a multi-agency support programme called the General Education Quality Improvement Program (GEQIP), which was supervised by the World Bank during the period 2009-2013, and which was mainly focused on "improving the teaching and learning conditions in primary and secondary education". That program was aligned with and contributed to Ethiopia's National Education Sector Development Plan IV (ESDP IV), which ran from 2011 to 2015. The UNICEF Country Programme that was the focus of this study was the one covering the years 2007 to 2011, particularly the "basic education" component.

2. The case study was an extension of a desk study of these two development-support programmes undertaken by the Development Portfolio Management Group core staff (covering all relevant program and Norad documents). It also included a review of documents – those at the country level, interviews with key informants from the Ministry of Education, UNICEF, the World Bank, other development agency personal, and a civil society organization (the National Teachers Union), plus visits to schools in Addis Ababa and Amhara District, including an interview with the school principals (see list of persons interviewed in Annex 2). The evaluation adopted as its conceptual framework a "Theory of Change" model (Figure 1), that DPMG framed drawing on contemporary education change literature.

3. **Country Context**. Ethiopia is the second most populous country in Sub-Saharan Africa. In 2009, the total population of the country was about 81.9 Million whereas in 2013 it has grown to 94.1 million. The average growth rate of the population between 2009 and 2013 was 3%, with only a slight increase in the percent urban (2%). Infant mortality significantly decreased during that period from 81 per 1000 to 64. Ethiopia also has one of the fastest growing economies in the world, for instance, its annual growth rate reaching 10% in 2013, helping it to reduce those living below the international poverty line from 40% in 2004 to 30% in 2011. Despite this, it remains one of the poorest countries in the world with a per capita income in 2013 of only US$ 370. At this point the country is still one of the largest donor recipients worldwide, receiving over $3 billion annually.

4. **Education Context.** Over the past two Education Sector Development Plans (II and III) Ethiopia went through rapid expansion of its basic education program, such that by 2012-2013 the net enrollment ratio for grades 1-4 (lower primary) reached 95.5% and for total primary (lower plus upper, grades 1-8) about 86%. These general figures mask large regional variations (the gross enrollment ratio) in Gambella district in 2012/13 above 125% compared to Afar district at 50%). Early childhood development also has been growing rapidly, but started at a very low base, so the coverage is still modest, with the GER being 26.1% in 2012/13 compared to 4.2% in 2008/09). Gender parity in enrollments was basically attained in primary education during this period in a global sense but not across all districts.

5. **Growth & Transformation Plan (2010/11-2014/15) & Education Sector Development Plan (2011-2015).** When the two programs were initiated the current Growth and Transformation Plan was being developed. It overall goals were national economic growth of 11-15% per year over the period and for education it adopted the Ethiopia's National Education Sector Development Plan IV goals, which aimed to ensure the achievement of MDGs in education and "equitable access to quality education in general".

6. **Outcomes.** The DPMG Theory of Change highlighted the three outcomes of interest to the Norwegian Ministry of Foreign Affairs for the evaluation, **quality educational outcomes**, **gender equality**, and **equity (e.g., for marginalized groups)**. In the conduct of this evaluation, we examined the GPE and

UNICEF programs to see to what extent they emphasized these outcomes and then tried to assess how effectively they did so.

Findings

- **GPE/GEQIP.** For **improved learning outcomes**, this program included as one of its main performance indicators a change in the percent of students scoring at least 50% on the countries National Learning Assessment for grades 4 and 8 (between 2007 and 2012/13). The results shows substantial increases at both grade levels (grade 4: 44% and grade 8: 80%), with changes for girls modestly higher than for boys at grade 4 and considerably higher (*4 times*) at grade 8. This indicator was also used as an indicator of **gender equality** (showing robust results); in addition, the evaluation examined male/female breakdowns for primary school completion rates. On the latter, results showed girls improving at a much faster rate than boys, to the point that by 2012-2013 they reached parity. Finally, for **equity**, the Program did not include a measure of equity for excluded groups in its results framework, but the Ministry's education management information system presented regional breakdowns on **net enrollment ratios**. Since regions vary greatly in terms of economic level and kinds of livelihoods (e.g., pastoralism vs service sector) this was used as a proxy for **equity**. This data shows higher than national level change in net enrollment ratios for all four "emerging regions" of the country.

- **UNICEF.** On **improved learning outcomes**, UNICEF also aspired to track improvement in the proportion of students achieving 50% or above on the National Learning Assessment, with a target of 85% (at least in its 300 targeted districts) but in the end was not able to present change scores. It did present a tracking of the number of schools adopting the Child Friendly School model, with its total of 895 coming close to its target of 1000. On **gender equality** UNICEF was very active in this field but presented only one indicator, the gender parity index which reached 0.96 in its targeted districts by 2011, somewhat below its target of 1.0 (it was not clear if the results was for those districts only or for the nation as a whole). Finally, for **equity**, where UNICEF was also very active, there were no performance indicators (as with GPE/UNICEF), but the national Education Information Management System results by region could also be used as a proxy outcome for its efforts.

- **Qualitative assessments.** The field study gave the team a chance to create qualitative data on outcomes, including those on two dimensions: a) answer to questions about the programs most valuable ("positive") contributions, and b) perceptions of the program's "upstream contributions."

    GPE/GEQIP. For GPE/GEQIP respondents indicated its most positive contributions were:
    - Its focus on the Ethiopian education sector's highest priority: improving the quality of education and student achievement;
    - The pooling of a sizable amount of funding from many agencies, which required -- and in fact resulted in -- carefully designed aid-management systems, including good coordination between the government and donors, solid financial management, joint supervision, and serious monitoring and evaluation.
    - The skillful GPE/GEQIP blend of existing government structures and a technical advisory team for the implementation of the project.

    UNICEF. In the case of UNICEF, the perceptions highlighted these positive effects:
    - Reaching the unreached through Alternative Basic Education and in vitalizing Early Childhood Development,
    - Creating knowledge and recommendations from its study on out-of-school children and drop outs:

- Being widely participatory and approachable: a perception among education sector leaders and senior bureaucrats that "the government and this UN agency can work together successfully."
- Its contribution to emergency education;
- Its decision to transform its **Child Friendly Schools model into something closer to the School Improvement Plan approach.**

Upstream Contributions are those related to policies, plans, regulations, guidelines, etc.

- GPE/GEQIP. GPE/GEQIP was recognized for its contributions to systemic design of quality assurance mechanisms at the school level (self-assessment methods and school report cards), a school inspection policy, the school grant mechanism; and ESDP IV.
- UNICEF. UNICEF is recognized for a long list, including strategies and policies for education in pastoral areas, studies and proactive systems for improving female participation and outcomes; drafting Early Childhood Development policies, standards, and approaches; initiating and supporting research on out-of-school children and the development of Ministry of Education guidelines; supporting design of the 2010-15 sector plan (an influence on almost every chapter).

7.   Interventions.   According to our Theory of Change, certain interventions are likely to have influence on the priority outcomes, and these were grouped into the following categories or thematic areas:

- Schools and classrooms
- Teachers
- Curriculum
- Student/teacher time on task
- Learning materials
- Language of instruction
- Teacher/school/system supervision
- Changes in the learning environment
- Child's background and resources (included early childhood education)
- Community and parents

8.   The evaluation then examined the two multilateral programs for how and how well they addressed these themes.

9.   GPE/GEQIP.   This program supported interventions in all of these thematic areas except for schools and classrooms and student/teacher time on task. In most cases the program met or exceeded its delivery targets. Three interventions were particularly strong and influential, **increasing the number of qualified teachers**, **producing/distributing learning materials based on the revised curriculum,** and **improved the learning environment (mainly through the School Improvement Plan/school grants program)**.

10.   UNICEF.   UNICEF's interventions were numerous and significant in their magnitude.   The indicators were not such that change over time could be well documented and the reaching of targets was often hard to determine; nevertheless it is clear that UNICEF made big contributions **in schools and classrooms** (especially for the Alternative Basic Education programs), the **learning environment** (mounting Child Friendly Schools which morphed into School Improvement Plan programs); supporting the augmentation of **children's resources** (school readiness) through Early Childhood Development programs, and its various **upstream contributions**.

Aid-Management, Financial and Enabling Conditions

11.   Aid Management. Several factors determine the quality and thus the value added of an agency's aid management.   Some are enabling conditions, such as how the agency's activities are financed. However, proximate factors are whether the agency manages the project cycle well.

12.   GPE. Several parties are accountable for the quality of GPE's management of donor aid: the GPE Board and its committees: its executive arm, the Secretariat; the Ethiopia Local Education Group; and the supervising entity (World Bank) for GPE's Project to Improve Education Quality in Ethiopia (PIEQM) (2010-2015).   For the Ethiopia grant application to GPE, the **Secretariat** commissioned an External Quality Review that was highly professional. The Review flagged not just strengths, but also weaknesses that could have threatened the success of the GPE project.

13.   Local Education Group. In Ethiopia the Local Education Group (LEG) is called the Education Technical Working Group (ETWG). Overall, the data reveal an actively engaged ETWG/LEG and a satisfied membership, with only minor changes proposed to improve functioning. The data show a surprising consensus among respondents on which functions the ETWG performs, their frequency, and the respondents' satisfaction with how the ETWG discharges these functions. In response to a question about how effectively members of the ETWG collaborate with each other, interviewees uniformly rated the collaboration as effective.

14.   ETWG members were unable to rate the performance of the supervising entity. Although Education for All-Fast-track Initiative had funded 40% of GEQIP I, the respondents noted that ETWG's annual Joint Review Missions monitored aspects of the implementation of the Government's overall sector plan, not GEQIP, which was monitored by the funding partners.

15.   The ETWG apparently did not appraise the version of the Education Sector Development Program on which the GEQIP is based, although it did re-endorse it. Prior to 2012, the GPE Secretariat did not systematically recommend that the Development Partners conduct an Education Sector Plan appraisal and endorsement before engaging on each new phase of a long-term education strategy. In the event the quality of the ESPD was mixed. The ESDP is comprehensive, addressing each sub-sector and cross-cutting initiative. There is a meta-matrix of key performance indicators, baseline data, and targets by year for the overall ESDP. The relationships between the activities, outputs, and true outcomes are generally logical and plausible. The targets are the primary flaw. Many are questionable or implausible. GPE's new financing model, introduced in 2015, where a share of the grant will be "held hostage" to progress on education sector plan (not project) targets, will penalize implausible targets in ESPs. The ESDP's M&E framework for the plan is well structured with a calendar of annual joint review missions and Annual Review Meetings that will focus attention on the ESDP's results framework and pressure the implementing units to generate credible M&E data.

16.   Based on an analysis of GEQIP documents used to evaluate the **Supervising Entity (World Bank)**, the quality of design is rated moderately unsatisfactory, primarily because of unresolved problems with its results framework and the unmitigated complexity of the operation relative to Government's capacities. To its credit, the Bank tried to scale back the operation, but unsuccessfully. Its outcome and output targets were considerably more realistic than those of the ESDP.

17.   The quality of supervision is rated fully satisfactory. Given problematic aspects of the design, it is to the credit of the Government, the World Bank as supervising entity, and the pooling Partners that the project succeeded as well as it did. In 4 ½ years the Government of Ethiopia disbursed 99.6% of almost US$436 million out of $437.7 million. As noted below, the ICR determined that most program outcome targets were achieved.

18.   The Project was under close supervision throughout its lifetime. The World Bank and the pooling Partners conducted regular and frequent joint missions on a quarterly basis during the early phases of implementation and bi-annually later on. These missions provided implementation support to the Ministry of Education (MoE) on various aspects of implementation, including procurement, financial management,

and M&E. The Bank's biannual Implementation Status Reports give numerous examples of actions taken by the World Bank and pooling Partners to resolve implementation and partnership problems. The World Bank conducted its required Implementation Completion Report after GEQIP I closed. It was thorough and satisfactory, finding that program outcome targets except for the Teacher Development Program were largely achieved.

19. **UNICEF program.** UNICEF's project cycle, the quality of which partly defines the quality of UNICEF's aid management for Ethiopia, starts with the United Nations Development Assistance Framework (UNDAF) for Ethiopia. The UNDAF structures aid activities by the large number of UN agencies active in Ethiopia. UNICEF's Country Program Document for 2007-2011 includes support to basic education that is aligned with 2007-2011 UNDAF priorities. UNICEF's Country Program Action Plan for 2007-2011 elaborates on themes in the Country Program Document. UNICEF reports on progress against the CPAP annually.

20. Project design falls between moderately unsatisfactory and unsatisfactory. In support of Ethiopia's Education Sector Development Program III (ESDP III), UNICEF's Country Program Action Plan focuses its basic education program on three projects and seeks specified outcome targets for its basic education program for 2007-2011--for example, the net intake rate for grade 1 increases from 51 per cent to 96 per cent of schools in targeted *woredas;* gender parity achieved in targeted *woredas*; and 85 per cent of children retained and attaining at least 50-per-cent pass result in grades 4 and 8 in the targeted *woredas.*

21. It is hard to discern from the upstream documents the design of the components or activities that UNICEF will use to achieve these objectives. The Country Program Action Plan should translate the Country Program Document's approach into operational activities, but its discussion of its basic education program remains at the strategic level. The Country Program Action Plan's results framework is organized around the three projects, but, the activities listed under a "project" are quite disparate and cannot be construed as a designed component.

22. UNICEF's results framework for the 2007-2011 program has to be patched together between upstream documents. The UNDAF results framework focuses more on outcomes, baselines, and targets; the CPAP, on outputs, indicators, baselines, and targets. The composite framework defines the outputs sought, with indicators, baselines, targets, data sources, and applicable risks/ assumptions. It identifies outcomes, baselines, and targets. However, UNICEF's theory of change or causal pathways from activities to outputs to outcomes cannot be constructed from the upstream documents or from the composite results framework. Some outcome targets seem unrealistic. These may be targets sought by the Government of Ethiopia in its ESDP III, but, even if they are, UNICEF should have examined their plausibility before adopting them as their accountabilities for their 2007-2011 plan.

23. UNICEF's supervision reporting falls between moderately satisfactory and moderately unsatisfactory, in good part because annual reports could not report against a clean and coherent results framework. Annual reports included helpful information, including the most recent Government data on outcomes.

24. However, targets shifted between the Country Program Action Plan and annual reports. For example, under "key milestones and expected results" the Country Program Action Plan expected that by the end of 2007 the NIR would be 81% and 96% by the end of 2008. In the 2007 annual report, the target date for reaching 96% had shifted to 2015. Some of the inferences about the effects of outputs seemed unwarranted.

25. **Financial Management.** The assessment, both using desk studies and some personal data collection at government offices, developed a credible overview of how public sector funds were managed during the period of these programs, covering the themes of **public financial management systems, budget credibility, predictability and control in budget execution, and fungibility and/or additionality,** plus a recent **ratio and trend analysis for the education sector.** The assessments were basically positive, with for example, the public financial management systems being seen as performing well (previous weakness having been dealt with), budget credibility receiving a B+ rating, predictability and

control of execution rated at C+ (unchanged from an earlier period), and fungibility/additionality showing increases in variance between actual and original budget expenditures but that the variance was largely due to capital expenditures (largely caused by the oil price increase). The ratio and trend analysis for the period is summarized in Table A.

Table A: Ratio and Trend Analysis

Category	2009/10	2010/11	2011/12	2012/13
% of national budget allocated to education	23.8%	24.9%	23.9%	22.5%
% of national recurrent budget for recurrent expenditures in education	29.5%	30.5%	31.6%	32.4%
% of education budget allocated to recurrent expenditures	56.9%	53.0%	54.6%	58.6%
% of GDP allocated to education	4.7%	4.5%	3.4%	3.3%

26.   **Enabling Conditions.**   The Theory of Change postulated that the following conditions would have an effect on the "ability" of the interventions to make a difference on the desired educational outcomes, grouped by **System-Level Enabling Conditions** and **Basic Education Enabling Conditions** as show in Table B.

Table B: Enabling Conditions

System-Level Enabling Conditions	Basic Education Enabling Conditions
• National/political commitment to improved quality and equity in BE, including ECE	• Community/ parental involvement and supports especially for girls/ disadvantaged
• High share of national budget for education	• Fee free policy, school feeding, scholarships for Disadvantaged
• Conflict/disaster sensitive mechanisms in place	
• Functioning Local Education Group	• Institutional capacity at national and local level (for planning and implementation)
• Appropriate multi-lingual policy in place & funded	
• Disaggregated Education Information Management System & learning assessments in place	• Strong budgetary support for BE, including ECE

27.   The team examined the current status of each one of these conditions and roughly determined how strong each was, using a three point hierarchy, **strong, moderate, weaker.** The results are summarized here shading the above list as follows: strong: green; medium: blue, and weaker, yellow.

28.   **Causal Connections**. Such connections were mapped out schematically in a figure showing both pathways and the strength of connections (see Figure 1). The results of the analysis are summarized as follows:

> *…our findings are that half of the six "system-level" enabling conditions were strong and the other half moderate. For the "basic education" conditions there were no strong ones: instead two were moderate and two were weaker. Our informants indicate there that there continue to be capacity constraints especially at the local (woreda) level and that funding for basic (k-8) is lower than the GPE standard. This suggests that effectiveness of the basic education improvement programs are constrained by low level capacity at the district level and tight funding.*
>
> *Concerning the connections between interventions and outcomes, first it can be seen that an important variable in our model, "student/teacher time on task," was not seriously address by either program. Also, there is some evidence of complementarity between the two programs, with Global Partnership providing more support for learning outcomes (in fact it has only two narrow arrows connected to equity), and UNICEF more for the equity ones, and, at the intervention level, Global Partnership*

covering teachers, learning materials and language of instruction, whereas UNICEF has supported more children's resources (Early Childhood Development), schools and classrooms (especially Alternative Basic Education Centers). Both Programs made upstream contributions, but UNICEF went both to learning outcomes and equity, whereas Global Partnership went mostly to learning outcomes. Of course, the two causal arrows in the schemata are those determined by the Global Partnership for qualified teachers and learning materials, which were provided on a massive scale.

Lessons Learned

29. Different questions were forwarded to respondents in order to widely explore the lessons learned from implementing both GPE/GEQIP and the UNICEF country programs. For the sake of brevity the list of lessons learned will cover the main points only, without elaboration:

- Need for program and geographic focus;

- The need for more effective Monitoring and Evaluation;

- Need for strengthened Implementation capacity, especially at the local level;

- *Need for a differentiated approach*es;

- Better planning of institutional arrangements;

- Effective communication with the field and timely budget release.

Conclusions

30. The main purpose of this evaluation is to "generate evaluation evidence on the relevance, efficiency and effectiveness of Norwegian aid to basic education through UNICEF and GPE, focusing particularly on the achievement of Norway's policy objectives quality of learning, gender equality and equity..." The first line of each conclusion is offered here, with more details available in the main body of the report:

31. *Outcomes*. Findings in Ethiopia reveal substantial improvements over a five year period (2009-2013) on all three of Norway's priority areas: improved learning outcomes, gender equality and equity (e.g., for marginalized children).

32. *Interventions*. **GPE/GEQIP** supported two large-scale interventions, teacher training and providing learning materials (textbooks) that were exceptional – their causal connections to learning outcomes were determined and were positive. Other successful GPE/GEQIP interventions with "implied connections" were *curricular reform*, and *improving the learning environment* (through SIPs and school grants), and a set of upstream contributions to the quality assurance mechanism and an inspection policy framework. **UNICEF'**s line of connection ran more to **equality issues**. In many cases its interventions were large in scale, like that in support of Alternative Basic Education Center's or of Early Childhood Development, but it was difficult to gauge their impact given the way they were tracked. The agency did create successful connections through its upstream contributions, however, that benefited all three outcome areas.

33. *Enabling conditions*. The success of the interventions is presumed by our Theory of Change to be <u>enabled</u> by numerous contextual conditions, both at the system-level (the education "system") and the more operational level ("basic education"). As a backdrop to all is the *national strategic context*, which during our evaluation period (2009-2013) can be characterized as one of strong economic growth (10% in 2013), political stability, and relative peace and security.

Recommendations

34. Given the positive results on priority outcomes, the reach and effectiveness of Program interventions, and their mostly "implied" connection to the outcomes, the two agency efforts under review here can be considered to have been beneficial to Ethiopia and its basic education programs and to have generally been good investments by Norway. However, for subsequent Programs (e.g., GEQIP stages or

new Country Programs), our findings prompt us to suggest some new ways of organizing or managing development cooperation.

