

While the field of international education has made great strides in recent years with raising the number and the quality of impact evaluations, their results are incomplete without cost data for these interventions. Policy-makers and donors cannot make fully informed decisions about the best way to invest limited resources without information about the costs of achieving desired outputs and outcomes through different interventions or delivery strategies. Evidence on cost of interventions is also critical for making responsible decisions about scaling and sustaining programs within country systems. However, differences in donor and national reporting systems currently hamper collection and analysis of cost data. Thus, opportunities for policy decision making that include cost information are often missed by national governments and international funders. This guidance note addresses this gap by introducing a common framework for collecting, analyzing, and using cost information across the global donor-supported education portfolio.

The BE² Steering Committee*

February, 2020

This document should be cited as follows:

Walls, Elena, Tulloch, Caitlin, and Holla, Alaka. 2020. "Cost Measurement Guidance Note for Donor-Funded Education Programming." Washington, DC: United States Agency for International Development, prepared for Building Evidence in Education (BE²).

Cover Design:

FORUM ONE

*The Building Evidence in Education (BE²) is a donor working group led by a Steering Committee composed of the Department for International Development (DFID), United Nations International Children's Emergency Fund (UNICEF), United States Agency for International Development (USAID), and The World Bank Group. The Mission of BE² is to harmonize research efforts across donors with the ultimate objective of improving availability and quality of data and evidence for decision making by national governments and funders.

Measuring Costs of Donor-Funded Education Programming

Foreword

The Building Evidence in Education (BE²) donor working group was launched in 2012 with the aim to engage bilateral and multilateral donors and foundations committed to:

- Strengthening donor research collaboration and coordination;
- Encouraging higher standards of commissioned research; and
- Promoting the availability and access to rigorous evidence.

The working group is led by a Steering Committee composed of the Department for International Development (DFID), United States Agency for International Development (USAID), The World Bank Group and a rotating representative of the United Nations (UN) organizations, currently the United Nations Educational, Scientific and Cultural Organization (UNESCO).

This series of Guidance Notes, prepared for the BE² working group by its members, provides tools and guidance for generating better evidence and leveraging existing evidence more effectively and efficiently. These Guidance Notes have benefited from the advice of BE² member organizations and are intended to serve as tools for researchers and commissioners of research.



This guidance note has been authored by Dr. Elena Walls (USAID), Caitlin Tulloch (International Rescue Committee), and Alaka Holla (The World Bank), for the BE² working group. The note is based in large part on USAID's approach to measuring costs of donor-funded education programming.

Many organizations and practitioners have provided input, e.g., during working sessions at the Comparative and International Education Society (CIES) conference, the Education and Development Forum – UKFIET, or other feedback opportunities. BE² thanks all its members and other contributors for comments provided to drafts of this guidance note, in particular Rachel Hinton, DFID.

This page intentionally left blank

TABLE OF CONTENTS

GLOSSARY	vii
EXECUTIVE SUMMARY	ix
1. INTRODUCTION	1
2. FRAMEWORK FOR MEASURING COST	4
3. COST DATA ANALYSIS APPROACH	8
3.1. Cost-economy analysis and sustainability	9
3.2. Cost-efficiency analysis	12
3.3. Cost-effectiveness analysis	13
3.4. Cost-benefit analysis	16
4. COST CAPTURE APPROACH	17
4.1. Cost capture approach for donor-funded interventions	17
4.2. Cost capture approach for government-funded programs	22
5. HOW TO USE COST ANALYSES	24
5.1. Learning from cost data analysis results	24
5.2. Understanding Cost Drivers	25
5.3 Considering Context	27
5.4 Considering Intervention Delivery	28
5.5. Applying Cost data analysis Results to Another Context	31
6. RECOMMENDATIONS FOR IMPLEMENTATION	33
References and Useful Resources	35
Annex A. Description of Reporting Categories	37
Annex B. Reporting on Government and Non-Government Contributions	42
Annex C. Reporting on Details of Outputs and Dosage	49

GLOSSARY

Cost, cost capture	<p>Cost is defined as totality of incurred monetary expenditure and in-kind contributions that produced a desired education output or outcome. Cost is different from price, which is the monetary value exchanged in a market transaction for one unit of a good or service; it is specific to the site and time of the transaction.</p> <p>Cost capture refers to a systematic way of recording the objectives of expenditure and corresponding contributions according to predefined categories.</p>
Cost category	<p>Cost category is defined here as class of costs incurrent to produce a particular kind of education outcome, such as trained teachers or teaching and learning materials.</p>
Effect, program effect, marginal effect	<p>Effect of a program is defined as changes on outcome variables of interest that can be attributed to a particular intervention, ideally derived from an impact evaluation using an experimental or quasi-experimental methodology. Marginal effect refers to the changes in the outcome variables attributed to the intervention holding all other explanatory variables (covariates) fixed.</p>
Equity	<p>Equity in education means that personal or social circumstances such as gender, disability, ethnic origin, or family background are not obstacles to accessing high-quality education services. The programmatic principle of equity refers to the idea that program resources must be used to compensate for existing barriers to education for marginalized groups.</p>
Expenditure	<p>Expenditure, in this guidance, is defined as the amount of money spent to develop and/or implement an education intervention, including labor, materials, travel, and other expenditures, as captured through the accounting system for the reporting period.</p>
Cost-benefit analysis	<p>Cost-benefit analysis, also known as Return on Investment (RoI), is a systematic approach to estimating a monetary value of all benefits produced by a program and comparing this monetary value to the total costs of the program.</p>
Cost-economy analysis	<p>Cost-economy analysis is a systematic way of assessing components of the program and prices at which inputs were purchased for each component to establish how well the value-for-money and equity principles were adhered to during the program implementation.</p>
Cost-effectiveness analysis	<p>Cost-effectiveness analysis is a systematic approach to calculating the ratio of the amount of “effect” a program achieves for a given amount of cost incurred, or the amount of cost required to achieve a given impact. (See definitions of “cost” and “effect” above.)</p>
Cost-efficiency analysis	<p>Cost-efficiency analysis is a type of economic analysis that is used to calculate the costs of producing outputs. The results are frequently</p>

	expressed as unit costs for producing outputs.
Ingredients	Ingredients data refers to a breakdown of labor, materials, rent, travel, and other elements for which the expenditure is reported into their respective units, quantities, and prices. In-kind contributions include labor, materials, and resources donated by other parties and should be included in the overall costs.
Unit cost	A unit cost is the total expenditure incurred to produce one unit of a particular product or service. Unit costs include all fixed costs, or overhead costs, and all variable costs, or direct material and labor costs.
Value-for-money	Value-for-money refers to the optimal way to expend resources to achieve intended outcomes. Interpreted broadly, value-for-money analysis refers to a combination of cost-economy, cost-efficiency, and cost-effectiveness analyses.

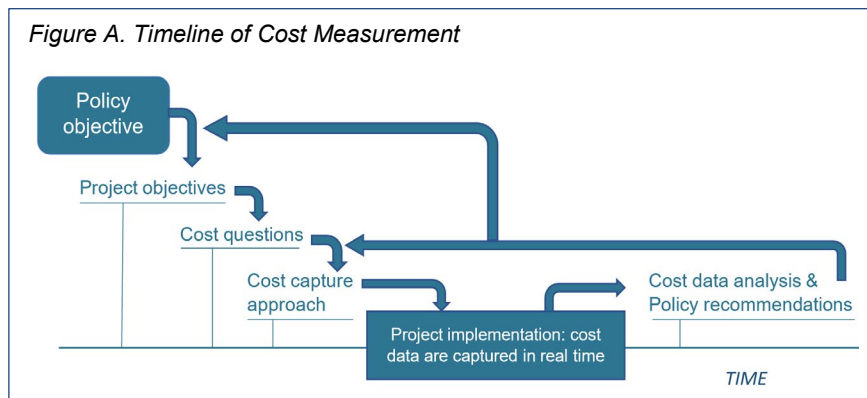
1. EXECUTIVE SUMMARY

The field of international education is increasingly data-driven. We have made great strides in recent years in increasing the number and the quality of impact evaluations, but it is difficult to translate their results into policy recommendations without data on the costs of these interventions. At present, funders and governments often cannot identify cost-effective investments because the data required to accurately measure costs are frequently absent. Even providers of education services have difficulty untangling costs of their own programs and estimating the total costs of a specific intervention. Where cost evidence is available, data collection metrics and cost data analysis methods are not standardized, which makes it difficult to compare data across different agencies and programs. Thus, national governments and international funders often miss critical opportunities to leverage information about the costs of interventions when allocating budgets.

The main objective of this guidance note is to establish a common framework for capturing, analyzing, and using cost information in education programming. In accordance with the BE² mission to improve coherence of donors' approaches to strengthen impact at national and global levels, this document aims to guide the alignment of financial reporting requirements to the needs of cost data analysis. In so doing, national governments and donors will be able to answer common cost-related questions. In turn, a better understanding of the cost structure of common interventions and the costs of delivering specific education components in low- and middle-income countries will help optimize the allocation of limited resources to achieve maximum reach and impact of education outcomes. This guidance can thus be useful to those who are commissioned to produce research, independent researchers and academics, implementation partners of multilateral donors, as well as national governments.

The framework for cost measurement put forth in this guidance builds on experience of DFID, USAID, and the World Bank with collecting and using cost data. It is also informed by the most up-to-date research on costing. The framework includes three interconnected elements: cost data analysis, cost data collection, and the use of cost data. It emphasizes how the objectives of cost measurement inform what cost data must be captured, which in turn determines which types of cost data analysis are possible. Much like impact evaluation design, the questions that cost measurement should answer need to be articulated before program activities are implemented, as these questions will inform how data on the costs of activities are collected and how results are documented throughout an intervention (Figure A). Without good cost capture mechanisms in place from the start of a project, accurate estimates of costs will be all but impossible.

Figure A. Timeline of Cost Measurement



The guidance note outlines several options for the types of cost data analysis that can be performed, each of which can help address different questions and may require somewhat different data. The types of cost data analysis methods described in the guidance are cost-economy, cost-efficiency, cost-effectiveness, and

cost-benefit. These can be used separately or in combination, depending on the cost data analysis objectives and questions asked.

This guidance note recommends an activity-based costing approach for the collection of cost data. Activities include any event, unit of work, or task with a specific goal — such as conducting teacher training, developing or producing books or other learning materials, or undertaking a learning assessment. The activity-based costing approach lends itself well to international education since the number of education-related tasks is fairly universal across countries and education systems. The guidance proposes a thematic categorization of activities fitting within different education levels, as defined under UN’s Classification of the Functions of Government. In addition to monetary expenditure on education activities, the guidance note proposes an approach to capture in-kind contributions of other parties.

Collecting and analyzing cost data also can help identify contextual and programmatic features that affect efficiency and effectiveness of donor and government investments in the education sector. An analysis of pricing, context, and intervention features is useful at any stage of an intervention: when managing an existing program, when planning to sustain a successful program, and when assessing a program’s viability in a new context. A close examination of context-specific cost drivers can help inform planning, scale up decisions, and help with the selection of an intervention. It is important to note that cost-efficiency and cost-effectiveness estimates are produced for a specific context, dosage, and scope of the intervention. Changes in dosage and scope can result in unpredictable changes in estimated costs. The guidance outlines recommendations for considerations that must be taken when applying results of cost data analysis to a different context.

The guidance note’s fundamental objective is to facilitate adoption of robust cost measurement practices and for the results to improve effectiveness of global investments in education development among funders and national governments. The closing section of the note presents recommended steps for instituting cost measurement practices in an international donor agency. Harmonizing cost measurement across the entire global sector will increase its value by allowing for comparisons of cost-efficiency and cost-effectiveness of interventions funded by different donors. Similar to international standards for evaluation studies, adopting clear standards for cost studies would allow us to build and use the knowledge base and ultimately improve the efficiency and effectiveness of international investments in education.

1. INTRODUCTION

KEY POINTS:

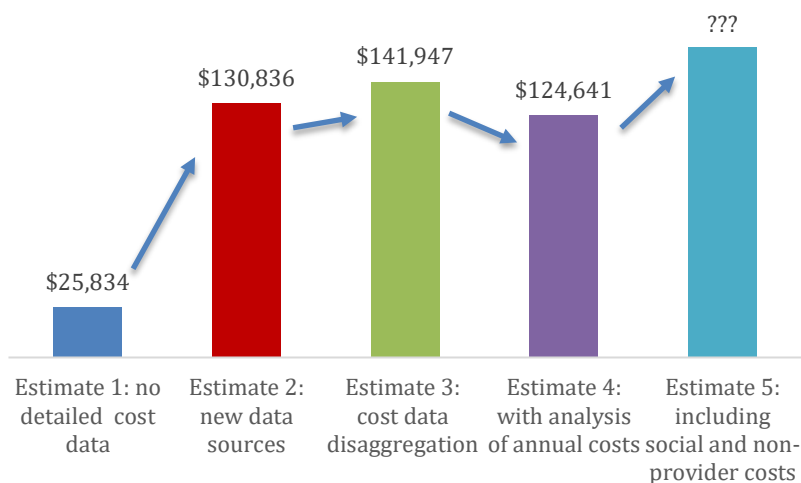
- International education funders need better cost data to improve results and value for money of their investments.
- Standardizing donor-required reporting would help streamline reporting across partner organizations and reduce reporting costs for governments.
- A common cost measurement framework is needed to help harmonize cost measurement approaches across donors as well as national governments that benefit from bilateral and multilateral donor investments.

With an increased emphasis on evidence-based programming, there is a need for better data on the impact and costs of donor-supported interventions in education. While the field made great strides in recent years with raising the number and quality of impact evaluations, its results are incomplete without cost data for these interventions. At present, it is difficult for funders or governments to use cost and effectiveness data to inform funding decisions because the cost evidence is not comparable and frequently incomplete.