1) ***Participatory approach to project planning****.* Projects like GPE/GEQIP or UNICEF-Country program are taking place within a societal context which is made up of multiple players or stakeholders. Since these stakeholders would participate in the project in different capacities, it is extremely useful to create an opportunity for them to actively participate in project planning from the inception. This would create clear impressions of what the project is all about, and what means and methods it is going to use. There will also be realistic understanding of the project. Both GPE/GEQIP and UNICEF suffered from "false expectations" by those who were involved in implementing the programs (especially at regional/district offices). The idea and goals of the project as well as its relationships with these levels of implementers should have been addressed from the start before they could grow into issues of concern.

2) ***Purposing and focusing projects in the planning and design phase****.* It has been indicated that the two supported aid programs lacked clear goals, purposes or geographic boundaries. They wanted to do everything all over the country. Some call them thinly-spread. The two Programs might have been more effective if they had put more resources into their lead interventions, such as alternative basic education and Early Childhood Development for UNICEF, and the School Improvement Plan/Grants for GPE/GEQIP.

3) ***Communication of project plans and implementation****.* A very long pipeline made delivering grants to school very complicated; sometimes officials along the way were not fully aware of what the funds were for and how to process them. This was another source of "false expectations." It would be advisable to shorten the pipeline if possible; also all actors along the way need to be fully briefed on the proper use of the funds and the proper ways of processing them.

4) ***Improvement of implementation capacity****.* The two Programs used different approaches to management (UNICEF has an almost government-like structure with eight bureaus in eight regions, whereas GPE/GEQIP relied on the Government of Ethiopia structure, but supported it with consultants). Both had their drawbacks: UNICEF witnessed a bloating of embedded staff whereas GPE/GEQIP faced the inefficiency of government structures. Program managers and planners can look at these different approaches and try to take the best and avoid the worst of both.

5) ***Establishment of professional monitoring and evaluation****.* Monitoring and evaluation was a big challenge for both Programs. For GPE there were challenges in defining the needs of marginalized groups and specifying performance indicators – including ones showing learning outcomes -- to reveal the change in such target populations. For UNICEF, this evaluation shows too little focus on outcomes and problems with tracking change over time (over the past two years there has been a growing "corporate" awareness of these issues and significant changes). Planning of solid, results-oriented M&E, involving experienced M&E specialists, should continue to be on both organizations' priority lists.

6) ***Sustainability of the programs****.* The legacy and continuity of the programs after their conclusion must be seen as an issue from the very start. In the current projects, there are worries as to how they will be sustained once the projects are over. The tightly organized GPE/GEQIP program does not appear to have a set of expectations for life at the end of 8 years of GEQIP. There appears to be an understanding among some that school grants will be sustained beyond GPE/GEQIP, but that needs to be more widely communicated and other sustainability issues dealt with (this applies equally to the World Bank, the GPE and UNICEF).

7) ***Efficient institutional arrangement****.* Development cooperation programs take place within a country's preexisting laws, rules, and regulations, which are often at odds with procedures and norms in the development agencies. Complex problems still arise from failure to confront this reality fully and openly. Consideration of ways to bridge such differences needs to be part of program preparation from the beginning.

8) ***Differentiated approaches***. Ethiopia is a decentralized country in economy, culture, geography and politics. Planning a program, as if the country were a centralized state, can easily lead to dysfunctionality. Participatory planning (see the first point) and a willingness to support different approaches in different regions are keys to success in this context.

I: Introduction: Objectives, methods, theory of change

1.1. *The Ethiopian Case study of the* Evaluation of Norway's Support to Basic Education through UNICEF and the Global Partnership for Education (Case no. 1100660), is one of four case studies conducted by the Development Portfolio Management Group (DPMG). The other three case studies being Madagascar, Malawi and Nepal. The main evaluation questions are: 1) What are the intended and unintended outputs and outcomes of the basic education initiatives that Norway's Ministry of Foreign Affairs (MFA) funds indirectly through two agents, GPE and UNICEF? and 2) What is the value-added for the MFA of using GPE and UNICEF as conduits for its investments? Annex 1 displays the Terms of Reference for this evaluation.

1.2. Two causal pathways are used to assess these questions: a) the research team's theory of change or causal path in Figure 1 below for improving three goals of interest to Norway's Ministry of Foreign Affairs (learning outcomes, gender equality, and equity), and b) the processes and quality assurance mechanisms that increase the probabilities of good aid management of the project/program cycle, diagrammed in Annex 5.

Figure 1: Theory of Change
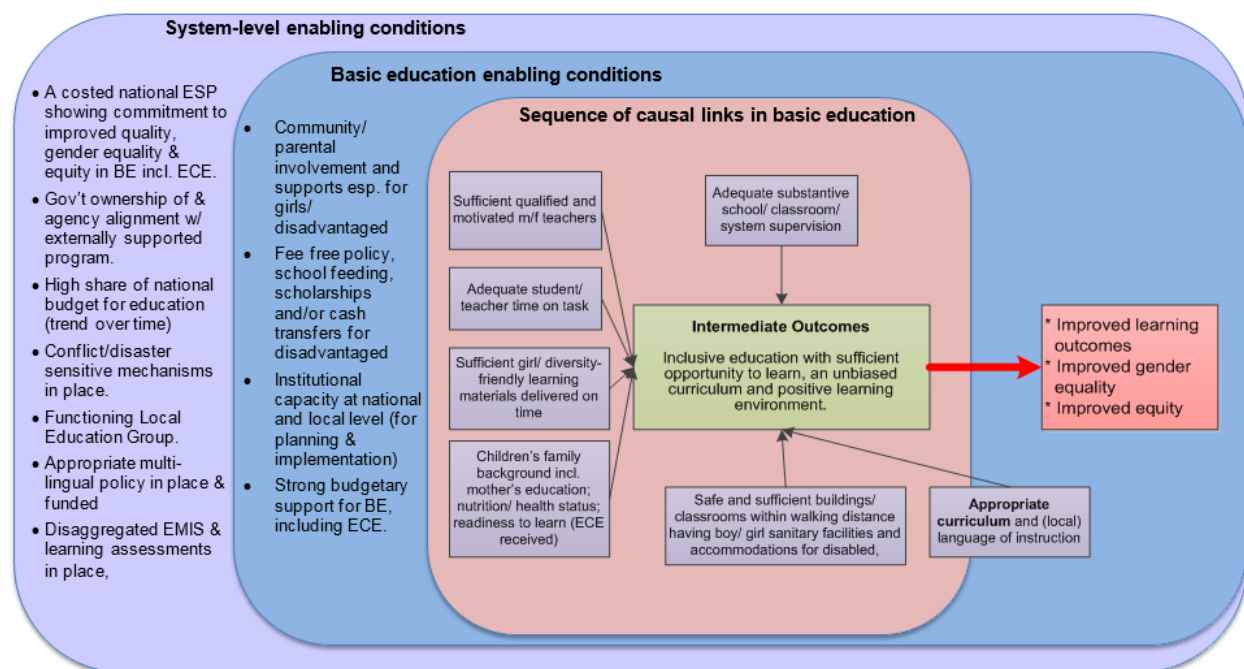
1.3. Each case study is based on multiple sources of evidence, including interviews with outside observers and the main parties connected to the GPE and UNICEF programs, such as members of the Local Education Group (including bilateral organizations and CSO members), supervising or managing entities for GPE programs, UNICEF staff, and Ministry of Education leaders and technical staff; and, where possible, school visits and school-based interviews. The fieldwork complemented and deepened the desk reviews of documents for each case. The case study teams consisted of two researchers, one senior and the other more junior, except in Nepal where they were both senior. Prior to the field work they were given the desk study results and invited to contribute to the field study design. The resulting field-study instrument, containing both standard items and customized ones, was jointly reviewed by the teams and the overall evaluation Team Leader during his one week visit to each country (except for Nepal) just prior to the field work, to assure full understanding or further adaptation of the items and inter-rater reliability. The Team

Leader participated in many of the early interviews in these locations to assure that data collection was conducted according to the agreed upon design and standards

1.4. The case studies assessed not only the outcome and contributing variables (interventions) in our Theory of Change, but also the contextual variables (enabling conditions) in the theory that can affect the outcomes and interventions, such as national and political commitment to basic education improvement and equity, the share of national budget going to education, institutional capacity, and private subsidies for basic education (e.g., free meals and scholarships). They also took into consideration various national strategic conditions such as the coup d'etat in Madagascar, Ethiopia's booming economy, Nepal's political turmoil, and, in Malawi, a major scandal at the Ministry of Finance that led to the suspension of aid by a number of donors. Each case study used process tracing, a method of checking whether the intervening steps in the theory of change were realized, to confirm the strength of the hypothesized pathways. This help establish whether inputs and outputs were related to outcomes through the processes outlined in the theory of change.

1.5. Qualitative and quantitative data sources were used. The assessment team reviewed documents, conducted interviews with participants/beneficiaries, and undertook direct observations at two schools (urban and rural). The analytical strategy involved triangulation of the perceptions of a purposive sample of 16 key informants from Ministry of Education (four offices), Local Education Group (LEG) members, including a representative of the Ethiopian Teachers Association, UNICEF Country Office and field staff, the World Bank (Addis Ababa office), as well as teachers at the schools visited.

1.6. The methodology had limits and potential biases. Educational outcomes cannot be directly or solely attributed to Government or donor-sponsored programs. Economic trends, natural disasters, and other factors can also affect outcomes. Each case study thus assesses not only the independent variables in the theory of change related to government and aid programs, but also the contextual variables (enabling conditions) in the theory that can affect outcomes, such as economic and political events, population growth, and governance. Each case uses process tracing, a method of checking whether the intervening steps in the theory of change were realized, to strengthen attribution of outcomes. This helps establish whether inputs and outputs were related to outcomes through the processes outlined in the theory of change. However, unlike statistical and experimental methods, process tracing cannot estimate the magnitude of the effects of each variable in the theory of change or of each enabling condition.

1.7. A second limitation is that the cases selected for study are not necessarily representative of the population of developing countries that received aid. Indeed, the case studies and desk studies were purposively sampled to include fragile states that pose difficult challenges for successful aid to basic education. Broad generalizations therefore cannot be made directly from any one case to the population of developing countries. It is possible, however, to make narrower, contingent generalizations from one or a few cases to subsets of cases that share key similarities. The challenges common to the several fragile states in the sample, for example, are more likely to be shared by other fragile states than by developing countries generally.

1.8. The case studies and desk studies are also constrained by the limitations of the available data. The programs being evaluated often tracked inputs and outputs, but they did not always have clear results frameworks, nor did they consistently measure baselines and outcomes. The programs being evaluated occurred some years ago (2009-2013). Some key players in the design and implementation of these programs, especially for the early part of the period from 2009 to 2013, were no longer available, or, if they were still in place, had to try to remember details. Respondents and authors of documents that were reviewed often had a natural interest in presenting events in a positive light. Where available, data from disinterested parties were triangulated with views that might be positively biased.

1.9. The report is organized into seven additional chapters. Chapter II summarizes Ethiopia's country and sector context. Chapter III describes the interventions of GPE and UNICEF programs during the review period. Chapter IV discusses the outcomes and project/program results. Chapter V discusses aid management, financial management and enabling conditions. Chapter VI discusses causal connections,

chapter VII identifies the lessons learned, and Chapter VIII focuses on the conclusions and recommendations resulting from the evaluation.

1.10. Qualitative and quantitative data sources were used. The assessment team reviewed documents, conducted interviews with participants/beneficiaries, and undertook direct observations at two schools (urban and rural). The analytical strategy involved triangulation of the perceptions of a purposive sample of 16 key informants from Ministry of Education (four offices), Local Education Group (LEG) members, including a representative of the Ethiopian Teachers Association, UNICEF Country Office and field staff, the World Bank (Addis Ababa office), as well as teachers at the schools visited.

II: Ethiopia: Country context

A. Strategic Context

2.1.   *Population*. Ethiopia is the second most populous country in Sub-Saharan Africa. In 2009, the total population of the country was about 81.9 Million whereas in 2013 it has grown to 94.1 million. The average growth rate of the population between 2009 and 2013 was 3%, with only a slight increase in the percent urban (2%). Infant mortality significantly decreased during that period from 81 per 1000 to 64.

2.2.   *Economy*. Ethiopia also has one of the fastest growing economies in the world, for instance, its annual growth rate reaching 10% in 2013, helping it to reduce those living below the international poverty line from 40% in 2004 to 30% in 2011. Despite this, it remains one of the poorest countries in the world with a per capita income in 2013 of only US$ 370.   At this point the country is still one of the largest donor recipients worldwide, receiving over $3 billion annually

2.3.   *Politics and conflicts*. The long term serving Prime Minster of Ethiopia, Meles Zenawi, passed away in 2012 after officially remaining in office over 17 years, and his successor took office peacefully. However Ethiopia is located in the Horn Africa, a region well known for its political instability.   A war with Eritrea at the turn of the century and civil wars in neighboring South Sudan and Somalia have brought large numbers of refugees' as well military tensions along its remote borders with these countries.

2.4.   *Regional disparities and volatility*. Ethiopia has a decentralized government structure where some regions are more prosperous than others. The less prosperous regions are mostly inhabited by ethnic minorities or pastoralists, and this sometimes leads to unrest. The main source of income across the country is smallholder agriculture, dominated by subsistence rain fed farming, using few inputs and characterized by low productivity. These farmers are extremely vulnerable to external shocks such as volatile global markets, drought and other natural disasters.[1]

B. Educational context

2.5.   *Structure*. The general education sub-sector in Ethiopia is made of up Preprimary, Primary, Secondary Education. The Preprimary education includes kindergartens ("schools") for 4-6 year olds mainly operated by NGOs, "zero" classes (coaching options at schools for 5-6 year-olds that have not had kindergarten to get them ready for first grade), and Child-to-Child Classes (older students interacting with preschoolers, stimulating more school readiness). Primary education is in two cycles, grades 1-4 and 5-8. Secondary Education is also divided into two cycles, general (grades 9-10) and preparatory (grades 11-12).

2.6.   *System Performance*.   Although there have been substantial improvements in opportunities for preprimary education, the sub-sector still remained small. In 2012/13, its Gross Enrolment Rate (GER) was 26.1% compared to 4.2% in 2008/09, which does show rapid expansion. At the primary school level, the overall Net Enrollment Ratio (NER) - male and female – increased for grades 1-4 from 88.1% in 2008/09 to 95.5% in 2012/13; and for grades 1-8 from 83% to 85.9% over the same period, a more modest increase. Increases for females were about the same as for males, but started out about 2-3 percentage points below

---

[1] Two examples show the government policy dimension of volatility. First, in 2003 the government adopted a resettlement policy to address rural poverty, moving 2 million farmers to new land, but they were not able to prosper there given unexpected challenges such as tropical diseases. Another possible source of social unrest is the government's 2011 announcement of a plan to build 4 large dams on the Nile, the largest of which (*Grand Renaissance*) would flood about 1,700 sq km of forest land near the Sudan border and displace about 20,000 residents. Its huge project cost (US$5 Billion) – likely to draw funds from other worthy sectors like education – and the tensions from dislocating so many local people (to say nothing of potential conflict with neighbor countries over use of scarce Nile waters), could lead to new rounds of turmoil.

them. The general figures mask large regional variations, e.g., in 2012/13 the primary GER Gambella was 126.6%, compared to 50% in Afar.

2.7. Enrollment ratios for secondary education were significantly lower than for primary, with NERs for lower secondary only reaching 19.4% in 2012/13 (from a base of 13.5% in 2008/09). The increase for girls (11.9 to 20.1 for that period was higher the overall average, with girls ending up with a slightly higher NER than boys. Using the common global definition of basic education as primary plus lower secondary, Ethiopia has a long way to go to before it reaches universal *basic education*. For upper secondary education, the NER stood at 5.2% in 2012/13 (5.2% males; 5.4% for males). Finally, the GER for higher education was recorded at 5.7% in 2012/2013, but only 30% of the enrollment was by females.

2.8. *Sector Plans and funding*. The Growth and Transformation Plan (2010-2014) provides the general framework for educational planning and development in Ethiopia. The Growth and Transformation Plan, as it is popularly known, is a national five –year plan created by the Ethiopian Government to improve the country's economy by achieving a projected Gross Domestics Product (GDP) growth of 11-15% per year from 2010 to 2015. The plan includes details of the cost (estimated at US$75–79 billion over the five years) and specific targets the government expects to hit at the end of the plan period. With reference to education and training, the plan endorses the goals of ESDP IV (Education Sector Development Plan IV). According to ESDP IV, the goal of the education and training sector over the life of the plan period is producing democratic, efficient and effective, knowledge based, inspired and innovative citizens who can contribute to the realization of the long term vision of making Ethiopia into a Middle Income Economy. Further, it focuses on educating and/or training a workforce that is demanded by industry, particularly the growing manufacturing industry, at all levels. ESDP IV has been developed to ensure equitable access to quality education at general, technical vocational education and training, and higher education levels; allowing these sub-sectors to have a strong linkage to, and interrelationship with, each other. The key objective over the five year period was to ensure the achievement of the educational MDGs

2.9. Ethiopia receives external funding support for its education from a variety of sources including from UNICEF and GPE. The multilateral support through GPE was initiated on 10 December 2008 and has become effective since January 2009. The agreement involved two FTI grants of US$78 million and US$90 million bringing the total US$ 418. This funding was packaged with IDA credit (General Education Quality Improvement Program I or GEQIP I), including as agency partners IDA (US$50m), Finland/Italy/ Netherlands/Sweden/UK (115M) and Ethiopia (US$84m). The life span of this agreement was January 2009 to 31 December 2013. (From now on we will refer to the GPE "project" as **GPE/GEQIP**). Similarly, UNICEF had a cooperative agreement with Ethiopia for a total of US$ 51,730 million in its country program for Ethiopia between 2007 and 2011.

III: GPE/GEQIP and UNICEF Program Description

A. GPE/GEQIP

3.1.   The overall objective of the GPE/GEQIP was to improve the quality of general education (Grades 1-12) throughout the country by providing a set of key inputs and interventions. Implementation was at both the federal level through the Ministry of Education (MOE) and in the 9 regions and 2 city administrations, which are responsible for management and financing of primary and secondary education. The developmental objectives of GEQIP I (2009-13) were to: (1) improve the teaching and learning conditions in primary and secondary education; and (2) improve the management, planning and budget capacity of the Ministry of Education and regional Education Bureaus of Education. GPE/GEQIP consisted of five components: (i) Curriculum, Textbooks, and Assessment; (ii) Teacher Development Program, including English Language Quality Improvement Program; (iii) School Improvement Program (SIP), including school grants; (iv) Management and Administration Program, including EMIS development; and (v) Program Coordination and Monitoring and Evaluation. The Program was designed to enhance coordination and synergy among components in order to maximize its effectiveness. While the project was primarily focused on the improvement of general education quality, improving the quality of schooling was expected to provide an incentive for the improvement of equity in access to quality education, especially for girls and disadvantaged children in rural areas; to ensure that children are promoted on time to higher grade levels (internal efficiency) and to ensure that children completed school (attainment) with quality learning outcomes. This project was also designed to contribute to poverty reduction and economic growth by supporting the quality and relevance of primary and secondary education. In short, GPE/GEQIP took a holistic approach to improving the quality of general education by adapting a widely-used school effectiveness model, which is many ways is similar to DPMG's Theory of Change.

3.2.   Concerning GPE/GEQIP interventions, the Program covered (in its Component 1): (i) revision of the curriculum for Grades 1-12; (ii) provision of textbooks and teacher guides based on the new curricula; and (iii) improvement of assessment and examination systems. The Second component (Teacher Development Program) supported the implementation of teacher educator and pre-service and in-service teacher development programs, focusing primarily on teachers in primary and secondary education, but also provided resources for the English Language Quality Improvement Program and the Alternative Basic Education program. Under Component three (School Improvement Program), schools completed self-assessments and developed school improvement plans and received school grants to implement their plans. The fourth component (Management and Administration Program) supported the institutional development of the MOE and regions, with a particular focus on enhancing the capacity of education officials to perform their roles, and strengthening the EMIS. Component five covered support to MOE and Regional offices for program coordination and monitoring and evaluation. GEQIP had a highly structured monitoring and evaluation system which was based on the results framework and specific performance indicators to measure both outcomes and determinants of those outcomes (including baseline, midline, and endline results).

B. UNICEF Country Programme (2007-2011)

3.3.   The key goal of UNICEF's country program has been to enable vulnerable children and women to progressively realize their rights to survival, development, protection and participation through support given to national and regional efforts to achieve the objectives of the plan for Accelerated and Sustained Development to End Poverty (PASDEP). This program had six components: (1) Young child adolescent and women's health; (2) Nutrition and food security; (3) Adolescent development, protection and HIV/AIDS; (4) Basic Education; (5) Water, Sanitation and Hygiene (WASH); and (6) Communication, gender and program coordination program.

3.4.   The aims of basic education revolved around policies and national standards to monitor school readiness in ECD and the development of learning programs, improvement of net enrollment in primary schooling in targeted *woredas*, improvement of gender parity, educational quality, and enabling children

under emergency conditions to continue their education. In the related WASH component the most important goal was that millions of children have equitable and sustainable access to and use of safe water, basic sanitation services and improved hygiene in drought-prone and underserved areas.

3.5. Results in terms of improved learning outcomes, gender equality and equity of marginalized groups were intended to be measured by performance indicators. With reference to learning/quality/improvement, four performance indicators were given: Grade 4 and 8 student achievements in NLA (National Learning Assessment), Primary School Completion Rate, Average Grade 4-8 repetition rates and establishment of model Child Friendly Schools. For NLA assessment the specific indicator of success was the proportion of students in Grades 4 & 8 who scored at least 50%. The target for primary school completion rate was set at 85% for 2011. The target for average Grades 4-8 repetition was planned to be 2% for 2011. (Most of these indicators and targets were to apply to UNICEF's 300 targeted districts).

IV: Outcomes and project/program results

4.1.   This section of the case study addresses directly Norway's Ministry of Foreign Affairs' three priority outcomes of improved student learning, improved gender equality and improved equity. These were also articulated as the main outcomes variables in DPMG's Theory of Change. The team's approach to them was to determine what outcomes (and outcome indicators) GPE/GEQIP and UNICEF included in their programmes that related to these three priority areas, and then to assess whether and how effectively the outcomes were reached.

A. GPE/GEQIP

i.   Priority Outcomes

4.2.   *Student learning*. GPE/GEQIP placed a high priority on student learning outcomes. Its main performance indicator for this was "percentage of students scoring at least 50% on National Learning Assessment in Grades 4 & 8." The following table shows the results on this indicator at the beginning of the Project (baseline) and at the end (2013).

4.3.   Table 1 shows that the overall student learning achievement at the endline increased significantly both for males and females relative to the baseline. The percentages are still not high (at the endline 70% to 80% were still not reaching 50% at the two levels, but the positive change for the middle to high achieving students is quite rapid (10 to 20 percentage points per year).

Table 1: Percent of grade 4 & 8 students scoring at least 50% on NLA (baseline and endline)

Gr	m/f	Base line	Endline*	Percent Change
4	F	20	28	40%
	M	21	29	38%
	Total	20	29	44%
8	F	6.1	16	163%
	M	14	20	43%
	Total	10	18	80%

*2012-2013

4.4.   *Gender equality*. The second priority outcome was gender equality. Table 1 was originally created to demonstrate improvements in **learning outcomes,** but, since it is disaggregated by gender, it can also be used to indicate improved **gender equality**. The table reveals grade 4 boys and girls improving at about the same rate and that gender equality has already been reached at that level. A the grade 8 level girls were performing at a very low level at baseline, especially relative to boys, but due an amazing 160% increase they came much nearer to parity with boys at the endline..

4.5.   The gender equality outcome was also measured by using the GPE/GEQIP indicator of primary school completion rate (Grades 4 and 8). This indicator, which usually measures internal efficiency, is used here as a proxy indicator for gender equality by the study team. As shown in Table 2 the results are disaggregated by gender, and provide both baseline and endline figures. The findings show both males and females improving over time, but the percent change for females is much – 5X – higher; for grade 8 female improvement relative to males' is higher still (15X). For both grades the differential rates of improvement bring females completion rates to parity with males by the endline results – an indicator of achieved gender equality. These result were confirmed during our visits to schools in Addis Ababa and Amhara district, that showed significantly higher enrolment rates for girls than for boys, which was

considered by one of the principals to be the result of a big push for gender equality starting about 4-5 years ago (for more details see Annex 4).

Table 2: Percent of grade 4 & 8 students completing primary school (baseline and endline)

Gr	m/f	Baseline	Endline*	Percent Change
4	F	56	74	32%
	M	69	73	6%
	Total	63	74	17%
8	F	33	52	58%
	M	50	52	4%
	Total	43	52	21%

4.6. *Equity*. The third priority outcome, equity for marginalized groups, was indirectly addressed by the GPE/GEQIP and there was *no indicator* in the Project for it. As mentioned above, there was an expectation that improved school quality would act as an incentive for girls and marginalized children to enroll (or re-enroll) in school, but the Project could be criticized (and was) for failing to assess whether this happened. Ironically, the data needed for this was being collected by the Ministry all along (published annually in its *Education Statistics Annual Abstract)* and this was **district breakdowns** of primary school enrollment and completion data, see Table 3.

Table 3: Regional breakdowns of primary NER (2008/9 and 2012/13)

Region	NER (%) 2008/2009 Girls	Total	NER (%) 2012/2013 Girls		Total
Tigray	98.1	96.9	93.1		92.0
Afar	23.2	24.4	44.9		41.5
Amhara	103.1	102.2	93.1		91.7
Oromiya	74.8	77.9	80.6		83.9
Somali	29.4	31.6	80.8		81.8
Benis-Gumuz	80.1	88.6	84.4		91.6
SNNPR	84.5	89.4	84.9		88.9
Gambella	69.7	75.2	93.2		98.0
Harari	83.6	91.9	69.1		74.8
Addis Ababa	74.4	76.1	67.8		69.4
Dire Dawa	70.2	73.4	81.3	90.5	85.9
National/Total	**81.3**	**83.0**	**87.7**	**84.1**	**85.9**

Source: *Federal Democratic Republic of Ethiopia Ministry of Education, Education Statistics Annual Abstract*
2008/2009 (2001E.C); 2012/2013 (2005 E.C)

4.7. The "emerging regions" in the country are the pastoralist areas of Afar and Somali, Benishangul-Gumuz and Gambella. As the above table shows, in all those regions there were improvements between 2008/09 and 2012/12; in fact, in the two most disadvantaged regions, Afar and Somali, the improvements were vast (in Afar the total percent increase was 70% and in Somali an amazing 160%*). (In Benishangul-Gumuz and Gambella they were a more modest 4% and 30%). This compares very favorably with the average change in NER at the national level, which was a modest 3%, with the levels even going down in some more advanced regions. In one of the emerging districts, Afar, the improvement for girls was even greater than that for boys – substantially so -- which, if the data are reliable, would be a good example of girls *overcoming multiple disadvantage* (e.g., being a girl in a disadvantaged area), something that should be examined more closely.

ii.   Other outcomes and unintended consequences

4.8.   *Other Outcomes*. The GPE/GEQIP did not specify any other outcomes besides those that were aligned with the Norwegian MFA's priority areas.

4.9.   *Unintended consequences*.   During field study, government officials, managers of the GPE project and the UNICEF team were asked to identify unintended consequences of education aid programs. The responses specifically for GPE/GEQIP were the following:

- The attitude and practice of the community towards contributing to schools were strengthened because of the GPE/GEQIP school grant program;

- The central government was motivated to increase its education budget over 2009 to 2013;

- The effective approach to leadership and management in GPE/GEQIP resulted not only in the improvement of management in general education, but also influenced leadership and m management in other education sub-sectors including higher education.

- Implementing agencies (regional education bureaus, and *woreda* education bureaus and institutions) believed that GEQIP covered all educational expenses;

- Concerns about the sustainability of project benefits and outcomes beyond the life of GEQIP have emerged among stakeholders

4.10.   The above findings show that GPE/GEQIP was perceived to have brought many positive unintended consequences. These include particularly an heightened interest within the community to contribute to school income, a decision by the central government to increase its funding levels for education, and the improved leadership and management in general education as well as in other sub-sectors including higher education.   On the negative side, unexpectedly the program created "false expectations" among some stakeholders possibly because of insufficient awareness building among them. If Regional Education and Woreda Education Bureaus developed "false expectations," it suggests that the main actors or owners of the GPE/GEQIP were very top down in their project development and planning approach. Another indication of faulty planning and communication were the worries about sustainability expressed by some.   This concern implied that GPE/GEQIP's interventions were considered very useful and acceptable to some stakeholders, but, at the same time, worries about sustainability grew up which should have been dealt with and communicated about at the initial planning stage. Some respondents claim that GPE/GEQIP was developed in collaboration with regions and MOE, but at the same time admit that planning and communication could have been organized better.

iii.   Perceptions of positive program effects

4.11.   *Both agencies*. In an open-ended interview question in field study interviews key informants were asked to indicate the biggest ways in which UNICEF and GPE had a "positive effect on basic education in their countries."   In responses both GPE and UNICEF were praised for having programs **based on the countries' strategic development plans and education sector plans**. Both were also recognized for **their focus on the priorities of the Ministry of Education**. The Ministry of Education was impressed by how both used participatory planning approaches and started from regions and districts (despite this, some false expectations grew up in both programmes).

4.12.   *Specific to GPE/GEQIP*. The following positive effects were noted:

- **Its focus on the Ethiopian education sector's highest priority: improving the quality of education and student achievement.** GPE/GEQIP's approach to doing this was to bring resources right down to the school level: school grants, new curriculum and textbooks, teacher development – all linked with improved teaching and learning at the school level.

- **The pooling of a sizable amount of funding from many agencies, which required -- and in fact resulted in -- carefully designed aid-management systems, including good coordination between the government and donors, solid financial management, joint supervision, and serious monitoring and evaluation.**

- **The skillful GPE/GEQIP blend of existing government structures and a technical advisory team for the implementation of the project**. From government, GPE/GEQIP I implementing agencies were MOE Directorates, 9 Region & 2 City Administration Education Bureaus with their *woreda's* and schools, 36 teacher training colleges, 10 Universities responsible for pre-service and in-service teacher program, all under the effective coordination of the Planning and Policy Analysis Department of the MOE. The technical advisory team of short and long-term consultants, specializing in project management, financial management, procurement and M&E, was hired and resided in MOE providing regular support to the implementing agencies. At the regional level institutional arrangements mirrored those at the federal level. Funding for schools followed a long pathway through federal, district, and woreda offices before reaching schools and colleges, but the system kept it flowing.

4.13. From the above presentation, **it is clear that GPE's perceived positive effect on basic education has to do with five points: clarity and alignment of program purpose; well organized program planning and management; use of existing government potential and structure; winning political support and system wide sense of ownership; and use of professional technical assistance**.

iv.   Upstream contributions

4.14. As part of its overall quality improvement strategy GPE/GEQIP was also designed to make upstream contributions at the systemic level. The most important ones were: a) system design **for quality assurance mechanisms** at the school level (e.g., school self-evaluation and report cards); b) a **school inspection policy** (framework and guidelines); c) systems for **licensing, relicensing and recruitment** of teachers); d) **school grant mechanisms**, e) a **gender and equity needs assessment**; and f) contributions to the drafting of the **Education Sector Development Plan IV**.

v.   GPE/GEQIP Intervention Results

4.15. Our Theory of Change sets out a set of school or school system variables /determinants that it postulates will together make a difference in the attainment of desired learning outcomes (in particular improved student learning, improved gender equality, and improved system equity). The first one (which we label an **intermediate outcome**) describes in general the kind of learning environment that is needed the produce these outcomes; the others are contributors to the creation of such an environment, as follows:

4.16.   Intermediate Outcome: Inclusive education with sufficient opportunity to learn, an unbiased curriculum, and a positive learning environment.

4.17.   Contributors:
- Safe and sufficient buildings/classrooms within walking distance having boy/girl sanitary facilities and accommodations for disabled;
- Sufficient qualified and motivated m/f teachers;
- Sufficient girl/ diversity-friendly learning materials delivered on time;
- Appropriate curriculum and (local) language of instruction;
- Adequate student/ teacher time on task;
- Adequate substantive school/ classroom/ system supervision;

- Children's nutrition/health status, readiness to learn (Early Childhood Education (ECE) received);

- Community/ parental involvement and supports esp. for girls/ disadvantaged.

4.18. The DPMG team determined that it was not practical to observe each the two projects/programs to see if they had these exact determinants in place, after all they were not explicitly designed to do so. The team decided therefore to specify the *general theme* of each of these determinants and observe what the projects/programs were doing in these thematic areas. The closer to our ToC version of the theme they came, the closer they were to implementing this particular theory of change.

4.19. The generic thematic areas (slightly adapted to field conditions) were:

- Schools and classrooms

- Teachers

- Curriculum

- Learning materials

- Language of instruction

- Teacher/school/system supervision

- Changes in the learning environment

- Child's background and resources (included early childhood education)

- Community and parents

4.20. This section shows what the GPE/GEQIP provided/implemented in these thematic areas and summarizes how well the "interventions" went.

Box 1: GPE Interventions

Schools and classrooms. The Program was organized to be comprehensive in its coverage of educational inputs/conditions that were conducive to educational quality improvement, but also looking globally across Ethiopian education development programs it was determined not to include features already covered by other programs. Schools and classrooms was a case in point: it was determined that this was covered by the "Promotion of Services" program funding by the Ministry of Education.

Teachers. GPE/GEQIP did focus seriously on enhancing the capacity of basic education teachers. Here are the performance indicators and results:

Performance Indicators: (a) the proportion of teachers with appropriate qualifications, Grades 1-8, and (b) policies developed on teacher training, professional development, licensing and relicensing to increase motivation.

Results: (a) grades 1-4 baseline 3.4%, endline 30.4%; grades 5-8 baseline 53.4%, endline 90.8%; (b) teacher policies for training, professional development, licensing and relicensing were developed.

Adequate student/teacher time on task. Not explicitly addressed by GPE/GEQIP and UNICEF.

Curriculum. This was also an emphasis of GPE/GEQPI with the following indicators and results:

Performance indicators: (a) percentage of schools/alternative basic education centers (ABECs) receiving curriculum orientation, (b) curriculum policy development.

> **Results:** (a) target 100%; achieved 100%; (b) based on poor Early Grade Reading Results it was decided that all mother tongue language textbooks (in 7 languages, grades 1-8) needed to be rewritten using more up-to-date pedagogy, which was done during GEQIP with support from USAID, with printing and distribution covered by GEQIP (done under GEQIP II).

Learning materials. This was also an important component of GPE/GEQIP, covering not only huge increases in numbers but also improved quality, based on curricular revisions and updating.[2]

> **Performance indicators**: (a) Pupil to new textbook ratio for primary core subjects (goal 1:1)

> **Results**: Endline pupil to textbook ratio for primary core subjects: 1:1.3, exceeding the target.[3]

Teacher/school/system supervision.

> **Performance indicators**: (a) training of principals (head teachers) and administrators through "Leadership and Management Training Program," (b) the percentage of woredas (districts) and regions reporting on a quarterly basis their physical/financial performance.

> **Results**: (a) 24,000 head teachers and bureaucrats trained compared to the targeted of 6,000;

> (b) *No data was available to assess this indicator.*

Changes in the Learning Environment. This intermediate objective was included in the program as follows

> **Performance indicators**: (a) the increased use of student-centered instruction (b) the percentage of schools/ABECs using school grants for school improvement program (SIP).

> **Results**: (a) there was an improvement in the use of student-centered methods, exceeding the project target; (b) school grants were used by 98% of schools and ABECs, compared to the goal of 90% of schools and 50% of ABECs. (This was confirmed by a visit to 2 schools, one in Addis Ababa and one in Amhara District, which revealed that use of grant funds had contributed to the learning environment of the schools.)

Child's background and resources. This is not a typical project input (although much research suggest that it should be) and did not form part of the GPE/GEQIP intervention.

Community and Parents. This became an important part of the GPE/GEQIP intervention:

> **Performance indicator**: PTAS/School Boards manage SIP program/operational budgets according to the guidelines

> **Results**: The targeted proportion of parent-student engagements and student–student meetings took place (75% for each). School visits to two schools confirmed the active participation of SIP committees and PTAs in the utilization of the funds according to guidelines; SIP committees were meeting at least every two weeks.

---

[2] The intervention also covered secondary school textbooks and teachers' guides (gr 9-12) but this evaluation only focuses on grades 1-8.

[3] On site assessment in 2013 showed that only 65% of student had the needed textbooks, but this was at least partially accounted for by the fact that many children leave their textbooks at home and some textbooks were not distributed until late 2013.

B. UNICEF country programme

i. Priority Outcomes

4.21. *Student learning*. A high proportion of grade 4 and grade 8 students attaining a 50% or better pass rate on the National Learning Assessment (NLA) in 300 targeted districts was a Country Programme *key result* for 2007 to 2011. Annual Reports actually gave composite scores, not pass rate results, and did not give figures for the 300 districts but for the nation as a whole. The scores (around 39% for both grades compared to the 85% target) were only reported for one year (2007), such that improvement could not be determined.[4] The other UNICEF indicator aligned with improved learning was the number of model Child Friendly Schools established. Even though this is an *output* indicator not an *outcome*, it does give some sense of the Programme's contribution to school quality. The goal for the 5 year country program was 1000 (see key results indicators for 2007) and actual number established was 895 (annual goals being exceeded once, in 2011). These figures do not tell us what proportion of primary schools in UNICEF target areas were converted to CFSs.

4.22. *Gender equality*. UNICEF was very committed to and involved with the improvement of gender equality, but did not establish strong outcome indicators for that. The best indicator was the gender parity index, which was to be assessed in the 300 target districts and to reach a level of 1.0 by 2011. The figures reported in the Annual Reports were country-wide and were for only 2010 and 2011. They were 0.93 and 0.96 respectively. The baseline was not clear, but the goal for 2007 was set at 0.87, so it would have been no higher than that, showing some progress.

4.23. *Equity*. More equitable access by and treatment of marginalized or vulnerable children were very important UNICEF goals, as succinctly stated in its Country Programme goal, but Country Programme documents did not specify any outcome indicators for these. There were, to be sure, important UNICEF interventions for vulnerable children (e.g., Alternative Basic Education Centers, and pre-school education and development programs) and important upstream work (e.g., ECE policies), and these will be described in forthcoming report sections.

ii. Other outcomes and unintended consequences

4.24. *Other outcomes*. In addition to addressing the three MFA priority outcomes the UNICEF Country Programme for Basic Education (2007-2011) address other main objectives, namely: a) developing/ producing policies and standards for Early Childhood Education; b) providing educational services to school children who are displaced due to emergencies (target: reaching at least 60% of them). For the first objective, UNICEF reports to have fully achieved the objective; for the second, there was no information about percent reaching in the Annual Reports, but it was reported that over 100,000 children were helped to continue their education despite floods and other emergencies in 2009 and 2010.

4.25. *Unintended consequences.* The case study informants mentioned the following unintended consequences for the UNICEF Base Education programme.

- UNICEF's involvement in so many areas of activities and its high visibility have created high expectations of its support throughout the Ministry and the country (the agency is seen as both credible and rich);

- UNICEF has become in the development agency community and to the government the "agency of first resort;"

---

[4] GEQIP used the same indicator and gave baseline and endline figures, which did show improvement over time (more or less the same time period as the UNICEF Country Programme period), so this case study relates UNICEF interventions to those national assessment findings.

- More careful oversight and scrutiny of financial contributions have led UNICEF to tighten the conditions for the disbursement of funds, leading some recipients to conclude that UNICEF has become less efficient than before; that it does not process funds "fast enough."

4.26. UNICEF in Ethiopia is a thinly spread organization both in terms of geography and the number of activities in which it is involved. It has 11 regional offices and several tasks to implement in 300 *woredas*. Its field presence and its wide scope are partly what has created the "first resort" image. Yet, it is obvious that it has limited resources, and that it could not escape from false expectations that it is a rich institution capable of providing answers to all problems.

4.27. Another unintended consequence (not noted by the informants, but observed by our team) of UNICEF'S being spread so thin – covering both wide-ranging field activities and upstream work – was not being able to give a full accounting of its results in its annual reporting (one respondent even mentioned its lack of a full set of good performance indicators for measuring program results from year to year). This lack of clarity about program outcomes is inevitably one of the reasons contributors are asking for better oversight and scrutiny of funds. Moreover, respondents noted that UNICEF budgets are falling even as it continues its wide-ranging agenda.

4.28. Here's one MOE opinion:

> For UNICEF the budget has been reducing year by year. Because of this, the activities planned under UNICEF program are minimized. Appropriate budget allocation or prioritization and more focused support based on the available budget would be essential.

iii. Perceptions of positive program effects

4.29. UNICEF is perceived to have had the biggest positive effect by reaching the unreached through ABE and in vitalizing Early Childhood Development (ECD). It was recognized that UNICEF has strengthened ABE by providing training for ABE facilitators and constructing centers, supporting the four emerging regions' implementation capacity-building programs, and recruiting and paying salaries for 14 capacity building experts. Moreover, UNICEF was acknowledged to have played a major role in the vitalization of Early Childhood Care and Development Programs in Ethiopia, through its drafting of ECD policy statements and standards (e.g., for assessing "school readiness") and designing and operating various ECD programmes, such as "child to child" and connecting ECD programmes to Child Friendly Schools. Also important was the awareness raising, community mobilization, and training programs in support of gender equity (e.g., helping the MOE to prepare manuals and materials on gender).