For example, a non-profit organization in Bangladesh was trying to figure out how much the pre-school program they were operating in rural areas to deliver early childhood education actually cost. The staff who worked on this assumed it would be a tedious accounting exercise. However, it quickly became apparent to everyone that collecting cost data was really an investigation, one

that takes you deep into the nuts-and-bolts of program design and implementation. And in this particular investigation, the initial findings were surprising: the initial “budget information” that the organization was keeping **was underestimating total costs of the program by a factor of five.** Moreover, the estimate model was also erroneous in many ways. It included general early childhood development activities that had nothing

Figure 1. Changes in a pre-school program cost estimates based on precision of available cost data (in USD)



to do with the preschool program; it also omitted some key costs of running the preschool program such as management and overhead and costs borne by the community.¹ Figure 1 shows the change in the cost estimates of the program delivery depending on how precise and complete the actual cost data for the program were, and the final costs will only be known once the data on the social and non-provider costs (such as community contributions) are collected and included in the cost estimates.

This example illustrates how even providers have difficulty untangling costs of specific programs they provide and estimating the total costs of their programs. In addition, differences in donor and national reporting systems can act as an obstacle in collecting comparable cost data. Where cost evidence is available, the data collection metrics and cost data analysis methods are not standardized, which makes it difficult to compare data across different agencies and programs. Thus, national governments and international funders often miss critical opportunities to leverage information about costs of interventions in making policy decisions.

Differences in financial reporting requirements across donors also result in high transaction costs for governments and non-governmental organizations (NGOs) that benefit from donor support. National reporting systems struggle to comply with reporting requirements imposed by multiple donors, and international NGOs have to create parallel systems to satisfy the financial reporting requirements of different donors (see Box 1). Moreover, differences in metrics and measurement approaches make using data generated through different reporting streams difficult to compare and use for management decision making, undermining the usefulness of this information for policy makers and program implementers alike.

Box 1. Complying with donor financial reporting systems at the International Rescue Committee

The International Rescue Committee (IRC) implements large-scale education programs, funded by donors including DFID, USAID, and private foundations. Each of these donors requires a different format for budgeting and reporting and have different definitions for which inputs can be counted as “support” vs. “programmatic” costs. Moreover, the donors who ask IRC to report cost-per-output of these programs use different output metrics across donors, and sometimes even across projects funded by that same donor. Because of fragmented finance reporting requirements, the IRC must maintain multiple versions of budgets for all of their projects—one that follows an internal template and one that is re-arranged to comply with donor charging requirements and templates. Because the budgets are not divided into consistent categories, staff cannot compare the cost per person of programmatic spending or materials spending across grants; budget lines have to be compared line-by-line to ensure comparability. And while staff spend hours assembling estimates of cost-per-output or cost-per-outcome for grants with value-for-money (VfM) reporting, these estimates cannot be meaningfully compared across projects because both the costs but also the definition of “outputs” and “outcomes” vary from grant to grant.

At an even more fundamental level, the cost of specific components of common interventions in the education sector—such as in-service teacher training or the production of books—and how

¹ Samuel Fishman, “Excuse me, can anyone tell me the cost of this education program?” (blog), World Bank Blogs, September 25, 2018, <https://blogs.worldbank.org/education/excuse-me-can-anyone-tell-me-cost-education-program>.

these costs vary in response to programmatic and contextual features, is poorly understood due to the lack of associated cost data. Costs of these essential components of education delivery are frequently bundled with the cost of other components and thus are not readily available for examination by policy makers. Developing standard mechanisms for capturing the costs of education interventions will increase transparency, allow for the linkage of costs to outcomes, and enable value-for-money analyses, thus providing a pathway toward resource optimization across programs and contexts.

At present donors have limited practical guidance for capturing, analyzing, and using cost data.² In accordance with the BE² mission to improve coherence of donors' approaches to strengthen impact at national and global levels, this document aims to guide the alignment of financial reporting requirements to the needs of cost data analysis. In so doing, national governments and donors will be able to answer common cost-related questions. A better understanding of the cost structure of donor-funded interventions and the costs of specific education delivery components in low- and middle-income countries will help optimize allocation of limited resources to achieve maximum reach and impact of education outcomes.

The guidance is not expected to be a comprehensive manual on all issues of use, capture, and analysis of cost data. Rather, the main objective of this guidance note is to establish a common framework for how donors think about capturing, analyzing, and using cost in donor-supported education programming. By "cost," we mean the value of resources allocated to support specific activities in or relating to the education sector, which may include direct expenditures by donors and contributions by other actors, in addition to pre-existing national expenditure in the education sector. By "education programming," we mean any activity that aims to improve access to and quality of formal and non-formal education at all levels, from early childhood to tertiary, including direct service delivery. This guidance can thus be useful to those who are commissioned to produce research, independent researchers and academics, as well as implementation partners of multilateral donors. We hope that the ideas put forth in this note will be of benefit to national governments as well, in their efforts to improve the efficiency and effectiveness of the education delivery in their countries.

The next section describes a few common objectives of costing exercises and presents a framework for cost measurement. Section 3 outlines what we can learn and what data is required for different types of cost data analyses, while Section 4 proposes processes for ensuring that the right data are collected for these analyses. Section 5 provides guidance for ensuring that the information coming out of cost data analyses is used appropriately in policy dialog within a country and globally among donors. Section 6 concludes with implementation recommendations.

² Some useful guidance can be found in Levin et al (2017), USAID's Cost Reporting Guidance (2018), Dhaliwal et al (2013), Vassal et al (2017), and World Bank (2019).

2. FRAMEWORK FOR MEASURING COST

KEY POINTS:

- International education funders have articulated policies that aim to promote improved value for money of education investments.
- A specific cost measurement guidance is needed to answer common education investment questions.
- The proposed framework specifies three distinct but interrelated parts: *cost analysis questions* that determine the *cost capture approach* and the subsequent *cost data analysis*.
- Real-time cost data capture is key to accurate cost measurement.

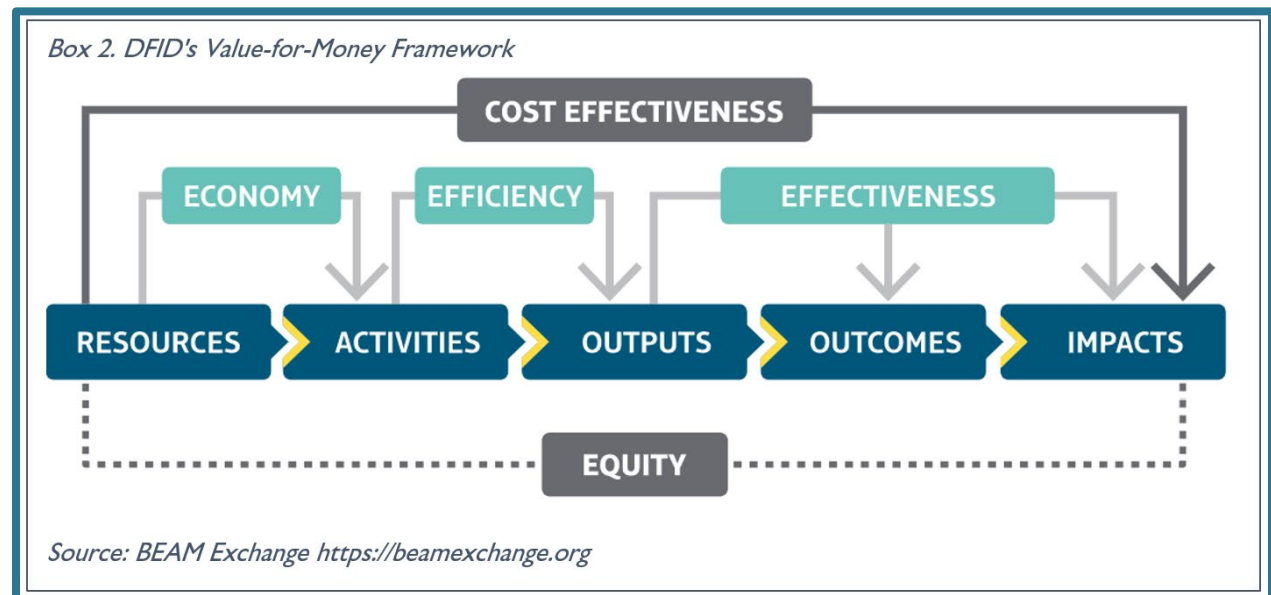
To improve the overall efficiency and effectiveness of donor-funded education interventions, cost data are just as important as data on effectiveness. Estimates of the cost associated with outcomes of education programming, as well as for specific unit costs of program components, would not only allow for improved transparency and accountability of donor-funded programming in the education sector but would also support the ongoing process of practice improvement. Routinely collecting data on actual costs of components of education programming will allow donors, national governments, and education practitioners to better understand the levels of investment needed to produce outcomes of interest and will inform strategic decisions relating to education programming.

Many donors have articulated policies that promote a better understanding of the costs of education programming vis-à-vis their results.³ United Kingdom Department for International Development (DFID) for example, established a framework for measuring costs and results of development programming focusing on three Es: economy, efficiency, and effectiveness⁴. The principle of “economy” refers to ensuring an appropriate balance of price and quality of the program inputs. The principle of “efficiency” refers to how many outputs are achieved for the cost of inputs used. The principle of “effectiveness” refers to the impact achieved relative to the cost of inputs. A fourth “E”—equity—has been added in recent years to ensure that value-for-money analysis accounts for additional resources needed to reach the most marginalized groups. Together, these elements in relation to results indicate whether a particular investment maximizes

³ For example, PEPFAR began collecting unit cost data in 2015. USAID’s Office of Education initiated cost reporting of its interventions in education in 2017, and Millennium Challenge Corporation is tracking expenditure by thematic categories as part of its accountability mandate.

⁴ Department for International Development. 2011. “DFID’s Approach to Value for Money (VfM),” p. 4. Report available on https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/49551/DFID-approach-value-money.pdf.

“value for money” for the donor (Box 2). Fundamentally, the value for money framework aims to provide information for more informed decision making within budget constraints.



The World Bank addresses value for money of all investments by including an economic analysis that addresses three questions:

1. *What is the project's development impact?* Answering this question requires comparing the expected stream of project benefits and expected costs (cost-benefit analysis), in addition to an explicit causal framework linking project activities to targeted outcomes.
2. *Is public sector provision or financing the appropriate vehicle?* This question asks if the public sector should be financing and/or implementing project interventions.
3. *What is the World Bank's value added?* This question seeks to determine the benefit from World Bank staff involvement and whether the proposed project design maximizes the impact of staff effort.

USAID's approach to value for money is largely rooted in the principles of system strengthening and the countries' journey to self-reliance. USAID has issued guidance aimed to standardize expenditure data, outputs data, and leveraged contributions data across its implementing partners. These data will enable USAID Missions and USAID/Washington to conduct routine analyses of cost data to inform our understanding of the current structure of expenditure across the USAID education portfolio, understanding of regional and program-related cost differences, establishing relationships between levels of investment and results, and understanding the role of such mediating factors as the amount of government and/or non-government support, among other things. The findings are expected to inform those who are having a dialog with host country governments on scaling and sustaining education interventions, as well as internally for project and activity planning and budgeting by USAID Missions.

This guidance note builds on the experience of DFID, USAID, and the World Bank with collecting and using cost data. It also is informed by the most up-to-date research that involves cost questions and attempts to address some of the more common challenges with collecting and

analyzing cost data. Some of the analytical methods (cost-efficiency, cost-effectiveness) are drawn from the research literature on economic evaluation, while others (specifically cost economy) are drawn instead from project management practices for development projects. This full range of analytical methods has been included both to match the “4 Es” framework popularized by DFID, and to emphasize how good economic evaluation builds on a basic understanding of costs.

The objectives of cost measurement establish the foundation for what cost data should be collected, how these data should be analyzed and reported, and finally, what type of education

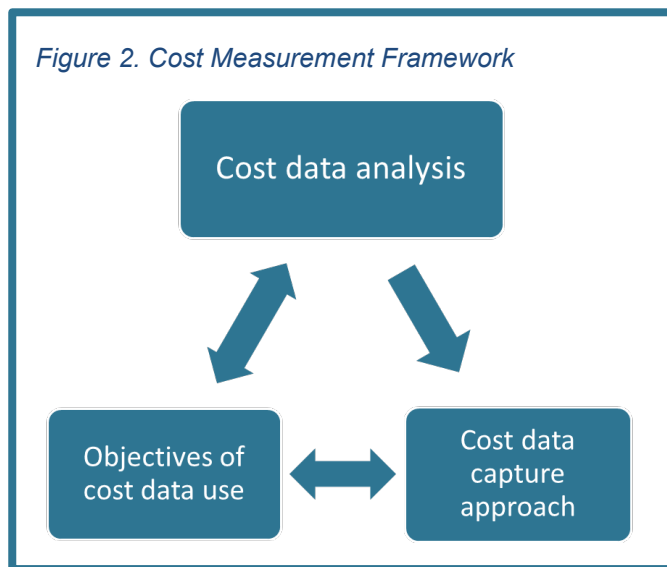
investment decisions these data can and should influence. If, for example, the goal of cost measurement is to identify the most cost-effective approach for achieving an education-related goal, then more than one program must be costed. If, however, the costing exercise is informing a future scale-up of an intervention, then detailed data on units, quantities, and prices must be collected, and fixed costs related to program set-up must be separated from more variable costs related to implementation.

In this guidance note, we use a framework depicted in Figure 2. The framework illustrates how the objectives of cost measurement inform what cost data must

be captured, which in turn determines what types of cost data analysis might be best suited for the purpose and type of data.

The relationship between analysis and objectives is bi-directional since, through reporting, we need to ensure that the cost-related learning questions identified at the onset are answered.