4.30. It was pointed out that UNICEF has supported an important study on out-of-school children, drop outs, and has developed outreach strategies for them based on recommendations. Some informants praised it for being widely participatory (although others do not – see "Lessons Learned"). There is a perception among education sector leaders and senior bureaucrats that the government and this UN agency can work together successfully. UNICEF's contribution to emergency education, its decision to transform its Child Friendly Schools model into something closer to the School Improvement Plan (SIP) approach, and its interventions in improvement of quality of education have also been mentioned among the agency's most important positive effects.

iv. Upstream Contributions

4.31. UNICEF informants are confident that many basic education innovations including education policies, regulations, laws and sector plans (or particular sections of them), and support for vulnerable children would not have happened or would not have happened to the degree they did had UNICEF not been there to push them. To accomplish this, UNICEF took the lead in many of the "upstream" efforts. Sometimes this upstream work has meant conducting research (like that on the out-of-school children), sponsoring study tours (like those to Iraq and Nigeria for a view of ECD programs), or sponsoring national gatherings, like the National Conference of Pastoralist Education. Among the most important upstream

contributions that UNICEF reported would not have happened -- or happened to the degree that they did -- without its support:

- Early Childhood Care and Development (policy formulation, strategy development, evaluation, curricular standards and guidelines probably would not have happened in government and could have stayed private.

- Alternative Basic Education (standard setting; support to out-of-school children)- probably would have happened but would not have expanded so far or have been so refined; now covered by EMIS

- Strategies and policies for education of children in pastoral areas - helped to germinate ideas through UNICEF sponsored study tours

- Mainstream gender equality concerns (both in the Ministry of Education and Ministry of Finance) - Almost all development agencies working on this, but UNICEF was more proactive than many; conducted studies, supported girls education forum.

- Contributions to the Education Sector Strategies through playing key roles among agencies- It was a coordinator among partner agencies and had a very strong voice and support in latest Education Sector Development Plan (ESDP).

- UNICEF was/is involved in every chapter of the ESDP V, as a member of many working groups and as co-chair of the Ed Tech Working Group in 2015.

- Provided support for formulating ABE standards Pastoralist Education Strategy

- Gender mainstreaming guideline for MOE and Gender in education Strategy.

- UNICEF initiated and supported a research on out of school children and based on it the Ministry of education developed guideline.

- Provided special support to the development of Mobile school in Pastoralist Communities Implementation Guideline

4.32. This upstream work has always meant UNICEF working within the Government goals, objectives and strategic plans. In fact, most policy and strategy documents the government produces do not have UNICEF's logo on the cover, but the MOE & its Directorates will readily acknowledge UNICEF's important upstream contributions.

v.   UNICEF Intervention Results

4.33. The key goal of the UNICEF' Country Programme in Ethiopia 2007-2011 was to enable vulnerable children and women to progressively realize their rights to survival, development through support given to national and regional efforts.   In that sense, it is very much aligned with the MFA's priority concerns of gender equality and equity for marginal groups.   The activities that it supported during the period were supportive of that goal. As with GPE/GEQIP to understand them in light of our Theory of Change we will sort the activities according to the themes that were listed under the GPE/GEQIP analysis (also see next paragraph).

4.34. Few of the Country Programme activities that relate to these themes had indicators with target values; and in few cases were any baseline values specified or needs analyses cited. Moreover, many of the indicators used were not specified in advance, instead many emerged as the programed moved along, often in response to annual reviews. Thus, the reports give a sense of what was delivered, but does not convey whether output met existing needs or represented a meaningful level of change. (More comments about this will be taken up in the Aid Management section of this report.) Also, the Programme plan often indicated a focus on 300 targeted districts, but it is not clear from the annual reports when this was the case.

Box 2: UNICEF Country Programme Interventions

Schools and classrooms. Under the Programme the following were constructed: separate latrines for girls in 199 schools; 879 ABEC centers (2010-2011); 4 boarding schools for girls, and 10 mobile schools for pastoralists. In addition to ABEC center construction, 1490 new centers were attached to existing buildings (2007-2009).

Teachers. There was a small amount of teacher training for ECD centers in 2009.

Curriculum. UNICEF helped design the learning programme for ECD centers.

Learning materials. Teacher Learning materials were provided for ABEC learners and equipment provided (at least at newly constructed sites in 2010 and 2012); some materials and facilities were provided for learning through Braille.

Language of instruction. There was nothing in the Programme re language of instruction

Teacher/school/system supervision. There was supervisor training at 62 schools; regional training on planning and management for 13,000 district and *woreda* managers; and limited email training at the central, district and school levels.

Changes in the learning environment. Over the Programme UNICEF reconfigured its Child Friendly Schools model into a system based on School Improvement Programs (SIP); in areas of low female educational attainment the Programme set up tutorials for vulnerable girls.

Child's background and resources (included early childhood education). UNICEF drafted ECD policies and standards (2007-2008); in 2009 the government endorsed a new ECD policy; ECD centers, either connected to existing schools (sometimes CFSs), or using the child-to-child approach, or introduced as a class "zero," were introduced in 474 locations, often with learning materials and facilitator training

Community and parents. Training of at least 5000 community members on Harmful Traditional Practices; community members and PTAs were training on developing SIPs; managing school grants.

4.35. This list of interventions shows what a large contribution the UNICEF country program made during 2007-2011, especially to communities of disadvantaged learners. While the agency has clearly been busy, the data give little indication as to how effective it has been, since these raw numbers give no indicator as to how close the agency has come to solving problems and filling gaps. Also of note is the showing that UNICEF covered some parts of our ToC model quite extensively and other parts scantily if at all (perhaps a reflection of its orientation to improving opportunities for disadvantaged learners). Comparing this list of interventions to that of the GPE/GEQIP it is apparent that they are quite complementary -- one addressing a theme quite extensively while the other almost not at all. More will be said about this in the causal connections section.

V: Aid management, financial management, and enabling conditions

A. Aid management

i.   GPE program

5.1.   Several factors determine the quality and thus the value added of an agency's aid management. Some are enabling conditions, such as how the agency's activities are financed. However, proximate factors are whether the agency's governance results in clear accountabilities and whether its management of the project cycle meet quality criteria. The schematic of the project cycle and questions that reveal the quality with which it is implemented is presented in Annex 5.

5.2.   **GPE.** Several parties are accountable for the quality of GPE's management of donor aid: the GPE Board and its committees; its executive arm, the Secretariat; the Ethiopia LEG; and the supervising entity (World Bank) for GPE's Project to Improve Education Quality in Ethiopia (PIEQM) (2010-2015). GPE has several unresolved governance problems that muddy accountabilities. Obviously, the performance of Government also defines the parties' opportunities for effectiveness, at least in part. However, the issue here is not Government's implementation performance, but how well the GPE players adjusted to the obstacles and possibilities presented by Government.

5.3.   In terms of quality assurance for grant applications to GPE, the **Secretariat** faces an incomplete contractual relationship. It is accountable for quality oversight, but can only rely on persuasion to obtain any redesign of the project on the basis of the findings of the quality review. In the Ethiopia case the Secretariat commissioned an External Quality Review[5] that was highly professional. The Review flagged not just strengths, but also weaknesses that could undermine the success of the GPE project. Examples of issues of concern were attention to equity and program complexity. From what the Panel could tell, key performance indicators were not designed to pick up two types of variability: a) ethnic, linguistic and socio-political variability that results in differences in educational demands; and b) variation in the success of the implementation of the GEQIP across regions and *woredas*—variations that can result in unacceptable inequities in the distribution of services and materials.

5.4.   The Panel flagged project complexity as portending implementation problems downstream. Even Phase I of the GEQIP was a large and complex program with 5 components, 12 sub-components, and 26 sub-sub-components. Some of the sub-sub-components were straightforward—e.g., a study, but most involved quite challenging activities. GEQIP was being implemented in a decentralized environment that involved several levels of government (federal, regional, zonal, and *woreda* levels), schools and their PTAs, universities, and Colleges of Teacher Education. Specifically, project funds would flow to the MOE, 11 Regional Education Boards (REBs and city administrations), 21 universities, 26 Colleges of Teacher Education, and about 22,900 schools (including ABECs). On paper the powers, implementation responsibilities, and lines of accountability of each player were clear. However, the capacity to plan, manage and monitor was weak, especially at the regional and *woreda* levels that are responsible for managing and financing primary and secondary education. Compounding the insufficient numbers of qualified staff was the constant turnover of key staff.

5.5.   **Local Education Group.** In Ethiopia the LEG is called the Education Technical Working Group (ETWG). The forum is co-chaired by the Ministry of Education and a lead donor on a rotational basis and meets on a monthly basis. Its membership consists of about 15 pooling and non-pooling donors in the education sector, but discussion is underway about adding civil society organizations active in the education sector. The Ethiopian Teacher's Association is now participating in the forum. NGOs sometimes participate in the Joint Review Missions and the Annual Reviews. Agencies use the forum to discuss the Government's

---

[5] In 2009 the GPE Secretariat proposed and the Board agreed to pilot independent reviews of the quality of the design of projects prior to their being submitted to GPE for funding. These were called External Quality Reviews, conducted by education experts experienced in development.

Education Sector Development Plan (ESDP) and priorities, to provide technical support for the sector, to conduct joint review, and share information. The forum is also used to create synergy and to do "donor mapping".

5.6. ***ETWG members rate the functioning of the ETWG positively.*** Table 4 summarizes the interview data on the performance of the ETWG relative to its functions. Overall, the data reveal an actively engaged ETWG/LEG and a satisfied membership, with only minor changes proposed to improve functioning. The data show a surprising consensus among respondents on which functions the ETWG performs, their frequency, and the respondents' satisfaction with how the ETWG discharges these functions. In response to a question about how effectively members of the ETWG collaborate with each other, interviewees uniformly rated the collaboration as effective.

Table 4: Interview Responses about Functions of the ETWG and members' satisfaction

Function	Engage in function	Frequency	Satisfaction rating
Sector planning and inter-agency coordination	Yes	Once a month	Satisfied
Donor harmonization and use of common procedures	Yes	NA	NA
Harmonize donor support around Government's sector plan?	Yes	NA	NA
Pool funding?	To some extent	NA	NA
Agree on common processes and procedures to reduce the burden on Government?	To some extent	NA	NA
Rely increasingly on Government systems for implementation, financial management, and procurement?	To some extent	NA	NA
Supervision of the education sector plan and projects designed to reach its goals	Yes	Annual	Satisfied
Policy dialogue and consensus	Yes	Varies. Depends on stage of planning.	Satisfied

5.7. ETWG members were unable to rate the performance of the Supervising Entity, the World Bank. Although EFA/FTI had funded 40% of GEQIP I, the respondents noted that ETWG's annual Joint Review Missions monitored aspects of the implementation of the Government's overall sector plan, not GEQIP, which was monitored by the funding partners.

5.8. ***ETWG did not appraise the Education Sector Development Program (ESDP)***. The ETWG apparently did not appraise the version of the ESDP on which the GEQIP is based, although it did re-endorse it. Prior to 2012, the GPE Secretariat did not systematically recommend that the Development Partners conduct an Education Sector Plan (ESP) appraisal and endorsement before engaging on each new phase of a long-term education strategy. In the event, the quality of the ESDP was mixed. The ESDP is comprehensive, addressing each sub-sector and cross-cutting initiative (e.g., gender and education or school health and nutrition). It is systematic, specifying for each sub-sector and cross-cutting program area expected program outcomes, key outcome targets, and a program matrix with component activities and indicators. There is a meta-matrix of key performance indicators, baseline data (except for student/textbook ratios), and targets by year for the overall ESDP.

5.9. The relationships between the activities, outputs, and outcomes are generally logical and plausible. The targets are the primary flaw. Many are questionable or implausible. For example, the percent of grade 1-8 teachers with a diploma is expected to increase from the 2009-10 baseline of 38.4% to 94.6% by 2014-15. The expected increases in the percent of students scoring at least 50% and at least 75% on core

subjects of the National Learning Assessment in a four year period are optimistic in the extreme. For example, the percent of 4th graders achieving at least 50% in the NLA core subjects in four years is expected to increase from a baseline of 41% to 75%. The percent expected to achieve at least 75% in the core subjects is expected to increase from 1.7% to 25% in this same time period. The learning outcome targets for grades 8, 10, and 12 are similarly optimistic. GPE's new financing model, introduced in 2015, where a share of the grant will be "held hostage" to demonstrate progress on ESP (not project) targets, will impose implausible targets on ESPD.

5.10. The ESDP's M&E framework for the plan is well structured with a calendar of annual joint review missions and Annual Review Meetings that will focus attention on the ESDP's results framework and pressure the implementing units to generate credible M&E data. Although the ESDP identifies country and sector risks well, estimates of the likelihood that risks internal and external to the sector will materialize are vague and poorly analyzed. It offers no mitigating strategies to deal with sector-specific risks.

5.11.   Supervising entity (World Bank). GEQIP is a two phase project-- GEQIP I and GEQIP II. Based on an analysis of GEQIP documents, the quality of design is rated moderately unsatisfactory; the quality of supervision, satisfactory.

Project preparation. The quality of the design of the GEQIP, despite some excellent aspects, was moderately unsatisfactory primarily because of unresolved problems with its results framework and the unmitigated complexity of the operation relative to Government's capacities. To its credit, the Bank tried to scale back the operation, but unsuccessfully. Government resisted attempts to size and phase activities in accord with its capacities, as evidenced by its very ambitious procurement plan for textbooks.

5.12. The results framework had multiple problems that obscured the picture of the achievements of GEQIP I.   Although the Bank approved three restructurings for the program, these all involved aligning closing dates for the EFA-FTI trust fund with the IDA closing date.   None was used to restructure the development objective (DO) indicators.   The midterm review was used to restructure the intermediate outcomes and intermediate outcome indicators, but by the last Implementation Status Report (ISR), some of these had shifted back.

5.13. The DO for the overall GEQIP is clearly stated, as is the DO for GEQIP I. The DOs overall and for the GEQIP I are well linked to the four key challenges identified in the PAD. However, often the causal path between activities, interim outcomes and their indicators, outcomes, and the DO cannot be traced or are not plausible.

5.14. The majority of the outcome and intermediate outcome level indicators were selected from the agreed Education Sector Development Program (ESDP) core sector performance indicators. However, the DO indicators used to measure the DO for the overall GEQIP and the DO for GEQIP I have serious relevance problems.   For example, outcome indicator 4 is stated as: "Consistent federal government commitment to education budget – a) Education expenditure as % of total government; and b) General education expenditure as % of total education budget." Achieving these outcomes is outside of the control of the sector and the GEQIP.

5.15. With one or two exceptions, the outcome indicators are measurable, and the intermediate outcome indicators had baselines.   The DO indicator targets for the overall GEQIP and GEQIP I are well selected. They are considerably more realistic than the ESDP IV, reflecting the slow pace at which some of the variables measured tend to change.   The intermediate outcomes and their indicators shift around during program implementation, are too numerous, and their scope is too detailed.   They vary from 11 to 14 intermediate outcomes and 18 to 23 indicators, obscuring the larger picture of how the program is progressing. They have a number of relevance problems and some measurement problems.

5.16. Since GEQIP I is complex, the results framework is complex. Thus, the M&E burden is significant and at risk of delays and data credibility problems. At the same time, since the majority of the outcome and intermediate outcome level indicators were selected from the agreed ESDP's core sector performance indicators, the M&E arrangements for GEQIP I build on country systems. M&E uses several data sources:

a) annual education statistical abstracts; b) education sector financing data from MoFED; c) annual and biannual reports consolidated from implementing agencies; d) independent evaluation, reviews and audits; e) comprehensive evaluation of school grants; and f) annual and biannual review missions conducted by the Development Partners. The planned timing of M&E data collection seems generally appropriate and realistic.

5.17. Compounding the scope problems with the GEQIP, Government tended to undermine efforts to build capacities. Historically, the government of Ethiopia had initially agreed to TA at the design stage and then resisted contracting the planned and needed project-funded TA. Of further concern was that in July 2009, just months prior to submitting the Project Paper to the Catalytic Fund, the MOE was restructured. The entire GEQIP team, including the Coordinator, was replaced. The MOE still did not have a full time Coordinator and had not yet hired several key consultants (including a procurement consultant) when its funding proposal was submitted to the Catalytic Fund.

5.18. *Project supervision.* Supervision was fully satisfactory. Given problematic aspects of the design, it is to the credit of the Government, the World Bank as supervising entity, and the pooling Partners that the project succeeded as well as it did. In 4 ½ years the government of Ethiopia disbursed 99.6% of almost US$436 million out of $437.7 million. As noted below, the ICR determined that most program outcome targets were achieved.

5.19. The Project was under close supervision throughout its lifetime. The World Bank and the pooling Partners conducted regular and frequent joint missions on a quarterly basis during the early phases of implementation and bi-annually later on. These missions provided implementation support to the MoE on various aspects of implementation, including procurement, financial management, and M&E. They reviewed the implementation of substantive components and provided expert technical support, adding technical specialists to the task team as needed. Field visits were conducted as part of the supervision missions, which allowed the Bank Task Team and pooling Partners to obtain firsthand feedback on the challenges faced by implementing agencies and from beneficiaries.

5.20. The biannual Implementation Status Reports give numerous examples of actions taken by the World Bank and pooling Partners to resolve implementation and partnership problems. For example, the December 2010 mission determined that the MOE was again proving reluctant to hire the technical advisory services required to implement the project as designed. The mission conveyed that if this trend continued, the project would have to be redesigned so that it could be implemented with existing capacity. This would require a sharp contraction in activities and funding. By May 2011 the project had turned around, with the State Minister himself becoming involved in the project and chairing monthly GEQIP implementation review meetings.

5.21. *Evaluation of GEQIP I at completion.* The World Bank conducted its required Implementation Completion Report after GEQIP I closed. It was thorough and satisfactory, finding that, with the exception of the Teacher Development Program, most program outcome targets were achieved.

ii.   UNICEF program

5.22. UNICEF's project cycle, the quality of which partly defines the quality of UNICEF's aid management for Ethiopia, starts with the United Nations Development Assistance Framework (UNDAF) for Ethiopia. The UNDAF structures aid activities by the large number of UN agencies active in Ethiopia. These include UNICEF, UNESCO, UNDP, UNAIDS, WFP, FAO, WHO, UNFPA, and UNHCR. UNICEF's Country Program Document (CPD) for 2007-2011 includes support to basic education that is aligned with 2007-2011 UNDAF priorities. UNICEF's Country Program Action Plan (CPAP) for 2007-2011 elaborates on themes in the CPD. UNICEF reports on progress against the CPAP annually.

5.23. **Project design.** Project design falls between moderately unsatisfactory and unsatisfactory. In support of Ethiopia's Education Sector Development Program III (ESDP III), UNICEF's CPAP focuses its basic education program on three projects: 1) Quality and Girls Education; 2) National Capacity Enhancement; and 3) Innovative and Complementary Education. It seeks these outcome targets for its

basic education program for 2007-2011: (a) policies and national standards to monitor school readiness in ECD and learning programs developed and implemented; (b) NIR in primary school increased from 51 per cent to 96 per cent of schools in targeted *woredas*; (c) gender parity achieved in targeted *woredas*; (d) improved quality of education achieved in targeted *woredas*, with 85 per cent of children retained and attaining at least 50-per-cent pass result in grades 4 and 8; and (e) 60 per cent of school-going children who are displaced in emergency situations able to continue their education.

5.24.   From the upstream documents it is hard to discern the design of the components or activities that UNICEF planned to use to achieve these objectives. The CPAP should translate the CPD's approach into operational activities, but its discussion of its basic education program remains at the strategic level, with statements such as these:

> "Institutions at all levels will be strengthened through training of education officials in planning, supervision, monitoring and reporting. UNICEF will also provide education materials, appropriate equipment and technical support to the decentralisation process in consultation with MoE. Duty bearers, including PTAs, will be sensitised to meet their obligations through communication, community dialogues and teachers training. Evidence-based advocacy and social mobilization will be utilized to create awareness on the importance of girls' education and achieving gender parity. Regions and woredas with low female participation in education will be supported to accelerate gender disparity reduction and improve education quality through: strengthening school cluster initiatives; supporting Girl/Child Friendly Schools; UNGEI (including a school WASH programme) and promotion of school, community and university linkages. Focus will be placed on issues of disabilities and diversity in school with policy advocacy, promotion of cost-effective inclusive approaches and support for national capacity to integrate the concerns of children with disabilities into the schooling system." (UNICEF CPAP for 2007-2011, p.17)

5.25.   The CPAP's results framework is organized around the three projects. However, the activities listed under a "project" are quite disparate and cannot be construed as a designed component. For example, just 3 activities listed under "National Capacity Enhancement" are "gender sensitive school improvement plan developed and implemented in 3000 schools of the 300 *woredas",* "federal, regional and 300 *woreda* education offices increase non-salary budgetary allocation of primary education through advocacy", and "all primary school children in 300 UNICEF targeted *woredas* have access to context-specific curriculum and teaching learning materials".

5.26.   UNICEF's results framework for the 2007-2011 program has to be patched together between upstream documents. The UNDAF results framework focuses more on outcomes, baselines, and targets; the CPAP, on outputs, indicators, baselines, and targets. However, UNICEF's theory of change or causal pathways from activities to outputs to outcomes cannot be constructed from the upstream documents or from the composite results framework.

5.27.   Some outcome targets seem implausible--for example, the net intake rate, or NIR, to grade 1 will increase from 60% in 2005 to 96% in 2008 The grade 1 dropout rate is projected to drop from an (undated) baseline in the UNDAF of 22% to 2% by 2008. The primary school completion rate is expected to increase to 85%, against a baseline of 52% in the 2007 CPAP. These may be targets sought by the Government of Ethiopia in its ESDP III, but, even if they are, UNICEF should have examined their plausibility before adopting them as their accountabilities for their 2007-2011 plan.

5.28.   **Supervision.** Supervision reporting falls between moderately satisfactory and moderately unsatisfactory, in good part because annual reports could not report against a clean and coherent results framework. Annual reports included helpful information, such as the resources used; the most recent Government data on outcomes; the activities pursued and outputs achieved; factors or constraints affecting performance and lessons learned; and monitoring, studies, and evaluations. The list of studies and comments on their results is particularly useful--for example, UNICEF conducted a Study on the Quality of Alternative Basic Education in Ethiopia, finding that the ABE program is deficient in quality and recommending a learning assessment that compares ABE and formal education students.

5.29. However, targets shifted between the CPAP and annual reports. For example, under "key milestones and expected results" the CPAP expected that by the end of 2007 the NIR would be 81% and 96% by the end of 2008. In the 2007 annual report, the target date for reaching 96% had shifted to 2015. The school completion target of 85% in the CPAP had changed to 80% by the 2010 Annual Report. Some of the inferences about the effects of outputs seemed unwarranted.

B. Financial management

5.30. ***Public Financial Management Systems***. The most recent Public Expenditure and Financial Accountability (PEFA) assessment for Ethiopia was conducted in 2010 for the fiscal year ended in 2009. When Ethiopia's public financial management (PFM) system is analyzed from the perspective of the six PEFA dimensions, the system in general performs well and the overall performance improved between the 2007 and 2010 assessments; the identified areas for improvement have been or in the process of being dealt with, which will strengthen the country's PFM system and better support the budgetary outcomes. The PFM of Ethiopia has recently been updated from the previous Integrated Budget Expenditure to a more comprehensive Integrated Financial Management, piloted in six federal ministries. With the successful completion of the pilot, the roll-out phase will commence and eventually include not only the central government but also the regions and the *woredas*.

5.31. ***Budget Credibility***. The 2010 PEFA performance assessment of the credibility of the federal budget indicates that the overall execution of expenditures during the years 2006/07-2007/08-2008/09, (the last years for which data is available), was good; the actual amounts spent for those years were not very different from budgeted amounts. The rating (based on Public Expenditure and Financial Accountability Program standards) was unchanged at B+ from the levels achieved in 2007.

5.32. ***Predictability and Control in Budget Execution***. In comparison to 2007, performance on taxes and customs improved slightly. In 2010, legislation for most major taxes was comprehensive and clear with taxpayers having access to up-to-date information on tax liabilities. Tax payers are registered in a data base, penalties for non-compliance exist, and tax audits are performed. However, there are significant tax arrears which prompted the establishment of the Ethiopian Revenue/Customs Authority. Cash flow management and forecasting for expenditures at the Federal level have greatly improved in 2010 in comparison to 2007, giving budget institutions a greater horizon for planning and committing expenditures. Debt data quality and reporting have also improved (slightly) and there are good payroll controls systems in place. Controls in public procurement are satisfactory, including a good procurement complaints mechanism. PEFA also shows that overall performance of internal audit has improved between 2007 and 2010. The rating for the Predictability and Control in Budget Execution remained unchanged from the 2007 levels at C+.

5.33. ***Fungibility and/or Additionality***. For the education sector, not only did the average variance between actual expenditures and the original budget increase from the 2007 to the 2010 assessment, the increased variance was largely due to the deviations in capital expenditures on the education. The large increase in the total budget variance between these two assessments is partially explained by the significant increase in the price of oil in 2007/08 which forced reallocations of the budget throughout government.

5.34. ***Ratio and Trend Analysis***. The following table demonstrates a generally positive trend in support to the education sector over the 4 fiscal years ending 2013; it also appears to demonstrate Government's commitment to the sector.

Table 5: Ratio and Trend Analysis

Category	2009/10	2010/11	2011/12	2012/13
% of national budget allocated to education	23.8%	24.9%	23.9%	22.5%
% of national recurrent budget for recurrent expenditures in education	29.5%	30.5%	31.6%	32.4%
% of education budget allocated to recurrent expenditures	56.9%	53.0%	54.6%	58.6%
% of GDP allocated to education	4.7%	4.5%	3.4%	3.3%

5.35.   The percentage of the total national budget and the percentage of the total national recurrent budget allocated to the education sector shows consistent levels of budget being allocated to the sector, and with regards to the recurrent budget that level of support has been increasing. Over these four years, the total budget has increased by approximately 106% while the total education budget has only increased by 95%. While this data could be interpreted as an indication of declining support for the sector, a more realistic interpretation would seem to be that other sectors were receiving increased levels of capital spending. This later interpretation is supported by the fact that the recurrent budget for the education sector has increased by about 100% over these four years while the national recurrent budget has only increased by about 83%.

5.36.   As indicated elsewhere in this report, the education sector's share of the average expenditures for developing countries was 17.3% in 2012. For Ethiopia, if the 6% budget variance between budgeted amounts and actual spending that occurred from 2007-09 continues, the sector's share of total expenditures will be well in excess of this reported average figure, and thus indicate Government's support for the education sector. The area which would seem to indicate a declining level of commitment to the education sector is with regard to the percentage of the GDP allocated to the sector. As indicated in the above table, the percentage of the GDP allocated to education has declined steadily over these four years from 4.7% to 3.3% in 2013. While this decline is disappointing, the decline may be due more to the fact that the overall economy has increased by approximately 175% over these four years than to a reduced commitment.

C. Enabling conditions

System-Level Enabling Conditions	Basic Education Enabling Conditions
• National/political commitment to improved quality and equity in BE, including ECE • High share of national budget for education • Conflict/disaster sensitive mechanisms in place • Functioning Local Education Group • Appropriate multi-lingual policy in place & funded • Disaggregated EMIS & learning assessments in place	• Community/ parental involvement and supports especially for girls/ disadvantaged • Fee free policy, school feeding, scholarships for Disadvantaged • Institutional capacity at national and local level (for planning and implementation) • Strong budgetary support for BE, including ECE

5.37.   Above are the enabling conditions postulated by the DPMG ToC, covering both system-level conditions and those specific to the basic education sub-sector.   The team's strove to collect information about all of these conditions, although in some cases information was not available.

i.   System-level enabling conditions

5.38.   *National/political commitment to improved quality and equity in BE, including ECE.* Federal Government commitment to equitable, high quality education is very strong. In 1994 the country issued its basic policy for education and training which emphasized issues of access, equity, quality and relevance. Since then as succession of five year education sector plans (ESDPs) have reinforced those goals; moreover, in 2011 the country released its overarching Growth and Transformation Plan which also emphasized improved (equitable) access, and in particular improvement of learning and the conditions influencing student achievement. For disadvantaged regions the government has been taking a holistic

approach spearheaded by a National Taskforce;[6] for preschool children, ESDP IV committed the country to expanded and improved ECCE.

5.39. *High share of national budget for education*. Over the period 2009/10 to 2012/13 the share of the national budget allocated to education was consistently between 22.5 and 25 percent.[7] Percent of GDP allocated to education declined from 4.7% to 3.3% during that period, the latter figure being almost at the 3.5% recommended by GPE.

5.40. *Conflict/disaster sensitive mechanisms in place*. The country is classified by the UN as a "fragile state," largely because of regional instability and frequent natural disasters. The country works with regional institutions like the Intergovernmental Authority on Development to prevent conflicts and launches military actions to neutralize possible spillovers of conflicts to its territory. For its frequent natural calamities, particularly drought, it has developed a National Commission for Preparedness and Prevention of disasters. In addition, UNICEF co-leads, with Save the Children, an "Emergencies Cluster" in Addis Ababa, which has helped the Ministry prepare for and direct relief to those who are caught up in disasters and/or conflicts.

5.41. *Functioning Local Education Group (called in Ethiopia the Education Technical Working Group or ETWG).* The evaluation team's survey in Addis Ababa shows a surprising consensus among respondents on which functions the ETWG performs, their frequency, and the respondents' satisfaction with how the ETWG discharges these functions. In response to a question about how effectively members of the ETWG collaborate with each other, interviewees uniformly rated the collaboration as effective ("1" on a 4 point scale). Asked how the ETWG could improve relative to these variables, the changes proposed were minor. Overall, the data reveal an actively engaged ETWG/LEG and a satisfied membership.

5.42. *Appropriate multi-lingual policy in place & funded*. The country's 1994 Education Language Policy recognizes the right of children to get instruction in their own mother tongues in primary education. The policy confirms Amharic the working language of government and English language of international communication. Learning these languages begins in primary school. Mother tongue instruction have been in 23 of the country's 77 local languages so far. Early grade reading assessments (in 2010) showed reading weaknesses and the government acquired external funding to help it create new reading textbooks in 7 languages.

5.43. *Disaggregated EMIS & learning assessments in place*. The country has a functioning Education Information Management System which has been producing a timely Education Statistics Annual Abstract since 1999. The abstract produces disaggregated data by gender, urban/rural and regions, and in recent years has been delivered within the reporting year.[8] Also, Ethiopia is one of the few countries in Africa that has regular, institutionalized learning assessments (e.g., the National Learning Assessment - NLA).

5.44. *Summary. This analysis was conducted against a backdrop (see Strategic Context) of strong economic* growth (reaching 9-10% in recent year), political stability, and relative peace and security. This helps to account for its unusually strong system-level enabling environment for improved basic education outcomes and equity, in which it has: a) enacted policies and sector plans over the past 20 years supportive of improved quality and equity in basic education, including – in ESDP IV -- ECCE; b) consistently devoted a high proportion of its national budget to education (between 22-25% over the past 5 years; the Sub-Saharan average being c. 18%); c) put into place disaster relief mechanisms which are supported by the donor community; d) participated in the establishment and satisfactory functioning of a Local Education Group; e) promoted – starting 20 years ago -- the acquisition of reading skills in the mother tongue,

---

[6]From 2008 a national taskforce chaired by the Prime Minister's office, and involving relevant ministries including MOE, Ministry of Heath, Ministry of Water and Energy, etc. was formed to provide technical assistance, monitor, and support the basic service provision including education for pastoralist communities.

[7]The budget share for education of 24.4% in 2010 compares favorably to the Sub-Saharan African average for that year (18%) was roughly equivalent to that in neighboring Kenya (23.8%), but around twice as high as that in Djibouti (12.3%), Uganda (11.6%).

[8] UNICEF has also been supporting work on EMIS at the district level; nevertheless, EMIS results are not used nearly as much as they could be for planning and decision-making.

strengthened recently by improved texts and instructional methods; and f) established an EMIS which has produced, since 1999, timely annual "abstracts," having disaggregations for gender, urban/rural, and district.

ii.   Basic Education Enabling Conditions

5.45. *Community/ parental involvement and supports, especially for girls/ disadvantaged.* Traditionally communities have helped with constructing, rehabilitating or upgrading schools, but this put the low income communities at a disadvantage, so the current ESDP has encouraged communities, e.g., through their PTAs, which are generally active, to support causes with fewer financial requirements, like decreasing absenteeism, school governance, and working on SIPs and related school grants. Girls' clubs were formed and other after school programs to encourage remaining in school.

5.46. *Fee free policy, school feeding, scholarships for disadvantaged.* Public education is fee free at the primary level, but families do incur expenses for uniforms, book rent, registration, etc. Also in some *woredas* a *community contribution* is collected with the provision that those who cannot pay will not be excluded. School feeding is practiced in 6 food insecure regions and this was planned for expansion in ESDP IV; also some girls and other disadvantaged have been assisted through scholarships, boarding and mobile schools.

5.47. *Institutional capacity at national and local level (for planning and implementation)*. Recent decentralization reforms have shifted significant education responsibilities to local *woredas* where newly empowered officials need strengthening. Developing institutional capacity at that level has been one of the most important goals of ESDP IV, although some strengthening was also found to be needed at the federal and district level.

5.48. *Strong budgetary support for BE, including ECE.* In 2007/08 the share of the education budget spend on primary education was 46%, on the low side of the GPE/FTI norm at the time (42-64%) (Public Expenditure Report, 2010). It was planned to go down to 37% during ESDP IV.

5.49. *Summary*. Our analysis of the enabling conditions in the Theory of Change showed areas of strengths but also some that were inclusive or questionable. On *community/parental involvement* the history shows a tradition of helping with buildings and having active PTAs with recent *redirection* towards School Improvement Plans. There has been some mobilization of community groups in support of girls' education, but it is not widespread. On school fees, education is fee-free through primary but there are some hidden costs (uniforms, book rent, registration, etc.) and some *woredas* authorize the collection of community contributions (but those who can't pay cannot be excluded). Some girls and other vulnerable groups have been assisted with scholarships and boarding /mobile schools but this is not yet widespread. Although some capacity development was still found to be needed at the federal and district levels, at the *woreda* level, where decentralization had brought new competence requirements, capacity development become an urgent priority under ESDP IV. As for percent of education budget for primary education, the figure was 46% in 2007/08 (on the low side of the GPE/FTI norms), and was planned to be around 37% during ESDP IV.
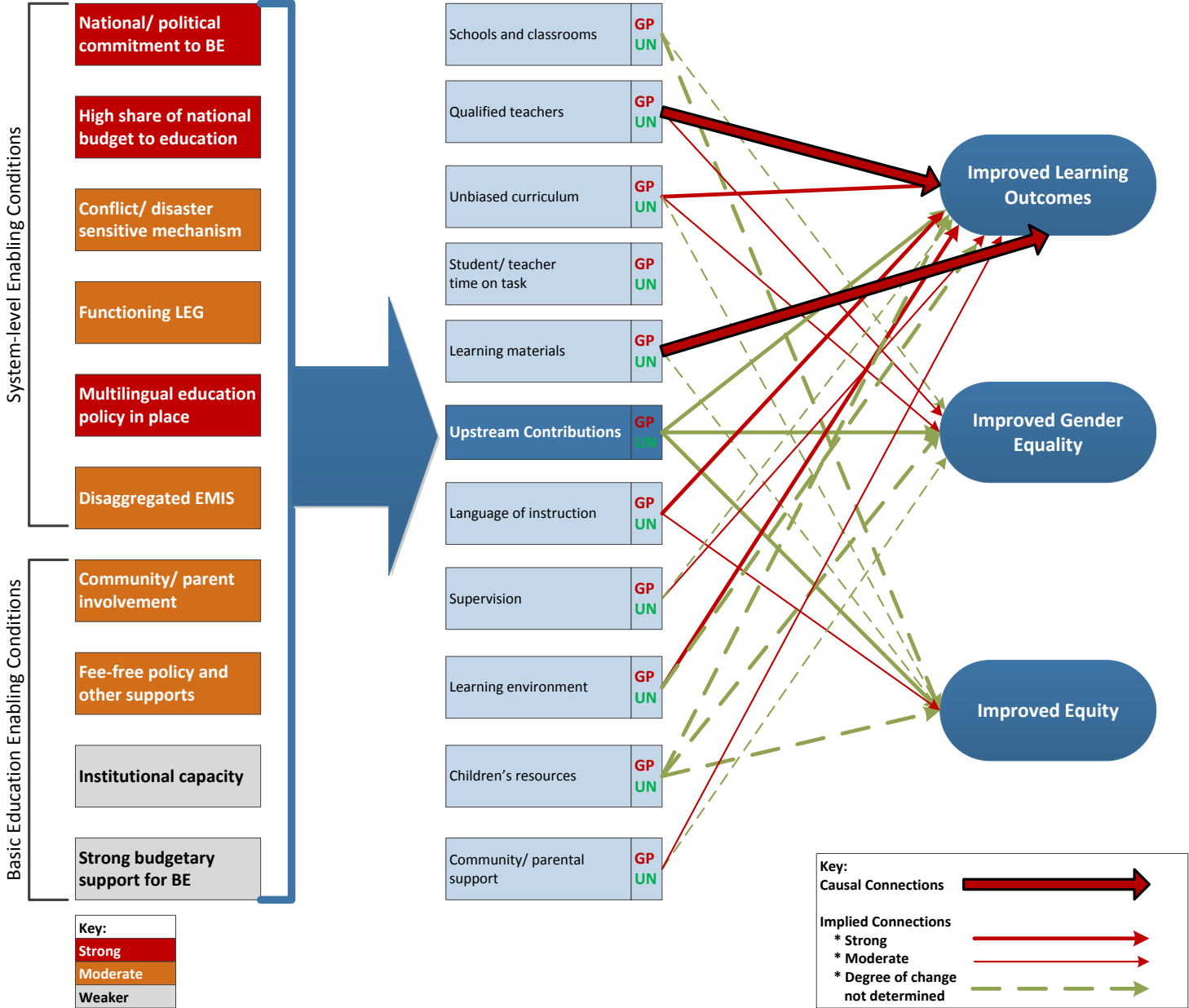
VI: Causal Connections

6.1.   Drawing from our Theory of Change, we created a figure to illustrate the connections between program interventions and program outcomes, and between those and the enabling conditions. Figure 1 shows lists the enabling conditions on the left, the interventions in the middle and the three "priority" outcomes on the right. Since neither the programs nor this evaluation used an experimental or quasi experimental design we are not able – with 2 exceptions -- to draw empirically-derived causal lines among the variables; instead we have drawn lines of "implied connections," by which we mean connections that were intended by the designers and/or which can be inferred logically by us, the evaluators. The elements in the figure are color coded, with the enabling conditions code according to their perceived (by us) strength in this national context, red for "strong," orange for "moderate," and gray for "weaker." For the interventions (the "themes" for our ToC) we line them up in boxes and then put small designation in the right hand corners of GP (global partnership, also colored red) and UN (UNICEF, colored green). The arrows from the interventions to the outcomes show the implied connections based on our observations, with the thicker arrows show relative strong implied connections and the thin one showing moderate ones. The two red arrows that are outlined in black show empirically-derived causal connections. The dotted lines for most of the UNICEF connections are to show that although the implied connections are there, the "degree of change" in the intervention could not be determined (given the lack of baseline values and/or change scores) – the exception to this are the UNICEF upstream contributions to which this does not apply.

6.2.   Looking to the left, our findings are that half of the six "system-level" enabling conditions were strong and the other half moderate. For the "basic education" conditions there were no strong ones: instead two were moderate and two were weaker. Our informants indicate there that there continue to be capacity constraints especially at the local (woreda) level and that funding for basic (k-8) is lower than the GPE standard. This suggests that effectiveness of the basic education improvement programs are constrained by low level capacity at the district level and tight funding.

6.3.   Concerning the connections between interventions and outcomes, first it can be seen that an important variable in our model, "student/teacher time on task," was not seriously addressed by either program. Also, there is some evidence of complementarity between the two programs, with GP providing more support for learning outcomes (in fact it has only two narrow arrows connected to equity), and UNICEF more for the equity ones, and, at the intervention level, GP covering teachers, learning materials and language of instruction, whereas UNICEF has supported more children's resources (ECD), schools and classrooms (especially ABECs). Both Programs made upstream contributions, but UNICEF went both to learning outcomes and equity, whereas GP went mostly to learning outcomes. Of course, the two causal arrows in the schemata are those determined by the GP for qualified teachers and learning materials, which were provided on a massive scale.