Figure 2. Cost Measurement Framework

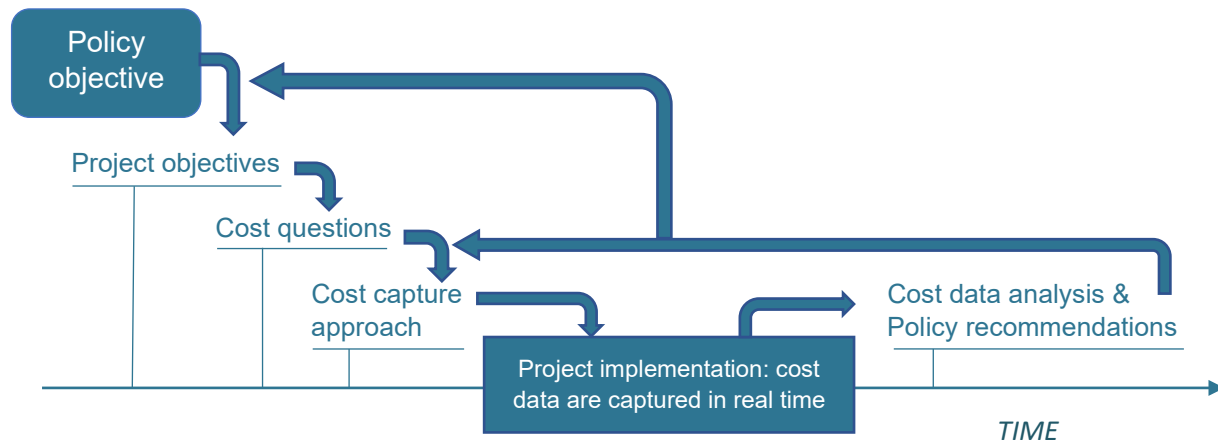


Box 3. Types of cost measurement objectives:

1. Identifying education interventions and system-strengthening approaches that tend to achieve greatest reach and impact per dollar spent in specific country contexts, so that funding can be dedicated to programs that achieve the greatest value-for-money.
2. Estimating unit costs for key education outputs to promote country-led sustainment options and identifying cost drivers and opportunities for improved efficiency and cost savings.
3. Improving budgeting for new country-level donor investments through cost data on the common education inputs and extant data on existing education expenditure;
4. Estimating optimal transfer sizes from central to more decentralized levels of government for national education programs;
5. Helping donors and national governments better understand the differences in expenditure for reaching the underserved and marginalized, both across and within countries.

The framework also has implications for the timing of the activities related to cost data. Specifically, cost data capture would ideally happen in real time, as an intervention is being implemented. Cost data analysis can be performed by analysts using retrospective data, but because of the difficulties of accurately recalling past efforts and the likelihood of not finding crucial data at an appropriate level of disaggregation after program completion, we recommend a real-time cost capture approach (Figure 3).

Figure 3. Timeline of Cost Measurement



Similar to impact evaluation design, the questions that the cost measurement should answer need to be articulated before program activities are implemented to inform how costs of activities are being captured and what results are being documented. Without good cost capture mechanisms in place from the start of a project, accurate estimates of costs will be all but impossible.

For the purpose of this guidance note, “cost” is defined as the totality of incurred monetary expenditure and in-kind contributions that produced a desired education output or outcome. “Expenditure” data include all labor, materials, travel, and all other expenditures, as captured through the accounting system for a reporting period. “Ingredients” data include a breakdown of labor, materials, rent, travel, and other elements for which the expenditure is reported into their respective units, quantities, and prices. In-kind contributions include volunteer labor, materials, and resources donated by other parties and should be included as a part of the overall costs.

3. COST DATA ANALYSIS APPROACH

KEY POINTS:

- Four methodologies of cost analyses are presented: cost-economy, cost-efficiency, cost-effectiveness, and cost-benefit.
- Each type requires specific data and answers particular kinds of policy and programmatic questions.
- Case studies illustrate the applicability of cost data analysis methodologies to specific investment problems.

With the right data in hand from cost reports, education donors and practitioners have several options for the types of cost data analysis they can do, each of which address slightly different questions and require somewhat different data. The choice of analysis type should depend on the policy or investment questions that were originally specified, since no single type of analysis is appropriate for answering all questions. Likewise, there is no single cost data analysis methodology that we should seek to apply to every program. Attempting to apply a particular methodology when the necessary data are not available might result in highly skewed estimates that are not useful for decision making. It is important to note, however, that it is possible to conduct multiple types of analyses simultaneously. For example, cost-economy analysis can be applied within a cost-efficiency or cost-effectiveness study.

Table 1 summarizes the main types of costing analyses and identifies what questions they can help answer, what data are required, and how the resulting estimates can be used in decision making. While these types of economic evaluation have been around for decades (and some, such as cost-utility analysis, were not included here), the BE2 framework builds on the “4 Es” put forth by DFID, and which are already familiar to many development funders and implementers. The framework identifies the types of analysis which tend to be most useful for informing policy decisions around education programs, and clarifies what type of policy question the different types of analysis are best suited to answer. The following sub-sections provide more detail on each type of analysis and include real case studies that used these analyses.

Table 1. Types of cost analyses

Analysis type	What questions can it answer?	What data do you need?	What can cost data analysis help achieve?
Cost-economy analysis	What did it cost to deliver this program? How much was spent on different activities? How much should the	Expenditure and contributions reports disaggregated by cost categories; ⁵ a method for allocating shared costs across	Establish basic value for money of the investment. Help budget for the new program. Help transition program or its

⁵ *Cost category* is defined here as class of costs incurrent to produce a particular kind of education outcome, such as trained teachers or teaching and learning materials.

Analysis type	What questions can it answer?	What data do you need?	What can cost data analysis help achieve?
	government budget for a scale up?	cost categories. ⁶ Output data, using common indicator for all programs.	elements to the government.
Cost-efficiency analysis	What did this program cost per <i>output</i> delivered? How does that compare to other <i>delivery methods</i> for this output?	Expenditure and contributions reports disaggregated by cost categories; method for allocating shared costs across cost categories. Output counts, using common indicator for all programs. Output counts, using common indicator for all programs, disaggregated by delivery methods.	Identify unit costs per output. Compare unit costs across delivery methods and identify value for money of the studied program. Identify delivery method that achieves the most outputs, within a given budget.
Cost-effectiveness analysis	What did this program cost per <i>outcome</i> deliver? How does that <i>compare</i> to other interventions that produce this outcome?	Expenditure and contribution reports disaggregated by cost categories; method for allocating shared costs across cost categories. Credible estimates of program's impact. Credible estimates of the cost-effectiveness of alternative interventions with the same outcome of interest.	Identify cost per unit of outcome. Compare costs of outcome across different interventions and identify value for money of the studied intervention. Identify the intervention that achieves the most outcome, within a given budget.
Cost-benefit/Rate-of-return analysis	How did the costs of this program compare to the monetary value of the benefits created?	Expenditure and contribution reports disaggregated by cost categories; means of allocating shared costs across cost categories. Credible estimates of program's impact on (multiple) outcomes. Economic valuation of the long-term benefits of the program.	Identify whether the studied program was "worth" the investment in monetary terms.

3.1. Cost-economy analysis and sustainability

Cost-economy analysis enables us to assess the total costs of a program, as well as its sustainability and the costs of scaling it up to other areas within a country. This type of analysis focuses on an examination of the components of the program, which inputs went into each

⁶ *Shared costs* refer to costs that support multiple tasks across the intervention. For example, costs of renting a space where project staff work on different tasks.

component, and at what price these inputs were purchased. While not typically thought of as a economic evaluation method, this kind of analysis is nonetheless important for ensuring cost control, and in forming the basis for other analyses which compare costs to outputs or outcomes. Detailed information on pricing of common inputs is very useful for future program design. Detailed data can also help establish whether running the program as designed is fiscally sustainable for local actors, by separating non-recurrent and donor-associated expenditure from recurrent program implementation costs (for example, ex-patriate management costs versus local management costs). Finally, cost-economy analysis can help answer equity-related questions, such as the impact of regional price differences for common education inputs on education access and quality for different population groups.

What can we learn? Knowing how much was spent on different activities within a program and what prices were paid for different inputs provides a good basis for assessing whether goods and services were procured in the most economical way. Since prices for common inputs vary significantly across countries (and often within countries), databases of local pricing would help increase precision of budget estimates for new programs. Cost-economy analysis can also help answer sustainability-related questions. An examination of the cost structure of the program helps isolate costs that would need to be borne if the program were sustained long-term, as well as those expenses that are associated with donor-supported programming, such as donor reporting or technical support travel.

What data do we need? Cost-economy analysis is not possible without detailed cost data. First, all donor program-related expenditure data must be captured in real time, using cost categories that are aligned with analysis questions. Second, operations costs must be separated from management costs, and intervention creation (non-recurrent) costs must be separated from intervention implementation (recurrent) costs. Field costs must be reported separately from foreign head office/expatriate costs. Finally, when an understanding equity implications of price differences (or other equity-related questions) is included as an objective, capturing costs according to additional sub-categories might be necessary. For example, if an objective of the cost-economy analysis is to better understand the price difference between building a school in relatively central locations versus in remote communities, the program will need to track costs of building remote schools separately.

Case Study. With the explosion of mobile phone ownership in low-income countries, donors are considering using them for education intervention delivery. But can mobile phone-based interventions be used to help improve student achievement? In 2015/2016 USAID co-funded⁷ implementation and an impact evaluation of the Makhalidwe Athu (MA) project to test whether using mobile phones would help improve student learning. The MA project was a nine-month pilot intervention aimed at improving the reading skills of 1,200 students in 2nd and 3rd grade in Zambia's Eastern province. The project provided reading materials in the predominant local language and supported reading activities through SMS text messaging to caregivers of children selected to participate in the pilot. Per the condition of the grant, the project captured costs according to USAID's recommended methodology (Table 2).

⁷ The project was funded by the All Children Reading Partners (USAID, World Vision, and the Australian Government), designed and implemented by Creative Associates, and evaluated by National Opinion Research Center (NORC) at the University of Chicago.

Table 2. Makhalidwe Athu Expenditure Data Using USAID Standard Cost Categories

Summary Expenditure* Data	Amount (USD)	Percent
USAID Category 1: General Management and Operations	333,913	39.6%
USAID Category 2: Monitoring, Evaluation and Reporting	118,479	14.0%
USAID Category 4: Teaching and Learning Materials	195,963	23.2%
USAID subcat 4: (Development)	92,252	10.9%
USAID subcat 4: (Promotion and Production)	11,191	1.3%
USAID subcat 4: (Distribution/Transmission)	92,520	11.0%
USAID Category 7: Parent/Community Involvement	146,496	17.4%
USAID subcat 7: (Development)	5,325	0.6%
USAID subcat 7: (Implementation)	141,171	16.7%
USAID Category 11: Complementary Activities	48,669	5.8%
GRAND TOTAL	843,520	100.0%

*The project did not collect contributions data.

The external impact evaluation found the project successful: it achieved moderate 0.27 SD effect size increase in student reading fluency, using difference-in-difference analysis of intervention and control group reading scores at baseline and endline. Computation of unit costs of this project resulted in more than USD700 per student. Recognizing that the small scale and the pilot nature of the project affect the unit cost quite greatly, USAID commissioned an analysis of project cost to establish how much it would cost to scale up the project nationally. The analysis found:

1. If scaled up, the project would be implemented through the Ministry of Education and thus not need to have start-up, close-out, USAID compliance, reporting, and other donor-funded, project-related expenses.
2. Since the intervention was already developed, development costs could be limited to a production and translation to local languages of new stories.
3. With economies of scale, many per learner costs could be reduced dramatically.

The final scale up costs per learner were estimated to be USD20.1 for students in Lusaka or Eastern Province, and USD21.6 for students in other parts of the country.

The evaluator highlighted several considerations regarding these estimated costs. First, MA was piloted in predominantly rural areas. Several things could change in urban areas. Second, cell phone ownership could be higher. This would imply that program's fixed costs could be lower per student, improving the cost-effectiveness of the program. Third, it is possible that the lack of reading materials is less of a problem in urban areas than in rural areas. This could increase the frequency of refusals because reading resources are not in such need; on the other hand, if children have better reading skills in urban areas than in rural areas, the impact of the activity

could be higher as the program had greater impacts on abler students, which would increase the effectiveness of the program per dollar invested.⁸

To conclude, having clear cost data in addition to impact data for this intervention provided crucial data for USAID and the government of Zambia in informing their decision about the scale up of the intervention. Without cost data, impact evaluation findings alone are insufficient to inform sustainment decisions in the context of tight budgets and competing priorities.

3.2. Cost-efficiency analysis

Achieving efficiency means maximizing the programmatic *outputs* achieved (e.g., number of teachers trained, number of books produced), given the *inputs* used in the program. Thus, cost-efficiency analysis provides a good measure of “operational” efficiency. However, there may be a trade-off between cost-efficiency and overall program effectiveness or equity. For example, a program can achieve a low cost of learning materials per student served if it does not translate materials into children’s mother tongue. This might be cost-efficient from the point of view of unit costs of materials, but such a program would unlikely achieve a high impact on literacy. Like with all cost data analysis, cost-efficiency results should therefore be interpreted alongside information about the appropriateness and quality of the program being studied.

What can we learn? Cost-efficiency analysis compares the costs of a program to the outputs created by that program. Such analysis is useful when choosing among alternative delivery models to deliver a given output, or to evaluate the value-for-money of a program during final evaluation. For instance, cost-efficiency analysis would reveal how much it cost per teacher per year if we provide professional coaching through one-on-one mentors, versus through teacher inquiry groups. Cost-efficiency analysis can also help to uncover how contextual or programmatic features drive the cost per output. For instance, how will the cost per teacher of a training program change based on the remoteness of schools? As these examples show, the value of cost-efficiency data is *comparative*: we can judge the performance of a given program or learn about how cost-efficiency is affected by programmatic/contextual features only when you have other programs’ cost-efficiency results to compare them to.