Figure 2: Enabling Conditions

Schools and classrooms — GP / UN
Qualified teachers — GP / UN
Unbiased curriculum — GP / UN
Student/ teacher time on task — GP / UN
Learning materials — GP / UN
Upstream Contributions — GP / UN
Language of instruction — GP / UN
Supervision — GP / UN
Learning environment — GP / UN
Children's resources — GP / UN
Community/ parental support — GP / UN

Improved Learning Outcomes
Improved Gender Equality
Improved Equity

System-level Enabling Conditions

National/ political commitment to BE
High share of national budget to education
Conflict/ disaster sensitive mechanism
Functioning LEG
Multilingual education policy in place
Disaggregated EMIS

Basic Education Enabling Conditions

Community/ parent involvement
Fee-free policy and other supports
Institutional capacity
Strong budgetary support for BE

Key:
Strong
Moderate
Weaker

Key:
Causal Connections

Implied Connections
* Strong
* Moderate
* Degree of change not determined

28

VII: Lessons Learned

7.1.   Different questions were forwarded to respondents in order to widely explore the lessons learned from implementing both GPE/GEQIP and the UNICEF country programs. The first question referred to the ways in which UNICEF and GPE could be more effective in promoting basic education in Ethiopia. The second question said "If you could re-allocate UNICEF and GPE funds spent on basic education in Ethiopia but had to stay in the same budget limit, where would you add funding and where would you reduce it?" The third question was about ways in which UNICEF and GPE need to adapt better to Ethiopia's history and culture regarding education.

7.2.   Responses to these questions implicated that several lessons have been drawn from implementing the two programs. Some of these include the need for programmatic and geographic area focus, the need for monitoring and evaluation, the need for capacity to implement change, the need to understand diversity and structural arrangement.

7.3.   *Need for program and geographic focus*. Respondents indicated that, while both program were relevant to Ethiopian problems they were both thinly spread. For instance, with reference to UNICEF a respondent said:

> *UNICEF could have done a better job supporting basic education, if it had stopped spreading itself so thin – focused on a few areas (both thematically and geographically (cutting down to 120 woredas from 300 in all 11 regions) (this is also a priority at the global level)… focused more on the least advantaged, like pastoralists… and in Oromia, help with a new initiative to design ABE for the second cycle of primary…*

7.4.   Similar views were aired about the GPE/GEQIP program, suggesting that it might have done better to concentrate more on the School Improvement Program.

7.5.   *The need for professional Monitoring and Evaluation*. The other most important issue was the need for proper monitoring and evaluation of projects. A respondent referred to GPE/GEQIP, saying the following:

> *There were some ways to make GEQIP I more effective and this are already under consideration in GEQIP II. There was no outcome indicator in GEQIP I for improving education provision for vulnerable groups (special needs, gender equality, pastoralist children).*

7.6.   Another respondent was referring to UNICEF's shortcomings in monitoring and evaluating its results:

> *Keeping the continuity under some intervention and trying to look the outcome it has achieved is essential, rather than simply discontinuing interventions. For instance, capacity building experts were recruited by UNICEF to support the emerging regions, but without evaluating what outcome this brought, it was stopped and now the Ministry took it under GEQIP.*

7.7.   *Need for Implementation capacity*. Low implementation capacity was one of the biggest challenges faced by GPE/GEQIP [see an earlier point about the recent decentralization to the *woreda* level), according to one informant. In order to minimize the challenge the project employed a team of technical consultants, and yet local capacity levels continued to pose challenges to the project.

7.8.   *Need for a differentiated approache*s. Some informants commented on the programs' lack of understanding of the wide diversity among schools in Ethiopia, as follows:

One formula for all schools has been used for school grants [amounts based on student numbers]. These similar models across the board didn't help schools in geographically disadvantaged areas since they have small numbers of students. The government doesn't like the idea of piloting.

7.9. ***Planning institutional arrangement***. An incompatibility of UNICEF systems with those of the Government has been reported, a perceived lack of flexibility in UNICEF's systems of finance, procurement and recruitment:

UNICEF's budget releasing system should be flexible. The experience is that now because of the liquidation problem of one single activity all the budget cannot be released and the planned activities will be postponed to another quarter. Because of this the ministry is now shifting some important basic education components to another donors and GEQIP II.

7.10.   Another interviewee took up the same subject in elaborate manner emphasizing the need for UNICEF to change and adopt the country's procurement system:

…it would be good [for] UNICEF to adapt its system to the country's procurement policy, such as by lengthening the liquidation period to 6 month…

7.11.   The most important lesson here is the need for more effective initial planning. It is well known that the two parties have different procurement systems. It would have been better to accommodate those differences before project start-up than to get stuck with them in the middle of implementation.

7.12.   ***Effective communication and timely budget release***: In GPE/GEQIP, one issue had to do with effective communication about and timely release of the school improvement funds to schools:

To make the program more effective, communication and budget release have to be improved. Sometimes there were communication irregularities and so utilizing the [best] existing communication means would [be helpful]… Unless schools are informed prior to starting the fiscal year about budget allocation and transfer, it is not possible for them to make appropriate academic planning or preparations.

7.13.   *This* concern is not astonishing given the long journey that money has to transfer to reach the schools. It is not surprising that one semester budget takes a year to arrive.

VIII: Conclusions and Recommendations

A. Conclusions

8.1.   The main purpose of this evaluation is to "generate evaluation evidence on the relevance, efficiency and effectiveness of Norwegian aid to basic education through UNICEF and GPE, focusing particularly on the achievement of Norway's policy objectives quality of learning, gender equality and equity..." This case study, covering Ethiopia, is one of the four that was to be conducted. It comprises a desk study plus live data collection during a field study period lasting about three weeks.

8.2.   *Outcomes*. Findings in Ethiopia reveal substantial improvements over a five year period (2009-2013) on all three of Norway's priority areas: improved learning outcomes, gender equality and equity (e.g., for marginalized children). On learning outcomes, the main indicator used (employed by both GPE and UNICEF) showed substantial improvements especially for girls, who in grade 4 attained parity with boys and in grade 8 almost did so. This indicator also showed movement towards gender equality. Another principal indicator of this, also used by both organizations, was percent of grade 4 & 8 completing primary school. The rate for girls was the same as for boys in the end, but to get there their rate of change was **5** and **15** times higher, for grades 4 & 8, respectively. Neither of the Programs had good indicators for "equity," but the national EMIS provides district breakdowns on *net enrollment ratios* and *primary school completion.* On the former, improvements in the two most disadvantaged regions, Afar and Somali, were revealed to be much higher than in the rest, and within them girls' rates improved more than boys', which suggests the overcoming of **compound inequalities** (those often experienced by girls in disadvantaged regions). **In sum**, by the indicators used by the GPE and UNICEF and/or by this evaluation, desirable changes in the priority outcomes were observed.

8.3.   During the period 2007 to 2013 both UNICEF and GPE/GEQIP mounted significant programs to influence such outcomes. Their interventions were mainly consistent with those in our team's Theory of Change, but with a certain unexpected complementarity: GPE/GEQIP covering some ToC themes (e.g., teacher qualifications and textbooks), and UNICEF others (e.g., children's resources, and alternative learning centers). Interestingly, neither of the Programs covered the ToC theme "student/teacher time on task." The Programs did employ experimental or quasi experimental designs (no control groups) and there were few attempts to assess causal connections between the interventions and desired outcomes. (Likewise, this brief study was not equipped to do causal modeling.)   Our Figure 1, summarizing connections between variables, shows lines of connection but, as we have explained, these are mainly "implied" connections not causal ones.

8.4.   *Interventions*. GPE/GEQIP supported two large-scale interventions, teacher training and providing learning materials (textbooks) that were exceptional – their causal connections to learning outcomes were determined and were positive. Other successful GPE/GEQIP interventions with "implied connections" were *curricular reform*, and *improving the learning environment* (through SIPs and school grants), and a set of upstream contributions to the quality assurance mechanism and an inspection policy framework. The strong GPE/GEQIP interventions were all related – by design -- to **learning outcomes.** There were only two narrow lines of connection with the **equity** outcome, an acknowledged weakness that is being addressed in GEQIP II. Common stakeholder perceptions about GPE/GEQIP were that it brought effective school improvement resources right down to the school level and that used carefully and well-designed aid management systems.

8.5.   In contrast UNICEF's lines of connection ran more to **equality issues**. In many cases its interventions were large in scale, like that in support of ABEC's or of ECD, but it was difficult to gauge their impact given the way they were tracked. The agency did create successful connections through its upstream contributions, however, that benefited all three outcome areas. UNICEF is regarded by managers and policy makers as a UN agency that they can work with successfully, and, in addition to the preceding, is appreciated for its strong support for emergency education and its flexibility in reorienting the Child Friendly Schools model towards the School Improvement Planning approach to quality improvement.

8.6. *Enabling conditions*. The success of the interventions is presumed by our ToC to be <u>enabled</u> by numerous contextual conditions, both at the system-level (the education "system") and the more operational level ("basic education"). As a backdrop to all is the *national strategic context*, which during our evaluation period (2009-2013) can be characterized as one of strong economic growth (10% in 2013), political stability, and relative peace and security. Of the postulated six system-level conditions, three were found by our study to be **strong** ("national/political commitment to basic education," high share of national budget to education," and "multi-lingual education policy in place") -- while the other three were found of to be **moderate** in strength. Concerning the basic education conditions, two of four were found to be **moderate** and the other two **weaker** (institutional capacity – particularly at the district level), and "strong budgetary support to BE"). The strong and medium ones could all have set the stage for effective implementation, and, where there were implementation shortfalls, weaknesses in financing and institutional capacity may have been at play to some extent.

8.7. During their lifetime these programs have also faced a variety of challenges. Among the most daunting ones, according to our informants (relevant to one or both of the Programs), were lack of programmatic and geographical focus; weaknesses in monitoring and evaluation; weaknesses in implementation capacity (especially at the local level); lack of differentiated approaches in support to schools; institutional incompatibilities in administrative procedures; and weaknesses in communications, e.g., about financial flows.

B. Recommendations

8.8. Given the positive results on priority outcomes, the reach and effectiveness of Program interventions, and their mostly "implied" connection to the outcomes, the two agency efforts under review here can be considered to have been beneficial to Ethiopia and its basic education programs and to have generally been good investments by Norway. However, for subsequent Programs (e.g., GEQIP stages or new Country Programs), our findings prompt us to suggest some new ways of organizing or managing development cooperation.

1) *Participatory approach to project planning*. Projects like GPE/GEQIP or UNICEF-Country program are taking place within a societal context which is made up of multiple players or stakeholders. Since these stakeholders would participate in the project in different capacities, it is extremely useful to create an opportunity for them to actively participate in project planning from the inception. This would create clear impressions of what the project is all about, and what means and methods it is going to use. There will also be realistic understanding of the project. Both GPE/GEQIP and UNICEF suffered from "false expectations" by those who were involved in implementing the programs (especially at regional/district offices). The idea and goals of the project as well as its relationships with these levels of implementers should have been addressed from the start before they could grow into issues of concern.

2) *Purposing and focusing projects in the planning and design phase*. It has been indicated that the two supported aid programs lacked clear goals, purposes or geographic boundaries. They wanted to do everything all over the country. Some call them thinly-spread. The two Programs might have been more effective if they had put more resources into their lead interventions, such as alternative basic education and ECD for UNICEF, and the SIP/Grants for GPE/GEQIP.

3) *Communication of project plans and implementation*. A very long pipeline made delivering grants to school very complicated; sometimes officials along the way were not fully aware of what the funds were for and how to process them. This was another source of "false expectations." It would be advisable to shorten the pipeline if possible; also all actors along the way need to be fully briefed on the proper use of the funds and the proper ways of processing them.

4) *Improvement of implementation capacity*. The two Programs used different approaches to management (UNICEF has an almost government-like structure with eight bureaus in eight regions, whereas GPE/GEQIP relied on the Government of Ethiopia's structure, but supported it with consultants). Both had their drawbacks: UNICEF witnessed a bloating of embedded staff whereas

GPE/GEQIP faced the inefficiency of government structures. Program managers and planners can look at these different approaches and try to take the best and avoid the worst of both.

5) ***Establishment of professional monitoring and evaluation***. Monitoring and evaluation was a big challenged for both Programs.   For GPE there were challenges in defining the needs of marginalized groups and specifying performance indicators – including ones showing learning outcomes -- to reveal change in such target populations. For UNICEF, this evaluation shows too little focus on outcomes and problems with tracking change over time (over the past two years there has been a growing "corporate" awareness of these issues and significant changes). Planning of solid, results-oriented M&E, involving experienced M&E specialists, should continue to be on both organizations' priority lists.

6) ***Sustainability of the programs***. The legacy and continuity of the programs after their conclusion must be seen as an issue from the very start. In the current projects, there are worries as to how they will be sustained once the projects are over. The tightly organized GPE/GEQIP program does not appear to have a set of expectations for life at the end of 8 years of GEQIP. There appears to be an understanding among some that school grants will be sustained beyond GPE/GEQIP, but that needs to be more widely communicated and other sustainability issues dealt with (this applies equally to the World Bank, the GPE and UNICEF).

7) ***Efficient institutional arrangement***. Development cooperation programs take place within a country's preexisting laws, rules, and regulations, which are often at odds with procedures and norms in the development agencies. Complex problems still arise from failure to confront this reality fully and openly. Consideration of ways to bridge such differences needs to be part of program preparation from the beginning.

8) ***Differentiated approaches***. Ethiopia is a decentralized country in economy, culture, geography and politics. Planning a program, as if the country were a centralized state, can easily lead to dysfunctionality. Participatory planning (see the first point) and a willingness to support different approaches in different regions are keys to success in this context.

References

Dane, F.C., 1990. Research Methods. Brooks/Cole Publishing Company, Belmont, CA.

Federal Democratic Republic Government of Ethiopia, 1994. Education and Training Policy. Addis Ababa

Federal Ministry of Education, 2010. *Education Sector Development Program IV*. Ministry of Education, Addis Ababa.

General Education Quality Improvement Program, Ministry of Education, 2009. *Program Implementation Manual, Volume I: Institutional Arrangements, Planning and Monitoring*. General Education Quality Improvement Program, Ministry of Education, Addis Ababa.

General Education Quality Improvement Program, Ministry of Education, 2009. P*rogram Implementation Manual*, Volume 4: *School Grants Guidelines*. General Education Quality Improvement Program, Ministry of Education, Addis Ababa.

Jennings, M., 2011. *Social Assessment for the Education Secto*r, Ethiopia. DfID.

Merriam, S.B., 1998. *Qualitative Research and Case Study Applications in Education*. Jossey-Bass Inc., San Francisco.

Ministry of Education, 2013. *Education Statistics Annual Abstract* 2012/2013. Education Management Information System, Ministry of Education, Addis Ababa.

Ministry of Education, 2011. *Pastoralist and semi Pastoralist Areas Alternative Basic Education Supervision* Guideline (Amharic Document). Ministry of Education, Addis Ababa.

Ministry of Education, 2010. *Education Sector Development Program IV (ESDP IV): Program Action Plan.* Federal Democratic Republic of Ethiopia, Addis Ababa.

Ministry of Education, 2010. *Pastoralist Areas Mobile Alternative Basic Education Implementation Guideline* (Amharic Document). Ministry of Education, Addis Ababa.

Ministry of Education, 2009. Education Statistics Annual Abstract. Education Management Information System, Ministry of Education, Addis Ababa.

Ministry of Education, 2008. General Education Quality Improvement Package. Ministry of Education, Addis Ababa.

Ministry of Education, UNICEF, 2014. *Capacity Gap Assessment of the Education Sector in the Four Developing Regional States* (Draft Report). Unpublished.

Ministry of Education, UNICEF Ethiopia. A *comprehensive strategy document to reach out-of- school children by 2015.* Unpublished.

Ministry of Education, USAID Ethiopia, 2013. Ethiopian 4th National Learning Assessment of Grades 4 and 8 Pupils- Data Analytic Report.

Ministry of Finance and Economic Development, 2010. F*ederal Democratic Republic of Ethiopia Growth and Transformation Plan 2010/11-2014/15*, Volume 1. Ministry of Finance and Economic Development, Addis Ababa.

Ministry of Finance and Economic Development, 2014. G*rowth and Transformation Plan Annual Progress* Report for F.Y. 2012/13. Ministry of Finance and Economic Development, Addis Ababa.

Patton, M.Q., 1987. *How to Use Qualitative Methods in Evaluation*. SAGE Publications, London.

UNICEF, 2006. *Ethiopia: Country Program Documen*t. Unpublished. UNICEF.

USAID Basic Education Program, General Education Quality Assurance and Examinations Agency, 2008. *Ethiopian Third National Learning Assessment of Grade 4 Students*. USAID, Ministry of Education, Addis Ababa.

USAID Basic Education Program, General Education Quality Assurance and Examinations Agency, 2008. *Ethiopian Third National Learning Assessment of Grade 8 Students*. USAID, Ministry of Education, Addis Ababa.

USAID Ethiopia, 2010. *Ethiopia Early Grade reading Assessment*. RTI International and the Center for Development Consulting, Addis Ababa.

World Bank, 2008. *Project Appraisal Document on a Proposed Credit in the Amount of SDR 33.5 Million (US$ 50.0 Million Equivalent) to the Federal Democratic Republic of Ethiopia in the Support of General Education Quality Improvement Program (GEQIP).* Report No.: 45140-ET. World Bank.

World Bank, 2014. *The Implementation Completion and Results (ICR) Report for GEQIP I.* Unpublished. World Bank.

Annex 1. Terms of Reference

Evaluation of Norwegian Multilateral Support to Basic Education Terms of Reference

1. Background

1.1 Global trends in education aid over the past decade
The overarching goals for education aid globally are the Education for All (EFA) goals and the Millennium Development Goals (MDG) 2 and 3. We are half a year from 2015, the year when the Millennium Development Goals are to be achieved. A lot of progress has been made since the start of the millennium. The global pre-primary education gross enrolment ratio increased from 33% in 1999 to 50% in 2011, equivalent to almost 60 million more children enrolled[9]. The number of children out of school fell almost by half by 2011 but has since stagnated.

Yet we will not meet the global education goals. 58 million children are still out of school, and poor quality schooling is a major obstacle to ensuring that adequate learning is taking place. Inequality in access and learning impede the achievement of quality education for all. One major reason for this is inadequate funding. Basic education is underfunded by USD26 billion per year[10]. The cost of such underfunding to the individual and to society includes lost well- being, productivity and health.

Domestic spending on education globally has on average increased from 4.6% to 5.1% of GNP from 1999 to 2011[11]. A suggested goal post-2015 is that countries should spend between 4 and 6% of their GNP on education[12]. Another international benchmark is that education should be allocated between 15 and 20% of the national budget, which for various reasons is the case in very few countries. Widening the tax base could help some countries meet the education goals, but especially the poorest countries will need external funding in addition[13].

Globally, the volume of financial aid for education has increased considerably since 2000, though it decreased by 10% from 2010 to 2012 (OECD). The education sector has a narrow donor base and is as such vulnerable to low aid predictability and delivery. In 2011, the top five funders of basic education[14] were the World Bank, the United Kingdom, The United States, EU Institutions and Germany[15]. UNICEF is one of the five most important multilateral channels in terms of total financing to education[16], and together, the multilateral agencies contributed 25% of total ODA to education over the past decade. The Global Partnership for Education (GPE) does not report to the OECD, but would be the fifth largest multilateral donor based on its own financial data[17]. Even so, the share that these agencies contribute to basic education has

---

[9] Education for All Global Monitoring Report 2013/4 "Teaching and learning: Achieving quality for all".

[10] Ibid.

[11] Ibid.

[12] The OECD average was 6.3% of GDP (GNP and GDP are not directly comparable. Information taken from http://www.oecd.org/edu/eag2013%20(eng)--FINAL%2020%20June%202013.pdf

[13] Rose, P. and L. Steer (2013): "Financing for Global Education. Opportunities for multilateral action. A report prepared for the UN Special Envoy for Global Education for the High-Level Roundtable on Learning for All: Coordinating the financing and delivery of education".

[14] Percentage share of donor's aid to basic education as a share of all donor's aid to basic education, source OECD/DAC.

[15] Rose, P. and L.Stee, op.cit.

[16] The largest multilateral donors as reported by the OECD-DAC in terms of total financing to education are the Asian Development Bank, The African Development Bank, The European Union Institutions, the World Bank and UNICEF.

[17] Rose, P. and L.Stee, op.cit. (footnote 3)

declined over the last decade relative to that of bilateral donors.

1.2 Trends in Norwegian aid to education over the past decade

The EFA goals and the MDGs 2 and 3 also guide Norwegian aid to education. Basic education is a priority, and two of the main goals for Norway are enhanced access to education and improved quality of the education provided[18]. There is a particular focus on girls' education and on provision of education in a safe learning environment, both during peace, and during war and conflict. Norwegian development cooperation is guided by the principle of a human rights- based approach.