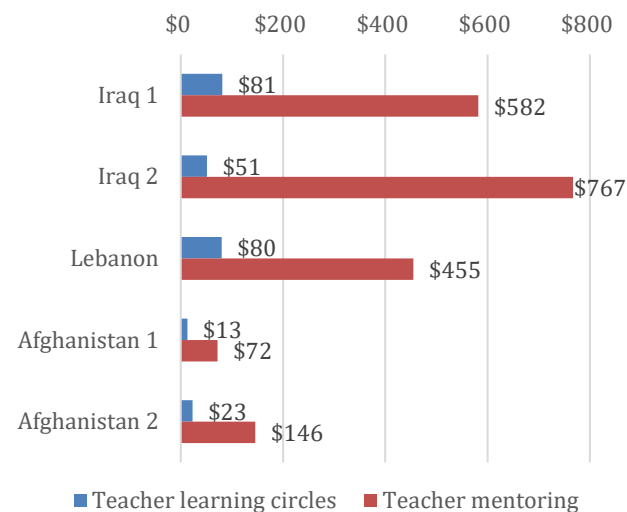
What data do we need? Because cost-efficiency analysis is particularly valuable when it is used for comparing programs, it is really important that the methodology for capturing and analyzing the cost data are the same in the compared programs. To ensure that the cost-efficiency analysis can be applied consistently, highly detailed cost and output data are needed. On one side of the analysis, expenditure data disaggregated by cost category and sub-category (if needed for analysis) as well as “ingredients” in each category (i.e., labor, fringe, travel, other direct costs, fees) are needed. On the other side, each program outputs must be counted using the same metrics. For example, if we are conducting a cost-efficiency analysis of two programs in which one measures the cost of a teacher training program using the students reached as the main metric, while the other program measures cost per teacher, the results cannot be meaningfully

⁸ NORC at the University of Chicago. 2017. “USAID Impact Evaluation of The Makhalidwe Athu Project (Zambia)” Report available on <https://allchildrenreading.org/wp-content/uploads/2019/07/USAID-MA-Zambia-Final-Impact-Evaluation-Report-002.pdf>

compared. Best practice when planning a cost-efficiency analysis is to identify analysis questions at the onset of the program implementation and align both cost and output capture for both programs. For more detailed guidance on how to conduct cost-efficiency analyses, see the references at the end of this guidance note.

Case study: The International Rescue Committee invests in several types of teacher development activities including face-to-face workshops, mentoring, and teacher learning circles (TLCs). The Education team expressed an interest in knowing how much it cost per teacher to provide ongoing professional support through two mechanisms: TLC's and one-on-one coaching. They conducted a comparative cost-efficiency analysis to explore the costs of running these various professional development programs across nine programs in Iraq, Lebanon, and Afghanistan (Figure 4). Most of the programs provided multiple types of support to the teachers they served, and the analysis had to separate the costs of face-to-face workshops, individual coaching, and TLCs to estimate the cost per unit of training for each of them.

Figure 4. Unit Costs of Teacher Learning Circles and Coaching (excluding support costs)



Comparing the two modalities of long-term support for teacher professional development (i.e., one-on-one coaching and TLCs), the results showed that TLCs cost \$49 per teacher per year on average while one-on-one mentoring cost \$423 per year. This provided valuable information for the planning of future programs.

3.3. Cost-effectiveness analysis

Cost-effectiveness is commonly understood as maximizing the outcomes achieved (e.g., increase in student literacy, increase in job readiness) per inputs used in the program. Cost-effectiveness analysis thus uses two pieces of information: the cost efficiency of a program (cost per output) and the effectiveness of the program (impact per output). These are usually combined in a ratio, such as the cost per increase in test scores, but that ratio can be split back into these two components to understand the results more clearly. For example, two programs with the same cost-effectiveness might have very different costs and impact. One could be high-cost and high-impact, while the other one could be low-cost and low-impact. Thus, the investment decision-maker needs both pieces of information to make an informed decision: cost per beneficiary and cost per effect.

What can we learn? Cost-effectiveness analysis compares the costs of a program to the outcomes created by that program. Such analysis is useful when trying to choose which intervention is likely to cause the greatest change in outcomes per dollar spent. In other words, cost-effectiveness combines the information from a cost-efficiency analysis with information from impact evaluations. For instance, a cost-effectiveness study could explore how many additional

days of school attendance are achieved for a given amount of spending from conditional cash transfer programs as compared to school feeding programs. Cost-effectiveness analysis relies on a rigorous estimate of the impact of that program from an impact evaluation using experimental or quasi-experimental methods. The availability of rigorous impact estimates is the major constraint on conducting cost-effectiveness analyses. Although it may be tempting to try to conduct cost-effectiveness analysis using rough estimates of impact, the results will merely reflect whatever assumptions of impact were used in calculations.

As with cost-efficiency, the value of cost-effectiveness data is comparative: the performance of a given program can only be judged when compared with another program's cost-effectiveness results.

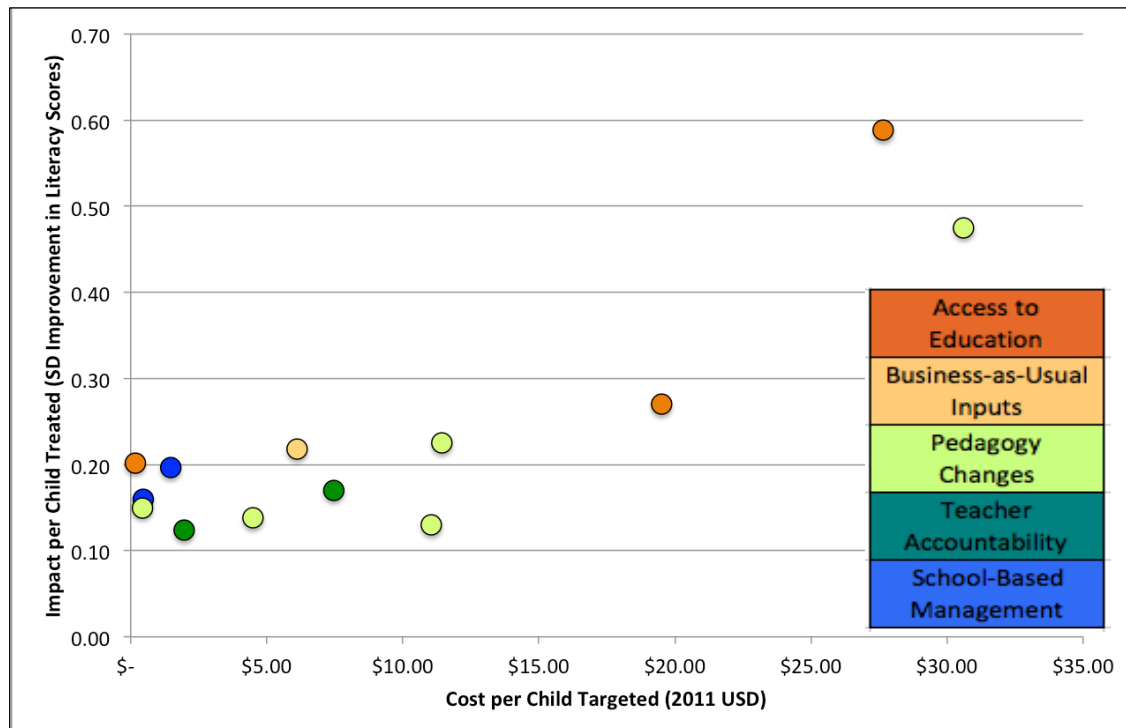
What data do we need? Similar to cost-efficiency analysis, the value of cost-effectiveness analysis comes from comparing programs. Therefore, it is imperative that the methodology for capturing and analyzing the cost data are the same in the compared programs. To ensure that the cost-effectiveness analysis can be applied consistently across the compared programs, highly detailed cost and output data are needed. Additionally, both compared interventions must be rigorously evaluated with a strong counterfactual. The outcomes must be measured using the same impact indicators across the compared programs. For example, if one program measures impact in percentage of students who reach the benchmark of proficiency on a national reading assessment while another program measures impact in improvement in early grade reading assessment (EGRA) scores, the results cannot be meaningfully compared.⁹ Best practice when planning a cost-effectiveness analysis is to standardize impact indicators and ensure that programs collect and report on cost, outputs, and outcomes following an established methodology. For further guidance, please see references at the end of this guidance note.

Case study: Over the past two decades, a new wave of randomized evaluations has examined how developing countries can help children who are not in school gain access to education and those in school improve their learning. A comparative cost-effectiveness analysis by Kremer, Brannen, and Glennerster reviewed the evidence on program impact and present the cost per standard deviation (SD) of change in literacy test scores for 30 programs that had been studied with impact evaluations. We use their data in Figure 5 to plot for each program both cost-per-beneficiary and impact per child.¹⁰

⁹ It should be noted that even EGRA scores might not be comparable across contexts or across sub-populations.

¹⁰ In the paper, Michael Kremer, Conner Brannen and Rachel Glennerster, "The Challenge of Education and Learning in the Developing World," *Science* (340, no. 6130 (2013): 297-300), instead plot impact per dollar spent.

Figure 5. Cost-Effectiveness of Student Learning Programs¹¹



In the figure, we see that the program with the highest impact per child is also the second most expensive program. We also see little relationship between spending and impact in the \$0-\$10 range. How can this information be used in decision-making? We might decide that based on the data presented in this figure any investment below \$20 is not “worth” it since it does not appear to produce a meaningfully, different result from “business as usual.” Therefore, we might advise policy makers to reduce the number of beneficiaries and increase the expenditure per beneficiary. Alternatively, we might know that anything above \$15 per child is not fiscally feasible. Then, we might want to limit how small of an impact we’re willing accept. For example, we might want to forgo anything below 0.12 standard deviations if we consider the effort – even if not very expensive – not worth it.

Note that the interpretation of these cost-effectiveness results did not focus on identifying the single program with the absolute highest impact per dollar and concluding that it was “the best.” Instead, comparative cost-effectiveness information allowed the researchers to look for patterns in the results and assess what was driving high cost-effectiveness. It also allowed the researchers to assess how reliably different models produced results across contexts. When using this information to make programmatic choices in a particular context, a policymaker would still need to pair it with information about whether approaches were feasible and appropriate in that country and education system.

¹¹ The figure is based on data from Michael Kremer, Conner Brannen and Rachel Glennerster, “The Challenge of Education and Learning in the Developing World,” *Science* 340, no. 6130 (2013): 297-300.

3.4. Cost-benefit analysis

Cost-benefit analyses (sometimes called a rate of return analysis) compare the total costs of an intervention (including costs to society) to the monetized value of the intervention's benefits (including social benefits accrued to those who did not directly benefit from the intervention). In contrast to the methods described above, it is not inherently a comparative methodology; we compare the costs of a program to the benefits of that same program.

What can we learn? Since both cost and the benefit are expressed in monetary terms, the comparison of cost and benefit is straightforward and the conclusion of whether the program was “worth it” is unambiguous. As such, cost-benefit analysis is well suited for certain types of questions. Two examples are programs with multiple outcomes that cannot be meaningfully measured with a single metric and when a program entails a large one-time investment that could be justified by impacts that accrue over a long period of time.

What data do we need? Conducting a cost-benefit analysis has stronger data requirements than other cost analyses and requires assumptions about how a program's effects will play out over time. In addition to the impact evaluation evidence that is required for cost-effectiveness analysis, cost-benefit analysis requires impact evaluation evidence for *multiple* outcomes (all possible outcomes influenced by the intervention under study) and typically needs impacts measured over longer periods of time. While it might be possible to calculate program costs and benefits using speculative estimates of impact taken from pre-post analyses or different contexts, this can also significantly increase the margin of error on results. However, one advantage of cost-benefit analysis is that comparative data from other programs is not required since the results from one program can be interpreted by itself and since cost-benefit analysis tells us whether or not (given our assumptions) the value of that program exceeded the cost.

Case study: Angrist et al. (2002) present a cost-benefit analysis in their evaluation of a school voucher program in Colombia – the *Programa de Ampliación de Cobertura de la Educación Secundaria* or PACES program. Vouchers for private secondary schools were distributed via lottery, and Angrist et al. exploited this randomized assignment of voucher offers to estimate the impact of vouchers on private school attendance, school completion, test scores, and the number of hours a student spends in employment.

To compute costs, they start with public expenditure on the program. The average value of the voucher received by lottery winners was \$74. The authors also recognize, however, that when students increase their private school attendance, they also free up slots (and therefore resources) in public schools. Since the average per-pupil cost of a public school slot was \$350 and since lottery winners were 14 percent less likely to attend public school after receiving a voucher, the government spent \$50 less in public expenditure per lottery winner. Thus, the total additional public spending per lottery winner was \$24 (\$74-\$50).

Households, however, might also have incurred costs as a result of the program. Some fraction of them had to pay for private school that they would not have done in the absence of the program. Similarly, some of them might have lost household income as their children reduced the amount of time spent working. It turns out that on average, lottery winners only spent \$52 more in school fees, despite receiving a voucher of \$74, on average. Their children earned an average of \$41 less through work, so lottery winners net contribution to the program was \$19 (\$52-\$74+\$41).

Total society-wide expenditure per year per household was therefore approximately \$43 (\$24+\$19). The program ran for three years and voucher take-up declined over the life of the program. (It was 88 percent in the second year and 49 percent in the third year.) Since lower take-up increases program costs per beneficiary and since the program ran for three years, an upper bound estimate of the total cost of the entire program is \$195 per household.

To monetize the benefits of the voucher program, Angrist et al (2002) start with the estimated effects of the program on school completion (an additional 0.12-0.16 years) and test scores (an improvement of 0.2 standard deviations). Assuming that the gain to the program come only from the economic return to 0.12 additional years of education and using data from a Colombian labor force survey on the returns to an additional year of education, the authors estimate an increase in lottery winners' earnings of \$36 per year. When, however, the authors consider that the 0.2 standard deviation increase in learning gains is roughly equivalent to an entire year of schooling (in a sample of Hispanic students in the United States that took the same exam), PACES might have raised earnings by \$300 per year. Since these earning increases accrue over students' entire working lives, while the costs of the program were limited to three years, the benefits of the voucher program undoubtedly outweighed the costs of the program.

4. COST CAPTURE APPROACH

KEY POINTS:

- **Successful cost measurement depends on availability of accurate and relevant cost data. This section presents activity-based costing as an easy to implement cost capture approach for donor-funded education programming.**
- **Options for integrating cost capture into government reporting systems are discussed.**

4.1. Cost capture approach for donor-funded interventions

Cost data analysis methods described in the previous section help address very specific cost-related questions to inform donor investment or a country's reform agenda. Articulating these questions is an essential first step, as shown in the cost measurement framework in Section 2. Questions must be based on how the investment is structured and what it is trying to accomplish. Suitable questions may include:

- Is this early childhood education program fiscally sustainable for the government to carry on after initial donor investments in curriculum development and training?
- Would a different scale-up strategy allow for greater coverage of children within a given budget?
- What changes in primary grade reading will this level of investment be able to achieve?
- What is the difference in costs of building a school in remote communities versus those close to urban areas?

As we argued in Section 2, which *question* we are asking will determine which cost data analysis method we will need to use, but it will also determine what *cost data* are needed.

This guidance recommends the **activity-based costing approach** for the cost capture step of cost measurement. This approach is based on activities, which are any event, unit of work, or task with a specific goal¹²—such as conducting teacher training, developing or producing books or other learning materials, engaging in a policy reform, or undertaking a learning assessment. The activity-based costing approach lends itself well to international education development since the number of education-related tasks is not large and is fairly universal across countries and education systems. Which activities should be costed will depend on policy interests and learning objectives that the national government or the donor is seeking to answer.

Activity-based costing approach requires establishing clear, unambiguous categories of tasks that will be used to categorize costs. There may be different ways to categorize education costs. Under UN's Classification of the Functions of Government,¹³ for example, the education sector includes the following broad categories of expenditure:

- Pre-primary and primary education
- Secondary education
- Post-secondary non-tertiary education
- Tertiary education
- Education not defined by level
- Subsidiary services to education
- Research and development in education
- Other education-related functions

These broad categories are useful for an overall understanding of the level of investments at each education level, but they do not help us understand the effect targeted investments or system tweaks may have on student achievement and other outcomes of interest at a given level of education in a specific country. For example, a donor might support primary grades teacher training. We will see that the overall level of inputs increased in the category of primary education, but it does not tell us where the investment was made and what results we are to expect. To achieve objectives outlined in the opening section of this guidance note, the donor community, their partners, as well as national governments need to have more granular data on the cost of components of the education system at any of the levels identified by the UN classification. So using our earlier example, if the government identifies early grade reading as an area in need of support and the donor community invests in new teaching and learning materials and teacher training, we would want to assess the impact of this investment by measuring learning gains in reading as opposed to measuring other metrics, such as learning gains in math or school readiness of children who have received pre-primary education. This level of detail about how the investment was used and what the results were is necessary for learning what works in

¹² "Activity-Based Costing (ABC)," Investopedia, accessed October 8, 2019, <https://www.investopedia.com/terms/a/abc.asp>.

¹³ "Manual on sources and methods for the compilation of COFOG statistics - Classification of the Functions of Government (COFOG) - 2011 edition," Eurostat, accessed October 8, 2019, <https://ec.europa.eu/eurostat/en/web/products-manuals-and-guidelines/-/KS-RA-11-013>.

education development and what types of interventions seem to produce better results. Detailed data on cost as well as results of donor-funded interventions is also a prerequisite for successful scale and sustainment of effective interventions by host governments.

An additional categorization of activities common to every education system can help achieve these objectives. This guidance puts forth the following activity-based categories for cost measurement that can be applied to most of the UN categories noted above:¹⁴

- Category 1. General operations, management, and reporting
- Category 2. Assessments and evaluations
- Category 3. Pre-service teacher training
- Category 4. In-service teacher training
- Category 5. Teaching and learning materials
- Category 6. Strengthening systems and accountability
- Category 7. Private sector engagement
- Category 8. Parents/Community engagement
- Category 9. Safe schools and infrastructure
- Category 10. Grants, scholarships, and cash transfers to individuals/families
- Category 11. Grants to organizations
- Category 12. Other

Examined alongside common education system performance indicators, such categorization of expenditure within each UN education sector level will allow us to achieve a better understanding of how resources are being allocated to improve system performance. Policy makers and donors will be able to use longitudinal data on donor investments and system response to more accurately assess the value of their investments and compare levels and trends across projects and country contexts. Such data will inform country education planning and help donors as well as governments prioritize future investments in education development.

Annex A of this guidance note defines each category, describes how to select relevant categories for measuring costs, and lists what sub-categories each category may include. Sub-categories can be used when we want to separate cost of creating an intervention (non-recurrent costs) from the costs of implementing an intervention (recurrent costs), to allow for more accurate computation of scale up and sustainment costs once the intervention is transferred to the host country government. Sub-categories are also essential when we want to compare cost-effectiveness of different modalities of the intervention. For example, if we want to compare cost-effectiveness of cascade model teacher training versus individual coaching, both would fall under “in-service teacher training,” so without dedicated sub-categories we would not be able to isolate costs of each modality.

The learning objectives of a costing analysis will inform the cost capture structure and the appropriate levels of required disaggregation. The steps below illustrate this process of selecting cost reporting categories.

Step 1. Based on the policy and strategic objectives guiding the investment, establish cost questions that could be answered using cost data from the implementation of the

¹⁴ These categories emerged from USAID’s experience with activity-based costing of its education portfolio. USAID’s cost guidance can be found here: <https://www.edu-links.org/resources/usaid-cost-reporting-guidance>.

intervention. For example, a strategic objective of the investment might be improving quality of education by strengthening pre-service teacher training. An associated cost question might be: “How much would it cost the Ministry of Education to sustain continuous professional development for teacher educators in pre-service teacher colleges?”. This question can be broken down into cost estimates for annual professional development events for teacher educators, estimates for production of new professional development materials, and estimates for sustainment of teacher educator evaluation system.

Step 2. Match objectives of the investment and cost questions with cost categories; select the most appropriate cost categories while minimizing the total number to reduce the reporting burden.¹⁵ In our example, this means the project would select “pre-service teacher training,” “teaching and learning materials,” and “strengthening systems and accountability” cost categories to capture costs associated with the intervention, in addition to the mandatory “general operations, management, and reporting” category.

Step 3. Determine which cost categories will need sub-categories. Sub-categories are critical for answering certain cost questions, however, like in Step 2, it is important to find a balance between collecting highly detailed data and the reporting burden. Common sub-categories within thematic cost categories could be “creation of intervention,” “implementation of intervention,” “policy,” and “research.” In our earlier example, we might want to know how much it costs to produce materials for the annual professional development for teacher educators so that the Ministry of Education can budget for this after the end of donor assistance. This means that the project would need to capture costs of developing materials separately from the costs of printing and distributing materials, using dedicated sub-categories in the “teaching and learning materials” cost category.

Step 4. Once the categories have been selected, the project management needs to train their staff on correctly using them. To ensure accuracy of data, costs must be captured in real time during the project implementation.

Capturing cost data on ingredients¹⁶ of an intervention includes both the project expenditure and in-kind contributions by government and non-governmental partners to that intervention. Capturing estimates¹⁷ of major monetary and non-monetary contributions to the intervention from other source includes specifically:

1. Estimates of host government contributions above the fixed costs of the education system;
2. Estimates of non-government, corporate, and individual contributions.

¹⁵ The description of activities included in each of the cost capture categories may help the implementing partner staff and its subcontractors to bill their time correctly. Such description should be developed in the first quarter of the project following the award and the selection of cost tracking categories and should be included in the project’s policies and procedures manual.

¹⁶ Ingredients data refers to a breakdown of labor, materials, rent, travel, and other elements for which the expenditure is reported. In-kind contributions include labor, materials, and resources donated by other parties and should be included as a part of the overall costs.

¹⁷ Estimates can be derived from a variety of sources, including verbal and written communication by authorized representatives of the host government; documentation and written communication by authorized representatives of NGOs and private entities; and implementing partner documentation.

We recommend that only contributions be captured that are essential for program implementation and that otherwise would have to be purchased. Only direct contributions that are distinctly different from the pre-project education delivery should be captured. For example, if a donor-funded intervention includes modification of classroom instruction by teachers, teacher salary is a part of existing fixed costs and should not be captured as part of intervention costs. On the other hand, if new teacher aids are hired to support implementation of the donor-funded intervention, the salaries of those teacher aids should be included. An example involving non-government contributions could be donations of equipment to a vocational training center. If the donated equipment is essential for the training of enrollees of the center and would have to be purchased otherwise, then it should be reported. If, on the other hand, it is not essential for achieving training objectives, then it does not need to be reported. Recommendations on how to capture and report on these contributions are provided in Annex B of this guidance note.

Additionally, if other donors are funding projects that aim to achieve similar results among the same population of beneficiaries, it is important to document the presence of these projects, the scope of their activities, and other relevant information. Reporting on these contributions should be based on written communication with contributors and on existing documentation, and it might be imprecise. We do not recommend that implementing partners be expected to engage in any verification activities to establish accuracy of contributions reported by others as it might be overly burdensome and detract resources from the primary objectives of the investment. Reporting on contributions of others is understood to be an estimate and should not be subject to an audit.¹⁸ When monetized, these estimates of host government, non-government, and other donor contributions will make overall cost estimates of the donor-funded activities more accurate.

Box 4 lists key principles of cost reporting.

¹⁸ Some donors have internal requirements pertaining to reporting on contributions. For example, under USAID regulations, contributions under an approved cost-share plan are auditable. Reporting on contributions as a part of total cost reporting is different and not intended to be audited.

Box 4. Recommended key principles¹⁹ of cost reporting:

1. Cost data are captured in real time, not retroactively.
2. To allow for comparisons across contexts, the main expenditure categories are standard and should not be re-named or collapsed. Sub-categories may be added to answer more nuanced cost data analysis questions. Categories should be selected based on investment objectives and anticipated tasks. When doing an evaluation, cost category selection should be informed by the evaluation design.
3. The entirety of the expenditure must be captured; expenditure should be reported in the category nearest to its intended result.
4. Both amount of expenditure *and* expenditure type (“ingredients”) should be captured. Ingredients include salaries and wages, equipment and supplies, rent, travel and per diem, participant costs, grants under contract/award, and overhead recovery.
5. “Project development” sub-category should be used to capture expenditure to create a new intervention.
6. The final list of categories and sub-categories should balance the desire for precision with the need to reduce burden of using too many different sub-categories.
7. Estimates of contributions of the government, NGOs, and private actors should be documented and reported.
8. Costs categories should be collected for a well-specified time period (for example, a school year or a specific cohort of beneficiaries like “students entering 4th grade in 2016”).
9. Details of the investment-funded outputs as well as available data on outcomes should be captured and reported alongside cost data, to enable unit cost and cost effectiveness analyses.
10. To accurately assess how much is spent on large categories and for each level of education, a cross-reference to the UN’s Classification of the Functions of Government category corresponding to the investment purpose should be included in all cost documentation associated with the investment. This will also permit comparisons across contexts.

4.2. Cost capture approach for government-funded programs

Official Development Assistance to national education budgets includes bilateral assistance from donors such as USAID and DFID as well as multilateral assistance from organizations such as the World Bank Group. One of the challenges that recipients of Official Development Assistance (ODA) face is differences in reporting requirements across donors. Harmonization of reporting requirements for grants and investment loans in the education sector would allow us to both reduce the reporting burden for the governments and allow for sector-level cost data analyses.

For government programs and large investment loans made by donors to governments, we propose using the same key principles of activity-based cost reporting outlined in Section 4.1. Costing a program implemented by a government, however, brings unique challenges, and some compromises may need to be made.

First, these types of investments and programs tend to support large multi-component projects implemented by multiple levels of government, and isolating intervention and activity costs in a standardized way is difficult and often cannot be done accurately by solely analyzing costs at the level of donor reporting (federal government to donor government). There are usually intra-

¹⁹ For additional information and clarification, please see Annexes.

governmental flows of funds and resources between different administrative levels of government, and these flows, as they relate to the set of investments and interventions in question, must be mapped in order to start thinking about capturing costs. After identifying policy goals, it is necessary to identify what data is available, at how many administrative levels of government data needs to be collected, and how realistic it will be to collect disaggregated and intervention specific data at each level.

Second, multiple levels of government rarely maintain activity-based reporting among themselves (as opposed to the reporting of very aggregated cost categories), which makes it difficult to obtain intervention-specific cost data at the level of donor reporting (the federal government). To collect costs, standardized reporting and cost collection systems need to be established at provincial/state, district/municipal, or even the community level. Getting these levels of government to adopt rigorous and common reporting practices can be challenging. Given the many levels of funding and expenditure reporting, it is critical to avoid double counting inputs to an intervention by counting the same inputs at different levels of government. We recommend trying to measure costs at the level of procurement and implementation, not at the level of transfers/financing.

When activity-based cost reporting proves too difficult to integrate into a government reporting structure, an alternative is to treat costing as a research exercise, rather than a reporting exercise. For example, setting up surveys for capturing costing data may be necessary. A member of the donor organization with a strong understanding of costs and a member of the government receiving the loan or grant working directly on the budgeting and expenditure reporting would need to commit some considerable time to developing a data collection effort.

Cost data collection can also be done alongside an impact evaluation given the time spent in the field to collect data for the evaluation. If an impact evaluation is planned, a cost evaluator should work with the principle investigators to integrate cost relevant questions into the baseline and endline surveys. Existing budget and expenditure data can then be disaggregated by a member of the impact evaluation team through interviews with project managers and finance and procurement personnel. Additionally, direct observations and qualitative interviews with communities may also be necessary to uncover information on frontline inputs to an intervention.

5. HOW TO USE COST ANALYSES

KEY POINTS:

- Comparative cost analyses help identify contextual and programmatic features affecting efficiency and effectiveness of investments.
- Many factors may affect external validity of cost analyses, including input pricing, existing capacity, and scale of implementation.
- Cost-efficiency and cost-effectiveness estimates are for a specific dosage and scope of the intervention. Changes in dosage and scope can result in unpredictable changes in results.
- Including equity considerations in cost capture may help estimate unit costs of reaching the most underserved.