Norwegian bilateral and multi-bilateral[19] aid to education increased from NOK 1293 million in 2004 to NOK 1690 million in 2013. Aid to education as a share of total aid peaked at 13.5% in 2006 and has since gradually decreased to reach 7.2% in 2013[20]. During the past decade, there has been a significant shift in the channelling of Norwegian aid to education from the bilateral to the multilateral channel. Of Norwegian bilateral and multi-bilateral aid to basic education, the latter increased from 30% in 2000, to 73% in 2013.

Most of the aid to education goes to the basic education sub-sector (86% in 2013). The total Norwegian multi-bilateral funds to basic education over the last five years amounts to NOK 3.79 billion. As illustrated in Figure 1, nearly three quarters (74%) of this has gone to UNICEF. Almost a quarter (24%) of the funds have gone to the GPE. For this reason, UNICEF/BEGE and GPE have been chosen as evaluation objects for this evaluation.

Figure 1: Norwegian Multi-bilateral aid to basic education by partner, 2009-2013[21]
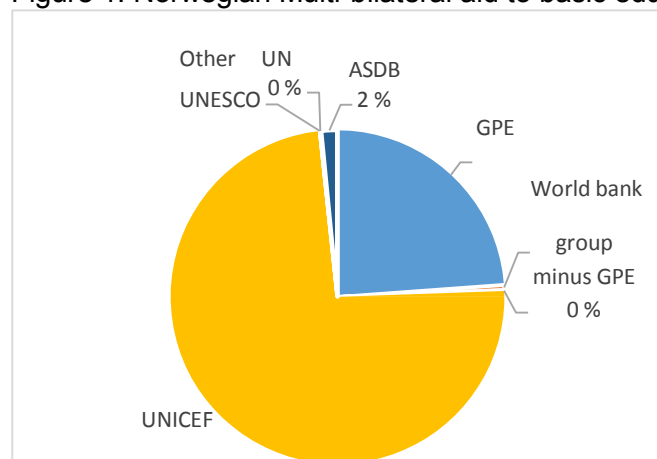
Source: Norad's Database

Allocations to UNICEF for basic education increased steadily in the beginning of the decade, and have stabilised around NOK 480-560 million per year during the latter half of the decade. Norwegian support to GPE started in 2003, and stabilised around NOK 100 million per year but

---

[18] Key document: http://www.regjeringen.no/pages/35167823/PDFS/PRP201120120001_UDDDDPDFS.pdf

[19] Bilateral aid here includes both government-to-government funds as well as funds from the Norwegian aid administration to/through NGOs and CSOs. Multi-bilateral aid includes both earmarked funds from the Norwegian administration to multilateral organisations (MO) centrally and funds from Norwegian embassies to the MO's local country offices. Pure multilateral funds (i.e. core funding) is not included in this evaluation.

[20] The share to education does not include core funding to multilateral organisations, of which some is used to support education, or expenses for administration in Norway.

[21] The two largest recipients in the "other UN" category up to and including 2009 are ILO (NOK 87 million) and the International Institute for Education Planning (NOK 83 million).

doubled in 2011 to reach NOK 200 million.

Norway is actively engaged in GPE as member of the Board, and participating in the constituency group as well as in one of the four committees advisory to the Board of Directors, namely the Country Grants and Performance Committee. Norway has bilateral annual meetings with UNICEF and participates in UNICEF's Executive Board and in working groups as relevant (e.g. the working group on Results Framework, 2014-2017).

1.3 UNICEF's Thematic Focus Area Basic Education and Gender Equality (BEGE)

In the period under review, UNCEF was guided by the second Medium-Term Strategic Plan (MTSP) 2006-2013[22]. According to the UNICEF Basic Education and Gender Equality Thematic Report for 2013, UNICEF aims to play a significant global leadership and advocacy role across the education sector, as well as working with key partners at the country level. UNICEF is committed to working for an evidence-based equity focus in education systems analysis and policymaking, for expanding coverage of basic education for the marginalised and for improving the quality of education.

UNICEF identifies five focus areas[23] that all receive "thematic funding"[24]. This evaluation concentrates on one of these; "Basic Education and Gender Equality (BEGE)".

The 2013 expenditure for BEGE was almost USD 713 million, with USD 112 coming from thematic contributions. Norway contributed almost 76% of the thematic funding for BEGE. Learning outcomes and equity including gender equality (the key focus areas in the current evaluation) accounted for the majority (72.2%) of expenditure for BEGE[25]. The contributions from Norway to UNICEF's Basic Education and Gender Equality for 2006–2013 have varied between USD 72 and 91 million per year[26].

1.4 The Global Partnership for Education (GPE)

The Global Partnership for Education (GPE)[27] is a global partnership of developing and donor countries, multilateral agencies, civil society organisations, the teaching profession, and private sector actors supporting the education sector in developing countries. It currently has 59 developing country partners. Focusing on coordinating action at country level, GPE does not operate as a traditional global fund. While it allocates funds to countries based on an agreed-on formula, it puts primary responsibility on national governments and in-country partners to mobilise and deliver support for education sector plans endorsed by the Local Education Group (LEG) and provides a global platform for mobilising additional resources nationally and internationally.

The LEG, intended to include all actors involved in the education sector, lies at the heart of the

---

[22] A new Strategic Plan (2014-2017) has just been instituted.

[23] The thematic focus areas as outlined in UNICEF's Medium Term Strategic Plan (MTSP) for 2006-2013 are: Young Child Survival and Development; Basic Education and Gender Equality; HIV/AIDS and Children; Child Protection from Violence, Exploitation and Abuse; Policy Advocacy and Partnership for Children's Rights; and Humanitarian Response.

[24] This is an alternative funding modality created to support the goals and objectives of the MTSP. It is more flexible than traditional earmarked funds (sometimes referred to as softly earmarked), and allows for longer term planning and sustainability of the programmes.

[25] The other two focus areas are "early learning" and "education in emergencies".

[26] UNICEF Thematic Report 2013, table page 48. Note that figures before and after 2012 cannot be compared.

[27] GPE started as the Education for All Fast Track Initiative in 2002, but was renamed the Global Partnership for Education in 2011 to reflect key changes in the governance structure.

GPE as a collaborative forum for policy dialogue, alignment and harmonisation of donor support to the national education plan. It seeks to keep all parties fully informed of progress and challenges, and collates and disseminates information, including on GPE partner and non-partner funding. The specific composition, title, and working arrangements of LEGs vary from country to country. When a program implementation grant is requested from the GPE, a supervising entity (SE) or a managing entity (ME) must be designated by the LEG[28]. The SE or ME will play a key role in the LEG, and in supporting implementation.

Following an evaluation published in 2010[29], the partnership was restructured and its mandate broadened. The largest donors to the partnership in terms of cumulative contributions by May 2014 are the UK (USD 857 million), the Netherlands (649 mill), Spain (353 mill), Australia (307 mill), Denmark (288 mill) and Norway (USD 285 mill)[30]. This year, the GPE's independent evaluation committee is commissioning an interim evaluation of the partnership. This evaluation will to the extent possible be coordinated with the GPE evaluation so that the two evaluations can complement, inform and support each other.

1.5 The difference between UNICEF and GPE

There are important differences between UNICEF and GPE, and how they engage in the education sector, which warrant some clarification. At the country level, UNICEF is involved from the national through to the school level contributing to both upstream policy and on-the-ground programme activities and outcomes. While UNICEF participates in the national policy dialogue, UNICEF's funding is often channelled outside the national education budget and targeted to specific groups and/or regions. UNICEF implements some projects directly, some through government and some through civil society. UNICEF has significant presence nationally and sub-nationally, and actively collaborate with government offices at all levels. GPE on the other hand has no direct in-country presence and builds on its partners, including UNICEF in certain countries, for implementation. The GPE Secretariat engages remotely or through periodic in-country short-term visits by secretariat staff or consultants.

2. Rationale, Purpose and Objectives

The current Norwegian government places education on top of the development agenda, and has recently launched a White Paper on Global Education[31]. Much of the funding for basic education is channelled through multilateral actors – notably through UNICEF and GPE. More knowledge on the relevance, effectiveness and efficiency[32] of these institutions will be important for future allocations of aid. The White Paper explicitly states that better results reporting and delivery is expected, and both UNICEF and GPE are potential candidates for substantial scaling up of Norwegian support to education. This is the rationale for assessing the degree to which Norwegian support to basic education through UNICEF and GPE provides quality results in an

---

[28] The SE or ME are a bilateral or multilateral development agency. The SE will transfer grant funds to the developing country government, who will implement the programme, whereas the ME will manage programme activities directly.

[29] See http://www.government.nl/documents-and-publications/reports/2010/02/01/mid-term-evaluation-of-the- efa-fast-track-initiative-final-synthesis-report-volume-5-appendices-vi-viii.html.

[30] See http://www.government.nl/documents-and-publications/reports/2010/02/01/mid-term-evaluation-of-the- efa-fast-track-initiative-final-synthesis-report-volume-5-appendices-vi-viii.html.

[31] http://www.regjeringen.no/nb/dep/ud/dok/regpubl/stmeld/2013-2014/Meld-St-25--20132014.html?id=762554.

[32] As defined by the OECD-DAC, see
http://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm

efficient and cost-effective manner.

The purpose of the evaluation is to facilitate more evidence based policy and programming decisions both in Norway and in UNICEF and GPE, with a dual focus on accountability and learning. This will be achieved through generating evaluation evidence on the relevance, efficiency and effectiveness of Norwegian aid to basic education through UNICEF and GPE, focusing particularly on the achievement of Norway's policy objectives quality of learning[33], gender equality and equity[34], and through increasing the knowledge base of basic education.

The objectives of the evaluation are to:

a) Assess the relevance and coherence of Norway's, UNICEF's, GPE's and selected national government's development objectives. Because a rights-based approach is a key principle for Norwegian development cooperation, the evaluation shall assess if and how this principle is followed by UNICEF, GPE and governments in the selected case countries.

b) Assess the efficiency and effectiveness of financial and technical inputs provided by UNICEF and GPE in generating results at the country level, with a particular focus on quality of learning, equity and gender equality.

c) Identify the added value, or comparative advantage, of GPE and UNICEF respectively. 'Added value' is defined as the degree to which UNICEF and GPE make a difference, positively or negatively, beyond the sheer volume of aid[35].

d) Identify good practices and lessons learned.

e) Provide evidence-based operational recommendations for consideration and action by decision makers and practitioners in Norway, in UNICEF and in GPE, and to the sector more generally.

3. Scope and Evaluation Questions

3.1 Scope
The evaluation covers all of GPE's and UNICEF's support to basic education during the period 2009-2013[36]. It will assess contributions of GPE and UNICEF to achieving results at national level, focusing on outputs and outcomes rather than impact[37], and emphasising the quality of

---

[33] Quality is defined here in line with UNICEF who sets out the desirable characteristics of learners, processes, content and systems. See http://www.unicef.org/education/index_quality.html. The aspect of learner achievement should carry particular weight.

[34] Equity is defined here as "all children hav[ing] an opportunity to survive, develop and reach their full potential without discrimination, bias or favouritism.", ref. UNICEF and consistent with the Convention of the Rights of the child. See http://www.unicef.org/about/partnerships/index_60239.html. This goes beyond equitable access to include equity in the quality of learning.

[35] This includes the way in which UNICEF and GPE interact with each other and with national governments, the quality of their technical inputs, additionality of funds, as well as other factors influencing whether results are achieved in an efficient and effective manner.

[36] To the extent that it is seen as relevant, activities spanning 2014 might also be included. Similarly, the evaluation team can argue for going further back in time.

[37] "Impact evaluation" here refers to rigorous evaluation design to identify the causal effect of an intervention or a policy/reform, including the use of a counterfactual comparison group.

learning, equity and gender equality dimensions of the basic education sector (pre-primary, primary, lower level secondary and including teachers' education and non-formal education[38]). Vocational-, adult- and informal education have been excluded as these areas do not constitute a substantial part of what is supported by either UNICEF or GPE.

GPE prioritises support to fragile states. Quality education for all is no less important in such contexts, and the evaluation will therefore assess how basic education sector plans cover allocation of resources to children who live in conflict- or disaster affected areas, and how they cover disaster risk reduction, conflict sensitivity and other measures relevant to such contexts. Beyond this, humanitarian aid is excluded from this evaluation.

As outlined below, the evaluation is planned with three separate but related parts: A Financial Assessment, a Results Assessment, and a Scoping Exercise for a potential Impact Evaluation. Each part is specified in detail under approach and methodology (Section 4). The Impact Evaluation might be commissioned in a separate tender following the Scoping Exercise.

The evaluation will include in-depth study in four pre-selected countries, and a desk review of 10 countries based on available documentation. Selection of countries for the desk review will be made during the inception phase.

The countries selected for in-depth studies are Malawi, Ethiopia, Madagascar and Nepal. The selection criteria were: the main geographical focus should be Africa; GPE and UNICEF had been present in the countries for some time and preferably since 2009; quality, equity and gender are important areas in national education plans and interventions; potential candidates for future increase in Norwegian development aid to education; at least one country is a fragile state.

3.2 Evaluation Questions

In response to the purpose and objectives of this evaluation, the team should design the evaluation to answer the questions outlined in this section. The questions are organised according to the different parts of the evaluation outlined below, although some may overlap.

1. What results[39] (outputs and outcomes) of basic education interventions have been achieved at the country level? What are the contextual and other factors contributing to or impeding progress on each goal? Have the interventions resulted in any unintended effects?

2. Given the different roles and mandates of UNICEF and GPE; how and to what extent do they complement each other?

3. To what extent are UNICEF and GPE working in ways that support national efforts towards fulfilling the relevant EFA goals in terms of 1) Quality of learning, 2) Gender equality and 3) Equity? This includes assessment of the quality of the technical inputs[40] and the extent to which the inputs are in accordance with the principles of aid

---

[38] Non-formal education should only be included to the extent that it is included in national education budgets.

[39] The focus should be on measures of quality (e.g. learner achievements, drop-out and repetition rates), equity (e.g. Benefit Incidence Analysis, Equity Gap) and gender, but general measures such as enrolment, completion and survival rates should also be included

[40] "Impact evaluation" here refers to rigorous evaluation design to identify the causal effect of an intervention or a policy/reform, including the use of a counterfactual comparison group.

effectiveness[41] and serve to strengthen the ability of governments to achieve their goals. The role of UNICEF and GPE vis-à-vis the education sector group in each country is key to answering these questions.

4. What have been the global patterns of financial allocations to basic education over the past five years? This is further specified in the methodology section.

5. To what degree is there stability and predictability of funding for education from national governments, UNICEF (and within UNICEF), GPE and other relevant actors, and in what ways does the degree of stability and predictability affect the ability to deliver results?

6. To what extent have resources been allocated and utilized in an efficient manner? This should include a minimum assessment of value-for-money, i.e. the extent to which the programme has obtained the maximum benefit from the outputs and outcomes it has produced within the resources available to it.

4. Approach and Methodology

4.1 Specific methodological considerations
The evaluation will consist of three parts:
1. A Results Assessment.
2. A Financial Assessment.
3. A Scoping Exercise: Preparation for a potential future Impact Evaluation[42].

All three parts are expected to inform and build upon each other.

For data collection purposes, visits to UNICEF and GPE headquarters are needed in addition to country visits to the four pre-selected countries.

Results Assessment
The evaluation shall document and assess results of the national basic education[43] interventions directly or indirectly supported by UNICEF and GPE at the country level, in relation to prevailing national policies.

In addition to in-depth country studies, this part of the evaluation shall include a desk study of 10 countries. The selection criteria will be similar to the criteria for the four in-depth case countries (see 3.1), and the countries will be selected during the inception phase. The desk review shall include results reporting from the relevant agency offices and/or governments, as well as review and analysis of relevant strategies, expenditure data, programme documentation, any reviews or evaluations, and a rapid review of available census or survey results to provide a general socioeconomic setting and a sense of educational status. The desk reviews should also include phone interviews with key personnel to allow for a deeper understanding of country processes.

---

[41] The focus should be on measures of quality (e.g. learner achievements, drop-out and repetition rates), equity (e.g. Benefit Incidence Analysis, Equity Gap) and gender, but general measures such as enrolment, completion and survival rates should also be included.
[42] The Impact Evaluation itself will be commissioned in a separate tender.
[43] As defined above.

The results assessment shall focus on three key areas in the basic education sector: Quality, Gender and Equity. These areas represent major obstacles to achieving the EFA goals[44], and it is important to assess the extent to which the relevant actors deliver results that make a difference on the ground. At the same time it is important to be aware of possible trade-offs between equity on the one hand and quality of learning on the other. Improving equity by including marginalised and poor students could potentially reduce the average level of learning in the short-term, unless extra resources are spent to counteract this, even though increased equity will pay off in the longer term. This perspective should be included in the analysis.

The education sector group in each country, be it a local education group (LEG) or its equivalent, plays a key role both in UNICEF's and in GPE's work at the country level. The focus in this evaluation should be to assess the value added of UNICEF and GPE to the group, or if relevant, in any other forum for dialogue with national government. This includes assessing how UNICEF and GPE contribute to the effectiveness of the education sector group and its role in achieving country results. Particular priority should be given to assessing the extent to which the national government takes leadership in the group including if and how it is supported to do so, and the role of any Joint Financing Agreement between donors. The role of GPE Supervising Entities and Managing Entities is also key.

GPE is currently formulating a new financing model, which could have implications for its work on statistics and results. The mentioned GPE evaluation will focus at the global level in addition to the country level, but it is important that the evaluation team is aware of these and other reform processes taking place at the global level.

Financial Assessment
A Financial Assessment study shall collect and analyse available statistics to establish the patterns of financial allocations to basic education, i.e. allocations to and from UNICEF and GPE and allocations to and within the selected countries. The Financial Assessment should be limited to the following:
- Characteristics of, and trends in Norwegian ODA funding to education, compared to ODA funding as a whole.
- For national education budgets: Provide an overview over the case countries' share of GDP allocated to education, the share of the national budget allocated to education, and the education budget going to basic education, for the period 2009-2013. This analysis should be related to availability of external funding, including but not limited to funding from UNICEF and GPE.
- For UNICEF and GPE:
  - Provide a simple overview of characteristics of, and trends in total funds received by the agency from donors (in general and for education), and in the agencies' allocations to education in different geographical areas. For UNICEF this should also include a specification of key focus areas[45] including BEGE and any further specification of BEGE funds.
  - Provide an overview of criteria for allocations of funds within UNICEF and to UNICEF's partners and to GPE-endorsed countries.
  - Provide an overview of flow of funds and identify any bottlenecks, e.g. caused by the timing of allocations to recipients (NGOs, national governments or others) or other factors.
  - Assess the fungibility and/or additionality of domestic and international funds

---

[44] Education for All Global Monitoring Report 2013/4 "Teaching and learning: Achieving quality for all".
[45] Key focus areas are specified by the Medium-Term Strategic Framework 2006-2013.

(e.g. for UNICEF how thematic funding influences thematic allocations of core funding and its relationship to non-thematic funding).

The Financial Assessment study shall base its findings on available statistics from each entity and country administrative data.

All data shall be cross-referenced in tables, graphs and text, analysing patterns within and between the categories over the past decade. Relevant categories for cross-referencing shall be identified by the evaluation team, and include as a minimum themes, sectors and countries.

The findings from the financial mapping shall be used as background data for the evaluation's wider analysis, especially with regard to assessment of the programme theories of change.

<u>Scoping Exercise: Preparation for a potential future Impact Evaluation</u>[46]
Given the recent reforms in GPE, and with reference to scoping study undertaken by White (2010) as part of the mid-term evaluation of EFA-FTI, the Evaluation Department does not foresee an impact evaluation of GPE at this point. In any case, an impact evaluation of GPE is planned by GPE for the years 2017 and 2018.

The scoping exercise should rather aim to identify what possibilities exist in terms of an Impact Evaluation preferably of a) a UNICEF intervention or alternatively of b) a reform or policy change in the basic education sector in one or more of the four selected case countries. If option b) is chosen, the reform or policy change should be one where UNICEF and/or GPE have played a major role, so that the Impact Evaluation can be combined with a contribution analysis. The Evaluation Department plans to use the information from the scoping exercise in the Terms of Reference for the Impact Evaluation. The scoping exercise should include information about any impact evaluations undertaken of the UNICEF interventions under review.

For potential candidates for a future Impact Evaluation, the scoping exercise could address questions and tasks such as;
- Mapping UNICEF basic education interventions including their duration. For each intervention: Has there been any major changes during the intervention period; what was the baseline situation?
- Which basic education sector reforms and/or major policy changes have taken place in the selected countries since 2009? How and to what extent have these reforms or policy changes been supported by UNICEF and/or GPE?
- What is the data needs and availability for analysing these changes, and what is the quality of the data, and needs for collection of primary data? Note that the Evaluation Department foresees that the main source of data will be secondary, and that primary data collection will be limited.

4.2 General methodological principles to be adhered to
The tender shall follow the OCED Development Assistance Committee's quality standards for development evaluation.

---

[46] "Impact Evaluation" here refers to rigorous evaluation design to identify the causal effect of an intervention or a policy/reform, ideally including the use of a counterfactual comparison group.

Details on evaluation methodology will be developed by tenderers in their proposals. The methodology should take cognisance of the data routinely collected (by GPE and/or UNICEF and/or other relevant actors), any previous evaluations and studies from the basic education sub-sector including literature on multilateral aid effectiveness, and relevant progress and other results reports at the global and country levels.

Proposals should include the appropriate treatment of gender and other equity considerations, both in terms of assessing the relevant DAC criteria, and in the data collection and analytical approaches of proposed methodologies.

The evaluation should be based on a theory /theories of change approach, explicitly linking inputs, activities and results, and identifying factors influencing successful outputs and outcomes within a range of contexts and factors that inhibit achievement of stated objectives. This will provide a framework for assessing the efficiency and effectiveness at the country level (and at the global level as relevant, for example when assessing technical inputs). Within the overall analytical framework, mapping of financial flows will be conducted both at the global and country levels. The GPE is currently establishing their own theories of change at the country level. UNICEF has developed a theory of change for BEGE[47].

The informants shall include a broad range of key representatives in the Norwegian aid administration, UNICEF and GPE as well as National Government representatives, donor representatives and other participants in the Local Education Groups (especially those taking the role of Supervising Entity or Managing Entity for GPE, as defined in section 1.4), in addition to relevant education staff (district officers, head teachers, teachers, etc.), parent-teachers associations and students.

The evaluation team shall develop an appropriate methodology that can respond to these Terms of Reference. The evaluation should draw on mixed methods. The methods adopted shall be described in detail in the tender, such as the following suggestive list (not exhaustive):
   a) Document search and reviews.
   b) Analysis of relevant databases and statistics for UNICEF, GPE and case countries.
   c) Interviews with key staff at Headquarters (Oslo, New York and Washington D.C).
   d) Interviews with key representatives of LEGs (or similar sector group if a LEG does not exist) in the selected countries, including government staff.
   e) Field visits to relevant intervention sites in the selected countries, including interviews with key officials, head teachers, teachers, parents and pupils.
   f) Document reviews including research.
   g) Surveys.
   h) Sampling.

The tender should describe the planned approach for the field studies, including how relevant

---

[47] http://www.unicef.org/parmo/files/FA2_Basic_Education_and_Gender_Equality.pdf, page 17.

beneficiaries/stakeholders will be selected for participation in groups and how groups will be organised (e.g. women only? children only? without authority figures?)

The evaluation shall demonstrate how triangulation of methods and multiple information sources are used to substantiate findings and assessments.

5. Deliverables

The deliverables are:
- Inception report not exceeding 20 pages to be approved by the Evaluation department
- Country reports for each pre-selected case country, including financial assessment and results assessment.
- Brief report presenting results of the scoping exercise for impact evaluation.
- Draft report.
- Final report not exceeding 60 pages excluding summary and annexes.
- Two policy briefs not exceeding 2 pages each, one targeting a wider audience and one targeting relevant personnel involved in development cooperation.
- Dissemination in Oslo.

All reports shall follow the Evaluation department's guidelines. All written material will be submitted electronically, and all supporting data will be made available to Norad. Norad retains all rights with respect to distribution, dissemination and publication of the deliverables.

6. Organisation

The evaluation will be commissioned and managed by Norad's evaluation department. Norad will be responsible for the final decisions concerning the Terms of Reference and the evaluation outputs.

A Reference Group will be constituted with separate Terms of Reference. It will include relevant staff from Norad, the Norwegian Ministry of Foreign Affairs, UNICEF, GPE as well as other relevant stakeholders. Reference Group members will be invited to comment on all evaluation outputs before finalisation.

The evaluation will be carried out by an independent team of consultants contracted by Norad's Evaluation department. The evaluation team is entitled to consult stakeholders pertinent to the assignment but it is not permitted to make any commitments on behalf of the Government of Norway, UNICEF or GPE. The evaluation team leader will report directly to Norad's Evaluation department.

Annex 2: List of those interviewed in the field

No.	Name	Sex	Organization	Position
1	**EshetuAsfaw-**	M	Ministry of Education	Director for Planning & Resource Mobilization Directorate
2	**Berhan Hailu**	M	Ministry of Education	GEQIP II Coordinator / Operation Officer /was GEQIP I Coordinator- Hired by World Bank
3	**Birhanu Arega**	M	Ministry of Education	Expert, Gender Directorate
4	**Yohannes Wogaso**	M	Ministry of Education	Special Support Senior Expert /Technical Assistance for Emerging Regions
5	**Mulumebet Tilahun**	F	Ministry of Education	Project preparation & Monitoring Senior Expert & UNICEF Focal person
6	**Thanh Thi Mai**	F	The World Bank	Senior Education Specialist
7	**Girma Woldetsadik**	M	The world Bank	Operations Officer
8	**Patrizia Digiovanni (Contacted but not part of the discussion)**	F	UNICEF / Ethiopia	Deputy Representative
9	**Lata Menon**	F	UNICEF / Ethiopia	Chief of Education
10	**SitotawYimam ,**	M	UNICEF / Ethiopia	Education Specialist
11	**Shumiye Mola**	M	UNICEF / Ethiopia	Monitoring and Evaluation Specialist,
12	**Maekelech Gidey**	F	UNICEF / Ethiopia	Education Specialist focal for ECCD
13	**Alembanchi**	F	UNICEF / Ethiopia	Education Officer (emergency education),
14	**Doriti**	F	UNICEF / Ethiopia	Education Officer
	Demeke Abate	M	Ministry of Finance & Economic Development	Senior Expert- Macro Economy Directorate
15	**Tizita Feleke**	F	Ministry of Finance & Economic Development	Expert- Macro Economy Directorate
16	**Bantiwalu Girma**	M	Yebiruh Tesfa Primary School- Addis Ababa	Principal(Head Teacher)
17	**Haymanot Gared**	F	Hidar 11 Primary School- Amhara Regional State	Principal (Head Teacher
18	**Rahel Yirgashewa**	F	DFID _ Ethiopia	Education Advisor

Annex 3: Data collection instrument

A. National Statistics Office or Report for 2013

1. **Percent of population below national poverty line in 2013:**
    (compare to the figure for 2011 which was 30%).

B. Education Statistics Office (EMIS) or Report (we need to update some
figures)**:**

2. **System parameters**

                                                                2009 2013
    *   **The average primary student-teacher ratio**   urban ____

        ____
                                                              rural   ____ ____
    *   **The average primary student-classroom ratio**   urban ____ ____
                                                              rural   ____ ____
    *   **The average primary school teacher absentee rate**   urban ____ ____
                                                              rural   ____ ____
    *   **The average expenditure per student (primary)**   urban ____ ____
                                                              Rural   ____ ____

3.   <u>NER gr 1-4</u>   <u>NER gr 1-8</u>

            <u>2009</u>   <u>2013</u>   <u>2009</u>   <u>2013</u>
    Female   _?__   ____   ____   ____

    Total   _89*_   ____   ____   ____
    *Based on ESDP 2011-2015; needs confirmation

4. Pre-primary access: we have UNICEF GER figures for 2009 (   ) **but not for 2013 ____. Also,
    not broken down by F/T ___ ____. Is it possible to find NER anywhere; if so, __
    _____ (with Female breakdown _____.)**

5. <u>Primary Completion Rate</u>. We have UNICEF figures for 2009 (gr 5/8; TF); **we need for year 2013
    ____ ____ ____ ____.)**

6. EMIS and disaggregated data.. There have been efforts to strengthen the Education Management
    Information System, yet many basic statistics (NER primary) are hard to find and breakdowns by
    gender, urban-rural, region, woredas, etc. are rarely given.
    *   **Why is it still difficult to get a full range of basic statistics?**
    *   **Is there a way to use the system to find basic breakdowns for use assessing equity issues,
        such as access for girls, cultural-linguistic minorities, children of very low income
        families, disabled children, and differences across regions, and *woredas*, etc.?**
    *   **What household surveys are available?**

C. Ministry of Education Officials

7. <u>General Questions</u> (*these will be repeated for the other groups of respondents*)
    *   **What are the ways in which UNICEF and GPE have had the biggest positive effect on basic
        education in Ethiopia?**
    *   **What are the ways in which UNICEF and GPE could be more effective in promoting basic
        education in Ethiopia?**
    *   **If you could re-allocate UNICEF and GPE funds spent on basic education in Ethiopia
        but had to stay in the same budget limit, where would you add funding and where would you
        reduce it?**

- **Are there ways in which UNICEF and GPE need to adapt better to Ethiopia's history and culture regarding education?**

8. Information about outreach programs for poor and vulnerable:
   - **The extent to which primary education is fee free?**
     If so, how large are any indirect costs (uniforms, books and materials, transportation, etc.)
   - **Existence and magnitude of a school feeding program _____**
     Who pays for it?
   - **Any girls' or mothers' clubs or other social support networks for girls and others)**
     _____
   - **Any boarding or mobile schools for the geographically challenged?**
     _____
   - **Any scholarship programs for girls or otherwise disadvantaged?**
     _____

9. **Did changes in aid programs and the conditions attached to them have unintended consequences?**

10. **Did unintended consequences of education aid programs include any of the following:**
    - incentives for teachers to go on strike;
    - incentives for unqualified people to become teachers just to earn a paycheck;
    - incentives to divert government spending away from education to other budget areas;
    - aid programs becoming a political weapon or resource in competition among competing factions in the government or different regions or ethnic groups;
    - incentives for diversion of funds to other purposes (including other legitimate educational purposes) at the level of individual schools;
    - unintended outcomes that were good, including development of strong management procedures that were then copied by education ministry or school personnel.

D. Ministry of Education Finance Office (or Ministry of Planning or Finance) (whatever government agency tracks government expenditures in education).

11. Public Financial Management Information
    a. **Is there any information more recent than the 2010 PEFA which has reviewed the public financial management system?**
    b. Budgetary and actual allocations to education
       Over the past five years,
       - **What percent of the GDP has been allocated to the education sector?**
       - **What percent of the national budget has been allocated to education?**
       - **What percent of the education budget has been allocated to primary education?**
       - **How has the education budget been allocated between development and recurrent expenditures?**
       - **What have been the actual level of expenditures for education as compared to the budgetary allocations?**
       - **What have been the actual level of expenditures for education as compared to the budgetary allocations?**
       - **When have budgetary releases occurred?**

E. General Education Quality Improvement Program I (2009-2013) (Project Management Unit or equivalent)

12. **General Questions**
- **What are the ways in which UNICEF and GPE have had the biggest positive effect on basic education in Ethiopia?**
- **What are the ways in which UNICEF and GPE could be more effective in promoting basic education in Ethiopia?**
- **If you could re-allocate UNICEF and GPE funds spent on basic education in Ethiopia but had to stay in the same budget limit, where would you add funding and where would you reduce it?**
- **Are there ways in which UNICEF and GPE need to adapt better to Ethiopia's history and culture regarding education?**

13. **Questions about development outcomes during 2009-2013:**
    a. For the learning outcome measure (% of students scoring at least 50%)
    - **Are the tests comparable from year to year? Yes __No __ If not, how is it possible to track improvement?**
    - **Could there also be breakdowns by region, urban-rural, cultural-linguistic group, etc., and those further broken down by gender? (If possible, please present the relevant figures)**
    b. **For primary completion rates (PCR):**
    - **Why are GEQIP and UNICEF figures so different? Which is more accurate and why?**
    - **Can PCR also be broken down, as suggested under item "a"?**
    c. **For teachers with appropriate qualifications: How wwas it possible for this to increase so much in such little time?**
    d. **On learning materials,: on site results who show books per student below 1 (0.65):**
    - **Has that been remedied (goal 1:1).**
    - **Were the expected number of mother-tongue textbooks distributed in 2013/14?**
    e. **What percentage of *woredas* have been reporting on physical and financial performance quarterly in recent years (at least 2013)?**
    f. **Any new data on the increased use of student-centered instruction? (possible school visit)**

Annex 4: School Visit Report: Yebiruh Tesfa and Hidar 11 Primary Schools

Two primary schools, one in Addis Ababa and the other in Amhara region, were visited to collect data on the effects of GPE/GEQIP and UNICEF particularly on the School Improvement Program (SIP) implementation, school grant utilization, and community participation. The schools were **Yebiruh Tesfa Primary School** in Addis Ababa, and **Hidar 11 Primary School** in Amhara Region. Data on the enrollment of boys and girls were also gathered to show the situation of **gender equality**, as well as data on the **use of mother tongue textbooks**. Although the report cannot be considered representative and used to generalize to all schools, it does gives some views about program intervention and outcomes at the grass root level and can be used to help confirm the validity of what was presented as progress at the central level.

Main findings

- The trend of enrolment in these two schools shows that the number of student decreased a little from 2009/10 to 2012/2013 (1039 to 912 in Addis Ababa and 511 to 497 in Amhara region). This condition is also seen in **net enrollment ratio** of the Addis Ababa and Amharic regions, whereas there were increases in other regions, which suggests there is some population mobility. Few students dropped out in these years.

- **School Improvement Program (SIP)** is one component of the GEQIP (also supported by UNICEF) and is confirmed to be operating in the schools. Findings in the two schools reveal that **School Improvement Plans (**three year strategic plans) were prepared by SIP committees in both locations, composed of parents, teachers and students (and for Hidar 11 *kebele* [sub-district] representatives). The overall planning and implementation of SIP was confirmed to be guided by the standard framework developed by MOE and regions, in part a result of the school/community training efforts made by GPE/GEQIP and UNICEF.   In fact, the *School Improvement Framework, School Improvement Program Implementation Manual*, and *School Improvement Guidelines* were the key references used at the school level, all with the purpose of helping the community improve school conditions and student learning outcomes. Following these guidelines data were collected from parents, teachers and students to identify and prioritize the school problems. Once the gaps were known, the committee prepared a strategic plan and communicated it to the school community, including Parent Student Teacher Association which is active in both locations. Examination of the plans shows that the SIP program has been addressing specific problems of each school in line with the national SIP framework. During the school visit both head teachers reported that the SIP had helped the school to address many challenges and improved student learning and retention.

- **GEQIP school grant and usage**

  The information from the visited schools confirms that the primary schools have been receiving School Grants from GEQIP to support implementation the SIP plans.

Name of School	School Grant in Birr						Remark
	2008/9	2009/10	2010/11	2011/12	2012/13	2013/14	
Yebiruh Tesfa primary school				40,000	41,000	42,000	Started receiving grants in the year 2011/12 after it was shifted from private school to governmental school
Hidar 11 Primary School	6645	7574	22033	45475	6315	25910	The school received more budget in 2011/2012 and didn't finished in the year. So, a reduced budget was allocated in 2012/13, given the carry-over from 2011/ 2012.

The head teachers reported that the grant allocation is based on the number of students and the schools regularly received these grants directly through the government financial structure at the Woreda level.

Main activities the school performed during these years are:

➢ Building and renewing toilet for female students (both schools)
➢ purchasing laboratory equipment and materials (both schools)
➢ Greening the school compound (Yebiruh Tesfa Primary School)
➢ purchasing reference books (both schools)
➢ Enriching pedagogical center of the school (Hidar 11 Primary School)

The SIP committee and PTSA participate in planning, monitoring implementation and controlling the budget by using it for the intended and prioritized activities and managing it efficiently. The head teachers reported that the GEQIP grant addressed many of the school problems and created conducive environment for students to learn.
Challenges raised at Hidar 11school is that the amount of money released fluctuates from year to year, and this creates problem in planning and implementation. The other issue is the time it is being released to the schools. Sometimes it comes very late so that the planned activities cannot be implemented on time.

- **Other funds.** Both schools also received support from USAID during this period. In 2010 **Yebiruh Tesfa primary school** were granted equipment for its the science laboratory. So, due to this and GEQIP support the school laboratory improved its standard and become a good laboratory where students gain practical skills. USAID gave financial grants to Hidar 11 school through 2011. These grants were mainly aimed at empowering female students and creating conducive environment for female students in the school. Accordingly, the school built toilet for female students and trainings were given for female students on life skill. The school also used the grant to reward outstanding female students at the end of every semester. These grants were used together with GEQIP funds to carry out the SIP plan of the school.

- **Community Participation in functioning school committee.** In both schools the school heads reported functioning of many committees where community members are active. The **School Improvement Plan (SIP) committee is mainly** to prepare the SIP plan and follow up its implementation. The committee has a regular meeting on the agreed date for example, every two weeks on Thursdays (in Yebiruh Tesfa), and every ten days (Hidar 11 School). Teachers, parents, supportive staffs and students are members and the principal is the leader. **The Parent, Student, and**

Teachers Association (PSTA) is a permanent structure that has a responsibility for the overall work of the school, including school planning, decision-making, implementation, disciplinary problems, etc. It is also involved in implementing the SIP. The head of the PSTA is nominated by the group.

- **Gender equity-** The enrollment of fewer girls than boys has been a typical characteristic of the Ethiopian school system. Because of this, a number of interventions, including some in GEQIP are targeting improved gender equality in enrollment as one outcome indicator. The reality in both visited schools showed a positive gender balance. In fact, in both schools girls have outnumbered the boys for the last three and four years. Although we found no data on school performance by boys and girls, there is a general sense the girls have been performing better than boys. According to one of the principals: "*Some four or five years ago the school [began] using every effort to help female students. However, female students are currently better than their male counterparts in almost all aspects" (Head teacher of Hidar 11 school).* Thus, now the schools are planning to support both males and females. The real equity problem at the schools is in the treatments of disabled students. The school lacks materials and equipment to support them. By arranging special classes, teachers are trying their best to support students with different disabilities.

Usage of mother tongue text books. Amharic mother tongue textbooks, prepared in 2010, are used in Hidar 11 primary schools. The new (2012) mother tongue textbooks are not yet being utilized. In 2007, the school conducted informal assessment to identify the mother tongue language of the students in the school and found that Amharic is the mother tongue for all the students in the school. Thus, for reading and other skills the school is using Amharic textbooks for lower primary grades.

Annex 5: Schematic of the project cycle and questions that reveal the quality with which it is implemented

Step
Decision to Invest in a Country
Does the agent have transparent standards for allocating its funds among candidate countries? These standards can be based on various criteria--the country's or sector's performance, its educational needs, or its fragile condition. The key questions are how these allocation standards are set, whether they are applied as intended, and whether they reflect the priorities of the agent's principals, including Norad. For example, given Norad's education goal of funding "education in crisis", is some funding allocated to countries in conflict or recovering from natural disasters?
Decision on Project Focus
This refers to the broad goals of the project, not to the selection of the investment sector or sub-sector. GPE and UNICEF/BEGE fund only the basic education sector, but they can select a range of project goals. The key questions are how these goals are arrived at and whether they reflect the priorities of the agent's principals, including Norad's several goals, such as educational quality, girls' education, instruction in the mother tongue, and equitable access to education.
Project Design
The quality of a project's design is strongly related to its ultimate effectiveness.[7] The design stage is also the best opportunity to maximize the efficiency of the project. The project design should include an economic analysis, which is most useful when conducted very early in the project cycle to weed out bad projects and bad project components. This analysis should evaluate the classic counterfactual questions.   For example, does the agent assess whether the program design is coherent and realistic relative to the country's past track record? Does it require a results framework with performance indicators, baseline data, and M&E arrangements? Does the agent require an economic analysis sensibly scaled to the data available, and does the agent use its findings to make decisions? Do implementation arrangements reflect a realistic evaluation of capacities? Have risks, such as sustainability and political economy risks, been identified and, if possible, addressed? Are proper safeguards in place to ensure that the project's money is spent on its intended beneficiaries?
Project Approval
Does this stage require a final tough review of the proposed project before money is contractually committed to it?
Project Implementation
Does the agent require and adequately fund systematic oversight of the project's progress, including at least annual measures of the status of outputs and outcomes? Does it routinely review these data to identify projects that are failing on one or more components? Since actions to turn problem projects around during implementation are found to markedly increase their odds of achieving their objectives, does it encourage proactive adjustments to the project design to rescue failing projects or project components? Is the agent willing to cancel projects that, for whatever reason, seem doomed?
Project Closure
Does the agent require an independent evaluation of the project's results, including an analysis of the apparent reasons for success or failure? Does it capture the lessons learned from these evaluations and use them to modify those policies and processes of the agent that affect the project cycle?