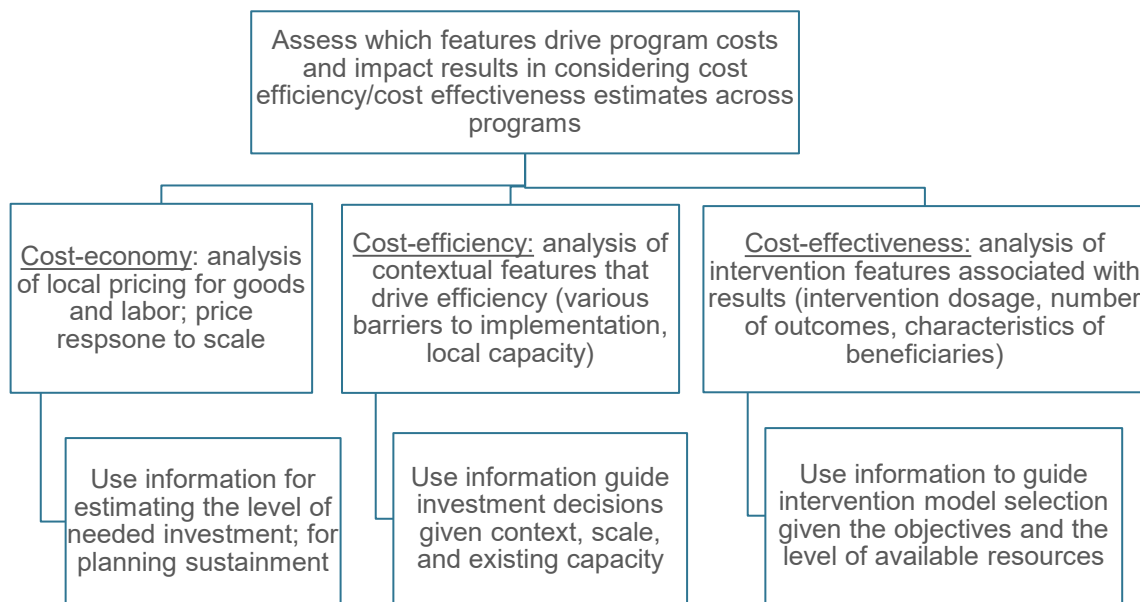
5.1. Learning from cost data analysis results

How results from cost analyses are used is as important as how they are conducted; misapplication of the results about a single program, or a comparison of several programs, can undermine the usefulness of cost analyses. Finding that one program is more cost-efficient or cost-effective than others doesn't mean that program is "the best" and should be replicated everywhere. Rather, **the importance of cost analyses is in uncovering what characteristics of different programs, or contexts, drive the cost or improve value-for-money of education programs.** Knowing not just whether or not a program is cost-efficient/cost-effective, but *why* is crucial for applying these insights well.

To take an extreme example, studies have shown that providing deworming medication to primary school children is a highly cost-effective means of reducing absenteeism. The underlying assumptions are that (1) in places where worm prevalence is high, this is a major driver of absenteeism, and (2) treating worm infections can be done cheaply through mass school distributions. The fact that deworming was extremely cost-effective in Western Kenya, however, doesn't guarantee that it would be cost-effective in an area where worm infections are not endemic. Rather than assuming that the results generalize to a new context, we would want to assess whether or not the conditions that drove cost-effectiveness in Western Kenya were true of the place we were hoping to see improvement on the same outcome.

Thus, while we have to apply consistent metrics and methods when conducting cost analyses, we should not necessarily be expecting to see consistent results across different programs in different contexts. As comparative cost data analysis becomes more routine in the education sector, our first priority should be to learn what features drive or constrain value for money. Sometimes the drivers will be contextual (e.g., low population density, high security costs), and while they can't be helped, data on the implications of these contextual factors will help to plan and budget better. In other cases, the drivers of value for money are changeable features of education programs, in which case we have derived insights about how to improve our programs to improve cost-effectiveness.

Figure 6. What we can learn from cost data analysis results



As discussed above, the point of comparative cost data analysis is not that some analytical method will guarantee the applicability of the results to every other situation. This is simply impossible because of the differences in price levels, institutional quality, population needs, among many other factors. Rather, the goal is to understand what features of context and intervention drive the costs and results of programs.

To return to the example of deworming: One of the key features that drove the cost-effectiveness of the Kenyan program was the feasibility of mass school-based deworming campaigns, which kept costs very low. In a worm-endemic country, you might be confident that worms were driving school absenteeism, but you would still need to assess whether or not the educational system could reasonably conduct a mass deworming campaign. If there were no reliable means of distributing drugs to schools, and you had to use other distribution channels, this would likely increase the cost of deworming campaigns and decrease the cost-effectiveness of deworming as an education intervention.

An analysis of pricing, context, and intervention features is useful at any stage of donor decision-making: when managing an existing program, when planning for sustainment of a successful program, and when assessing program's viability in a new context. The next three sections will explore them in greater detail.

5.2. Understanding Cost Drivers

Input prices vary greatly and sometimes unexpectedly. The common cost drivers of education inputs include how well infrastructure is developed, safety of movement, accessibility/proximity of the country to major international hubs and routes, and availability of qualified labor. The latter in

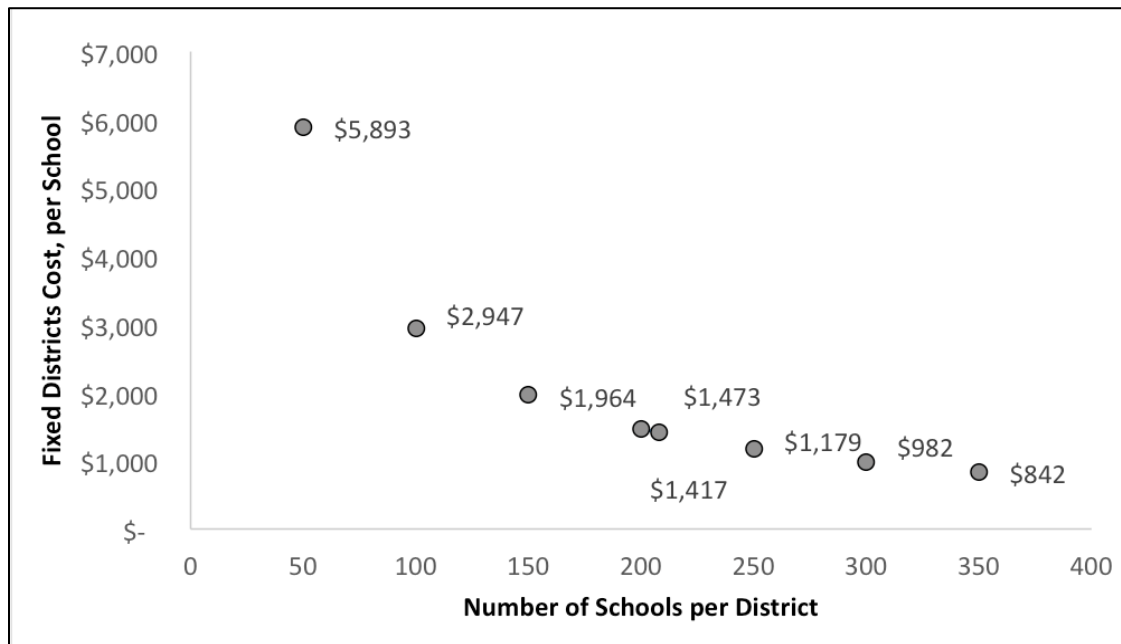
particular is a major cost driver of donor-funded education interventions. Lastly, program scale will impact the overall cost and the unit costs per output and outcome.

- Price of inputs: Price levels vary from country to country, in particular the price of goods such as workbooks, vehicles, venue rentals, etc. While it may be too difficult to figure out the exact price of every good in the place in which you're hoping to run a program, you can identify the top five inputs that make up most of your program cost (so long as you have good cost capture data) and find information on the prices of those goods in the new context. This will help assess if the cost is likely to be higher or lower than in the originally studied context.
- Price of labor: Price levels vary from country to country, and the price of labor (i.e., wages) is particularly important to consider. In some countries, especially post-conflict countries, the prices of skilled labor (e.g., master trainers, program managers) may be unusually high. In other countries, especially rapidly growing economies like India, the price of skilled labor may not be as high. While it may be too difficult to figure out the exact wages of every position necessary to run a program, you can identify the top five staff positions that make up most of your program cost (so long as you have good cost capture data) and find information on the likely wages for those positions in a new context. This will help you assess if the cost is likely to be higher or lower than in the originally studied context.
- Program scale: The unit cost of produced outputs and outcomes (e.g., cost per unit of improvement in student learning outcome of interest, cost per trained teacher) may change quite a lot when a program is implemented at a small scale versus a large scale. This is because of the start-up costs and the costs for developing/piloting a new intervention. Such costs will be incurred regardless of the scale of the program. When programs reach only dozens of schools and hundreds of learners, these costs aren't "spread out" over very many outputs/outcomes, and so the unit cost per output/outcome tends to be quite high. But as the number of reached students and schools increases, the total cost per output/outcome can become much lower.

Figure 7 illustrates this point by showing how the program cost per beneficiary drops as the number of schools participating in the program increases. When the program reached only 50 schools, the total cost per school was nearly \$6,000. However, when the number of schools in the program reached 350, the cost per school dropped to less than \$1,000.

Figure 1. The effect of program scale on a program cost per beneficiary²⁰

²⁰ Data are from IRC cost-efficiency analysis of USAID's Pakistan Reading Program.



5.3 Considering Context

Context affects implementation costs and results in many ways. Economic, political, and cultural factors interact with the intervention and need to be clearly understood when planning and managing an education program and when considering adapting a program to a new context. In-depth understanding of context factors affecting the program will help interpret results of cost-efficiency analysis and evaluate how they might change if applied to a different context.

- **Barriers to education:** Results of any program are contingent upon beneficiary participation in the program. If target beneficiaries cannot access the program because of systemic or circumstantial barriers to education, it will not matter how good the program was—we will not see a result. Barriers to education must be considered as a part of the existing education system analysis. Cost of the program implementation might change considerably if access issues are to be addressed.
- **Cultural or social barriers to implementing this model:** Additionally, a consideration must be given to potential barriers in implementing this specific model. There may be cultural constraints associated with program implementation, such as pre-existing socio-economic, political, religious, and linguistic conditions that conflict with the fundamental assumptions of the program. For example, a mother tongue-based reading program might be successful in improving early-grade reading in one context while parents might prefer to have English language-based program in another context.
- **Local capacity:** The costs of a program will be influenced by the existing capacity of the local institutions. In the deworming example, above, the low cost of mass deworming campaigns in Western Kenya was due to the capacity to distribute deworming tablets throughout the existing school system. When considering the cost drivers of the existing program and how the costs of a program may change if implemented in a new context,

the relationship between cost-efficiency and local capacity needs to be well-understood. We need to consider what existing structures or capacities the program relied on in the context in which it was studied, how a change in those capacities may affect the program delivery, and whether those are present in the new context. If they are not available, we may conclude that the program is not feasible in that new context, or want to figure out how much it would cost to build those capabilities and include it in our assessment of value-for-money.

5.4 Considering Intervention Delivery

The details of intervention delivery are critically important when analyzing program's cost-effectiveness, comparing cost-effectiveness estimates of multiple programs, and considering implementation of a program in a new context. Three aspects of intervention delivery are of particular importance: program dosage, how many different outcomes the program is trying to achieve, and what beneficiaries the program is reaching.

Program dosage. Any assessed program achieved its results because it was able to interact with beneficiaries long enough to effect a change in knowledge, attitudes, or skills. Any intervention has an optimal dosage under which the program achieves maximum results. If we reduce the dosage, we will not get the improvement we are looking for; if we increase the dosage, we are likely to spend more money than necessary. To help assess whether or not we can reach the necessary dosage of treatment that drove high value-for-money in a prior program, we need to assess the probable structure and scale of the new program. If we have a small amount of funding that is expected to cover a large population, then we can only afford a program with a fairly low cost per beneficiary. For example, a behavior change communication campaign is a low-cost, low-impact intervention that can be used to reach larger number of beneficiaries with a small budget and small effect.

An intervention that was found to be cost-effective but had a high dosage of treatment and high cost per beneficiary would not work the same way if the cost per beneficiary were reduced. For example, if teacher coaching was found to be effective at a price point of \$100 per teacher per year, then a reduced budget of \$50 per teacher per year will likely not produce half of the original impact on student learning. In fact, the intervention would need to be tested again to measure the impact and cost-effectiveness given the new dosage. A "safer" bet might be to keep the intervention intact but reduce coverage in response to budget constraints.

Box 5. Relationships between level of investment and level of results is non-linear

One particular challenge in dealing with cost-effectiveness estimates is that they are often presented as single ratios: \$200 per teacher, or \$550 per additional proficient reader. This might make it seem as if you would get that average return at any level of expenditure. Let's say I manage a program in which the total price of inputs is \$100 per learner and which had a 10percent average gain in student scores on a standardized assessment. Suppose I want to expand the program to include more children even if it means they may not benefit as much. I assume that if I cut the "amount" of program a learner receives by half, I might expect that it would cost \$50 per child and we will have 5% growth in learning instead of 10percent in the original study. Right? Wrong. This is a misinterpretation of the original cost-effectiveness estimate - costs and impacts do not scale up and down in a linear manner, nor do they scale up and down at the same rate. The section above on scale briefly covers how the presence of start-up costs and fixed costs means that the average cost per person treated may not scale in a linear fashion.

The issue of relationship between dosage and results requires close attention when program effectiveness is considered. To take a stylized example, we could give five children one book, or we can divide those five books among 100 children. In the former case, these five children can read the book and learn something from it. In the latter case, each child receives only 1/20 of the pages in the book - a dosage of treatment that is unlikely to help them learn anything. Thus, while the cost-per-child of the second program is 1/20 the cost of the first program, the impact of the second program is not equal to 1/20 the impact of the first program; it is zero. Therefore, we should not assume that the relationship between the cost and the effect will be linear, and commonly used cost-efficiency and cost-effectiveness estimates should be interpreted accordingly.

Programs with multiple outcomes. One general challenge in comparative cost data analysis is assessing programs with multiple outcomes. The result of cost-efficiency and cost-effectiveness analyses of different programs is the cost per output or outcome achieved. For instance, we might want to compare cost-effectiveness of two primary grade reading programs, one of which focuses on training teachers in reading pedagogy while the other trains teachers in both reading and math. The latter program might be more expensive per unit of outcome in reading achievement but might be less expensive if we consider both sets of outcomes. Similarly, providing scholarships for girls may improve their attendance at schools and the cost effectiveness of such a program could be expressed in terms of cost per additional years of attendance. But scholarships may also free up household income which could in turn improve the nutritional status of children living in the household, an additional outcome that is "purchased" with the same dollars as the increased attendance. While the cost effectiveness measure accurately captures the value (in terms of *attendance*) for the money spent, it does not capture the *entire* value of the program.

There are three paths for how to think about comparative cost data analysis of multiple-outcome programs:

- 1) Interpret cost efficiency and cost effectiveness findings alongside other critical information: Comparative cost data analysis is well suited to identifying which programs make the greatest progress toward one particular goal, but they are not and should not be used as a holistic measure of a program's value. Rather than trying to

- fit every possible dimension of program quality into the value-for-money metric, we instead recommend pairing information about cost-efficiency/cost-effectiveness with information about other impacts this program may have, as part of the wider value-for-money analysis of the program.
- 2) Create an index measure for outcomes: It is possible to create an index measure of the different outcomes of a program, weighting them by how valuable they are considered to be. For example, the health field uses “Disability Adjusted Life Years” as an index measure of the impacts of reductions in different diseases, based on the “disability weight” of how unpleasant it is to have different diseases. While appealing in theory, cost analyses based on index measures tend to be difficult to interpret, and it is almost impossible to find other programs to compare them with. Also, there are no agreed upon index measures that encompass all of the possible outcomes in education.
 - 3) Conduct cost-benefit analysis as well: Cost-benefit analysis, where a monetary value is assigned to all of the different impacts, is the one methodology which is suited to capturing impacts on multiple outcomes. However, as discussed above, cost-benefit analysis requires more assumptions about the monetized value of benefits and how that might change over time. As such, results from cost-benefit analyses are difficult to compare across programs, and cost benefit analysis should only be done when the relevant questions are about the overall “return” of a program.

The problem of using cost data analysis to assess programs with multiple outcomes is analogous to the problem of comparing houses that you are considering purchasing. The cost per square foot is a standard, intuitive metric that can be applied to all of the homes you might consider buying. But this metric hardly captures all of the relevant features for buying your new home: you also care about how nice the interior is, how big of a yard it has, where it is located, etc. And while you could in theory come up with a quality-adjusted, yard-adjusted, neighborhood-adjusted measure of cost per square foot, it will become increasingly difficult to calculate and interpret that measure. In the end, most people consider the cost per square foot, *alongside* other information about quality and location, rather than trying to pack it all into a single measure.

Thus, the challenge of multiple-outcome programs can be seen as a part of understanding cost data analysis results as one of many pieces of information about a program. Cost data analysis cannot and should not be used to encapsulate all of the information about the value of a program. Some programs may have benefits that are not captured in cost-efficiency or cost-effectiveness analyses, some programs may be feasible or appropriate only in certain environments, and so forth. Cost data analysis is not a substitute for these other crucial decision-making inputs; it is a complement to them.

Characteristics of beneficiaries: The costs and the results of an intervention will greatly depend on the type of beneficiaries it is trying to reach. Improving learning outcomes of marginalized students, for example, could be significantly more expensive than improving learning outcomes of students who are not considered at-risk. You might be able to reach this group by tweaking the program slightly in a way that could increase cost. For reaching young girls, this might mean modifying school construction to include sex-separated bathrooms. While this increases the costs of construction, it also ensures that the benefits can be experienced by both boys and girls. In other cases, you may actually need specific program models to reach particularly vulnerable

groups because typical program models are simply not effective for this population. On the other hand, these groups often benefit more from untargeted interventions, as is the case for girls' education (Evans and Yuan, 2019, 8). Geography, disability status, language, parental literacy, and socio-economic status are all important factors that could influence both the cost and the effectiveness of an education intervention in either direction. Most donors place an emphasis on reaching the poorest and most vulnerable with their investments and would not want to achieve cost-efficiency in their programming by targeting those easiest to reach.

While there are some nascent methods for quantifying the value of equity in value-for-money calculations, these rely on a number of assumptions and are more complex than most donor agencies can achieve in practice. Instead of trying to quantify "equity" as part of the value-for-money metric, it is recommended instead to think about reaching marginalized and vulnerable populations as a specific sub-population, which can impact program costs and results. To take a concrete example, let's think about adolescent girls and the barriers they might face in accessing education. On the one hand, some studies find higher impacts among more marginalized sub-populations (e.g., the experience with conditional cash transfers in Burkina Faso in Akresh et al. 2013 and multiple cases cited in Evans and Yuan 2019), while on the other hand, some studies find higher impacts for the subpopulations who were initially more prepared (e.g., Muralidharan and Sundararaman, 2011 and Dulfo et al. 2017).

If there is a trade-off between serving more people who are less vulnerable and serving fewer but more vulnerable people, with a given amount of resources, then how any individual donor agency negotiates that trade-off depends on how much value they place on equity. Are you willing to pay twice as much per outcome to reach vulnerable people? What about five times as much? What about 10 times as much? Cost data analysis will not give you the answer to these value judgments, but it will help you to quantify the trade-off you face in a particular situation.

5.5. Applying Cost data analysis Results to Another Context

Rather than thinking about how well a cost-efficiency/cost-effectiveness estimate will apply in another context, it is easier to break this out into two separate questions:

- 1) How well will the output/outcome results apply in other places?
- 2) How well will the costs of this program apply in other places?

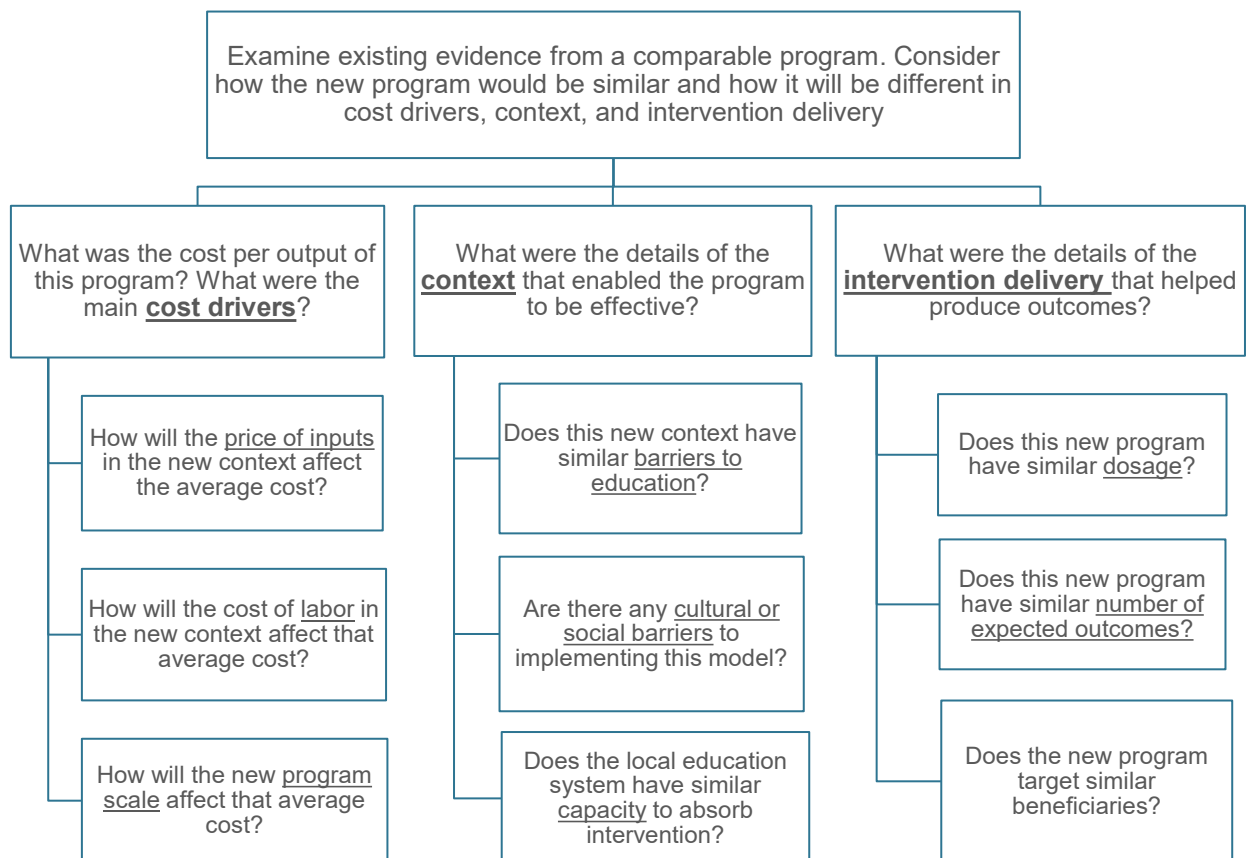
How might outcomes change? The outcomes of any intervention result from an interaction of the intervention with the context in which it is being implemented. In other words, the context of the education system largely determines whether the intervention will "succeed" or "fail." For example, a teacher training program might be very effective in a stable education system where teacher turnover is low and attendance is high, but the same intervention might fail in a system with high teacher absence and turnover. An education system analysis will help determine whether the program's theory of change is applicable to the country context. See Bates and Glennerster, 2017, for a helpful framework²¹ to determine when results might generalize to new

²¹ Mary Ann Bates and Rachel Glennerster, "The Generalizability Puzzle," *Stanford Social Innovation Review* Summer 2017.

contexts and the BE² guidance note “Assessing the Strength of Evidence in the Education Sector” for more details on assessing the external validity of information about program effectiveness.

How might program costs change? It is not really a question of “if” program costs will change from one place to another, but “how” they will change. There are a few key features that seem to drive program costs across programs that will help you understand whether the average cost is likely to be higher or lower in your new setting than in the place the program was originally studied. Figure 8 outlines the main categories of factors that might influence the applicability of data from another context: pricing of labor and goods, program scale, and local context features.

Figure 8. Assessing a program’s viability in another context



Regardless of how well the analysis of program’s viability in another context is performed, the cost-efficiency and cost-effectiveness estimates will be just projections until the program is fully tested in a new context. We recommend using results of cost data analysis alongside other considerations, being cognizant of their potential limitations.

6. RECOMMENDATIONS FOR IMPLEMENTATION

KEY POINTS:

- Principles laid out in this guidance can be adapted for any donor agency.
- Cost learning agenda is a foundation for the development of an organization-specific cost measurement approach.
- The routine and appropriate use of cost data analysis results is central to the success of institutionalizing cost measurement within an organization

Regardless of specific investment objectives, the allocation of resources in donor investments is typically expected to maximize return on investment, be that in fulfilling humanitarian objectives, social and economic development, or environmental protection goals. An improved understanding of costs of donor-supported education interventions, alongside with robust evaluation evidence, will help optimize value for money of international investments in education systems.

Principles of cost measurement presented in this guidance could be adapted for any international donor agency. We recommend an adaptation process be guided by pre-existing organizational priorities that best align with the principles of cost measurement laid out in the guidance. For example, if an agency's main objective in the education sector is improved access to quality education for girls in low-income countries, then cost measurement can be harnessed to estimate how many girls can meaningfully benefit from available resources. If the main objective of a development agency is to support holistic education development from early childhood education through tertiary education, linking cost measurement to achieving these objectives will help garner support for cost measurement among stakeholders. Thus, the first step of implementing this guidance is to link cost measurement to pre-existing organizational priorities. To do so, we recommend articulating a cost learning agenda that will help an international development agency improve value for money of its investments. Bringing together organizational stakeholders to develop concrete cost-related questions will help build internal support for necessary changes. Assigning staff to move implementation of cost measurement within the organization and allocating necessary resources to make organizational change are also a part of this first step.

Cost learning agenda questions will become the foundation for the next step – a development of an organizational approach to cost measurement. The approach would be more likely to succeed if it is built on existing operational systems rather than an entirely new reporting paradigm. The activity-based costing approach presented in this guidance can be adapted for implementation based on the existing organizational structure and priorities. Once an approach to cost measurement is developed, it is important to allow time and resources to build organizational capacity for its implementation. Piloting the new approach in a few countries/programs is useful in working out the details and building internal support. Since internal adoption is likely to take some time, we recommend a purposeful, incremental rollout of the cost measurement approach, allowing learning and reflection to take place throughout the process.

Cost measurement is more likely to gain traction if cost data analysis results are available and useful to field staff as well as headquarters management. Allocating resources for including cost analyses in evaluation studies, as well as in routine production of cost reports throughout project implementation will help stakeholders see the value of cost measurement.

This guidance is produced by Building Evidence in Education working group with the hope that donors will support its implementation in their organizations. Harmonizing cost measurement across the entire global sector will increase its value by allowing for comparisons of cost-efficiency and cost-effectiveness of interventions funded by different donors. Similar to international standards for evaluation studies, adopting clear standards for cost studies would allow us to build and use knowledge base and ultimately improve value for money of international investments in education.

References and Useful Resources

Angrist, Joshua, Eric Bettinger, Erik Bloom, Elizabeth King, and Michael Kremer. "Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment," *American Economic Review* 92, no. 5 (2002): 1535-1558.

Department for International Development. 2011. "DFID's Approach to Value for Money (VfM)." p. 4. Report available on https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/49551/DFID-approach-value-money.pdf.

Dhaliwal, Iqbal, Esther Duflo, Rachel Glennerster, and Caitlin Tulloch. 2013. "Comparative Cost-Effectiveness Analysis to Inform Policy in Developing Countries." In *Education Policy in Developing Countries*, edited by Paul Glewwe. University of Chicago Press.

Duflo, Esther, Abhijit Banerjee. 2017. *Handbook of Field Experiments*. Business and Economics.

Evans, David, and Anna Popova. 2014. "Cost-Effectiveness Measurement in Development: Accounting for Local Costs and Noisy Impacts." Policy Research Working Paper 7027. The World Bank Group.

Evans, David, and Fei Yuan. 2019. "We Can Learn a Lot about Improving Girls' Education from Interventions That Don't Target Girls", Center for Global Development Blog, March 2019. <https://www.cgdev.org/blog/we-can-learn-lot-about-improving-girls-education-interventions-dont-target-girls>.

Fishman, Samuel. 2018. "Excuse me, can anyone tell me the cost of this education program?" (blog), World Bank Blogs, September 25, 2018. <https://blogs.worldbank.org/education/excuse-me-can-anyone-tell-me-cost-education-program>

Glennerster, Rachel, and Kudzai Takavarasha. 2013. "Running Randomized Evaluations: A Practical Guide." Edited by Princeton University Press.

Kremer, Michael, Conner Brannen, and Rachel Glennerster. 2013. "The Challenge of Education and Learning in the Developing World." *Science* 340 (6130): 297-300. Levin, Henry, M. 1995. "Cost-Effectiveness Analysis." In *International Encyclopedia of Economics of Education*, 2: ed, Ed. by Martin Carnoy; Oxford: Pergamon: 381- 386.

Levin, Henry, M., and Patrick J. McEwan. 2001. *Cost-Effectiveness Analysis: Methods and Applications*. Vol. 21. Thousand Oaks, CA: Sage.

Levin, Henry, M., Patrick J. McEwan, Clive R. Belfield, A. Brooks Bowden, Robert D. Shand. 2017. *Economic Evaluation in Education: Cost-Effectiveness and Benefit-Cost Analysis*. SAGE Publications, Inc; Third edition

Manual on sources and methods for the compilation of COFOG statistics - Classification of the Functions of Government (COFOG). 2011. Eurostat. <https://ec.europa.eu/eurostat/en/web/products-manuals-and-guidelines/-/KS-RA-11-013>.

McEwan, Patrick J. 2012. "Cost-Effectiveness Analysis of Education and Health Interventions in Developing Countries." *Journal of Development Effectiveness* 4 (2): 189-213.

———. 2015. "Improving Learning in Primary Schools of Developing Countries: A Meta-Analysis of Randomized Experiments."

<https://journals.sagepub.com/doi/abs/10.3102/0034654314553127>

Karthik Muralidharan and Venkatesh Sundararaman, "Teacher Performance Pay: Experimental Evidence from India," *Journal of Political Economy* 119, no. 1 (February 2011): 39-77.

<https://www.journals.uchicago.edu/doi/abs/10.1086/659655>

Piper, Benjamin, Jepkemei, Evelyn, Kwayumba, Dunston, and Kennedy Kibukho. 2015. "Kenya's ICT policy in practice: The effectiveness of tablets and e-readers in improving student outcomes." *FIRE - Forum for International Research in Education* 2 (1): 3-18.

Tulloch, Caitlin. *Cost data analysis Methodology at the IRC*. Guidance Note 1. International Rescue Committee. Report available at:

<https://www.rescue.org/sites/default/files/document/963/guidancenote1-metholology.pdf>.

USAID Office of Education (2018). *Cost Reporting Guidance for USAID-Funded Education Activities*. Report available at: <https://www.edu-links.org/resources/usaids-cost-reporting-guidance>.

Annex A. Description of Reporting Categories

This Annex offers a detailed description of cost categories and illustrative sub-categories for reporting. Table 1 offers examples of illustrative sub-categories for each of the 10 standard reporting categories:

Table A-1. Standard Expenditure Reporting Categories

#	Standard Category	Illustrative Sub-Categories
1.	General operations, management, and reporting	a. General operations b. General management c. Donor reporting
2.	Assessments and evaluations	a. Learning assessments and evaluations b. Citizen-led assessments c. Research and learning agendas
3.	Higher education/Pre-service teacher training	a. Intervention design (non-recurrent costs) b. Intervention implementation (recurrent costs) c. Research
4.	In-service teacher training	a. Intervention design (non-recurrent costs) b. Intervention implementation (recurrent costs) c. Research
5.	Teaching and learning materials (TLM)	a. New TLM development and testing (non-recurrent costs) b. Production and distribution of TLMs (recurrent costs) c. Research
6.	System strengthening and accountability	a. Policy revision/development b. Capacity building c. Data systems strengthening (monitoring, education management information systems [EMIS]) d. Research
7.	Private sector engagement	a. Public-private partnership development b. Internships/direct placement c. Labor market assessments d. Research
8.	Parents/community engagement	a. Intervention design (non-recurrent costs) b. Intervention implementation (recurrent costs) c. Research
9.	Safe schools and infrastructure	a. Permanent schools/classrooms (construction, furniture, supplies) b. Temporary schools/classrooms (construction, furniture, supplies) c. School policies relating to safety d. Research
10.	Grants, scholarships, and cash transfers to individuals/families	a. Scholarships b. Grants c. Cash transfers
11.	Grants to organizations	a. Grants to government institutions b. Grants to non-government organizations c. Payments for direct service delivery
12.	Other (if applicable)	As appropriate

Category 1. General operations, management, and donor reporting

The “general operations, management, and reporting” category includes labor and other expenditures relating to coordination of human and material resources to achieve overall activity goals; shared program costs such as the cost of the rent for in-country facilities, utilities, communication services, legal fees, software licenses, liability insurance, vehicle pool, security costs, and other expenditures relating to general management and operations; and expenditures related to routine performance monitoring activities that are designed to track progress against performance management targets and report to the donor. Data collection, analysis, and reporting activities designed to inform the general management of the activity should also be included. This category also includes the cost of starting up the activity and closing it down. Sub-categories may include “program management,” “program operations,” and “donor reporting.”

Labor and other costs relating to management of specific components and deliverables (e.g., Technical Director, Component Manager, and technical assistance relating to specific components) should not be included in the general management and operations category. Management of specific technical activities relating to components of the program should be reported in the relevant technical expenditure capture category. Labor and other costs associated with large-scale assessments and evaluations should be reported under Assessments and evaluations (Category 2).

Category 2. Assessments and evaluations

This category should include costs of summative and formative evaluations and large-scale assessments and research activities not linked to specific technical questions. Those data collection, analysis, and reporting activities that are designed to support specific intervention components should be reported in the appropriate cost reporting category.

Routine data collection designed to improve instruction and local accountability (such as classroom formative assessments, school report card development, and implementation) should be reported under the “system strengthening and accountability” category; labor market assessments to inform the development of interventions should be reported under the “private sector engagement” category. Institutional and organizational assessments implemented to better understand system gaps should be captured under the “systems strengthening and accountability” category. Needs assessments should be reported under the main technical category, which it is designed to inform or split across main technical categories of the activity. Sub-categories may include “learning assessments and evaluation,” “citizen-led assessments,” and “research and learning agenda.”

Category 3. Higher education/Pre-service teacher training

This category of expenditure reporting includes all labor, materials, and other resources expended for improving organizational capacity of higher education institutions to train next generation professionals and implement research. This also includes training future teachers or para-professional educators who are currently enrolled in/attending pre-service teacher training institutes. The objective of these trainings must be to improve knowledge and practice of higher education instructors/faculty through strengthening of curricula or instructional approaches in the targeted institutions. “Training” is understood as a transfer of knowledge in content area of instruction, assessments, pedagogical practices, capacity to provide a healing and supportive

environment for students, language competency, use of instructional technology, and pre-service practicum, among other areas.

All tasks related to improving in-service teacher training should be coded in In-service teacher training (Category 4). Parent training should be reported under the “parent/community engagement” category; training of the government officials to support policy development should be reported under the “system strengthening and accountability.” Payments to teacher training institutions to support facility upgrades should be reported under the “Block grants” cost reporting category.

Standard sub-categories may include “intervention design (non-recurrent costs),” “intervention implementation (recurrent costs),” or “research.”

Category 4. In-service teacher training

This category of expenditure reporting includes all labor, materials, and other resources expended for training teaching force, including facilitators, instructors, principals, coaches, mentors, tertiary level instructors/professors, with the objective of improving knowledge and practice of instruction. “Training” is understood as a transfer of knowledge in content areas of instruction, assessments, pedagogical practices, capacity to provide a healing and supportive environment for students, language competency, use of instructional technology, among other areas. Examples of labor and other costs in this cost category include training master trainers in any type of education programming in a cascade model, training facilitators/instructors in alternative education programs/alternative learning programs and workforce development activities, training administrators and officials who oversee/support in-service teaching force, in-service teacher training policy development/revisions, and research into what works in in-service teacher training.

All tasks related to improving pre-service teacher training should be coded in Higher education/Pre-service teacher training (Category 3). Parent training should be reported under the “parent/community engagement” category; training of the government officials to support policy development should be reported under “system/accountability strengthening.” Payments to organizations to support teacher incentives for direct service delivery should be reported under the “Block grants” category. Training teachers in new teacher accountability policies should be reported under “system/accountability strengthening.”

Standard sub-categories may include “intervention design (non-recurrent costs),” “intervention implementation (recurrent costs),” or “research.”

Category 5. Teaching and Learning Materials (TLM)

This category includes expenditures on all labor, materials, and other resources that have the end goal of delivering teaching and learning materials into the hands of the final beneficiaries for whom they are designed, including virtual materials. Examples include development or revision of curricula for reading or other subjects, including for alternative basic education/alternative learning programs; development or revisions of workforce development/entrepreneurship curricula; conflict prevention/peace building or civics education curricula; development, piloting, production and distribution of books, guides, scripted lessons, formative assessments, and other materials that are used to support instruction; and learning in schools or non-formal learning centers. The intended users could be for master trainers, teachers, facilitator, mentors, coaches, students,

learners, parents, local organizations, individuals, or other entities. Also included here are development, piloting, production and distribution of books and other materials designed for use in schools, learning centers, workforce development centers, communities, through mobile libraries, parent-teacher associations, and other means; development, piloting and distribution of instructional software and content; and procurement and distribution of instructional technology (e.g., cell phones with speakers, MP3 players with speakers, tablets, e-readers, radios, solar panels or generators to power instructional technology). Development/revisions of pre-service teacher training curricula should be reported under the Higher education/Pre-service teacher training (Category 3).

Standard sub-categories may include “intervention design (non-recurrent costs),” “intervention implementation (recurrent costs),” or “research.”

Category 6. System strengthening and accountability

This category of cost reporting includes all labor, materials, and other direct costs that have the end goal of strengthened systems and improving accountability of the host country government institutions to provide high-quality education services. Examples include professional and organizational development for ministries, state/provincial and local/district educational agencies, and local NGOs; establish/strengthen national education policies relating to accountability, including frameworks or standards, such as reading proficiency benchmarks or student or instructor standards; and youth employment policy, non-formal education policy, or other education-related policies. This category can also include establishing/strengthening national assessment system; introducing/revising laws, policies, regulations, or guidelines on national, regional, or sub-regional levels that affect the provision of and access to education and the process of education delivery; developing, strengthening, or implementing EMIS; and developing/implementing education system accountability systems, including school report cards.

Standard sub-categories may include “policy reform,” “capacity development,” “data system strengthening (including EMIS),” and “research.”

Category 7. Private Sector Engagement

This category includes expenditure on all labor, materials, and other resources that have the end goal of increasing private sector involvement in support of the activity objectives, including private sector assessments for tailoring an intervention approach to existing private sector conditions. This category is particularly relevant for workforce development activities. Examples include labor market assessments; development of private/public partnerships; internships/placement activities; job fairs; development of an industry-recognized skills certification exam/system; development and support for websites and online platforms for connecting the private sector with education institutions and job seekers, including job match software, portals for connecting various job seekers and industries, websites for disseminating market and job-related materials for entrepreneurs, job seekers, and training institutions; and information systems to support market information dissemination implemented as a part of the workforce development programming.

Development, production, and distribution of curricular materials for entrepreneurship training, teaching, and learning materials for vocational institutes, etc., are to be reported under Teaching and learning materials (Category 5). Activities for institutional accountability strengthening, such

as development of an instructor evaluation system, revision of graduation criteria, should be reported under the System strengthening and accountability (Category 6).

Standard sub-categories may include “labor market assessments,” “public-private partnerships,” “internships and direct placement of young people in employment,” and “research.”

Category 8. Parent and Community Involvement

This category includes expenditures on all labor, materials, and other resources that have the end goal of involving families and communities in support of the activity objectives. For example, in youth development activities, this may include community awareness activities on youth issues or the creation of opportunities for youth to meaningfully participate in community affairs. In activities aiming to increase access to education, this may include back-to-school or on-time enrollment campaigns. In early grade reading activities this would mean increasing parental support for reading, community involvement in school governance, and improvement in the accountability of the local education system to communities and families. Training of volunteers in after-school tutoring and other activities designed to support learning may also be included in this category. Examples include behavior change communication campaigns to contribute to education outcomes (such as promoting on-time enrollment), change of livelihoods practices relating to climate change, and youth-led community activities; training school-based management committees and parent-teacher association members; and training community-based organizations to support learning outside of school; and training community volunteers to support learning outside of school.

Training for citizen-led assessments should be reported under the Assessments and evaluations (Category 2). Development and production of learning materials designed to support learning outside of school should be reported under Teaching and learning materials (Category 5). Cash transfers designed to incentivize families to send their children to school should be reported under the Scholarships, grants, and cash transfers (Category 10).

Standard sub-categories may include “intervention design (non-recurrent costs),” “intervention implementation (recurrent costs),” and “research.”

Category 9. Safe schools and infrastructure

This category includes expenditures on all labor, materials, and other resources that have the end goal of improving safety of schools for all children and teaching staff. Examples include school safety/infrastructure surveys, revisions of school policies relating to safety, updating school infrastructure, and activities designed to prevent school-based and gender-based violence. It is recommended that this reporting category be selected if it constitutes a distinct component of an activity’s scope. In cases in which safety topics are integrated into the general teacher training or policy work, it is not necessary to create separate sub-categories to capture these activities.

Expenditure associated with construction of new schools or refurbishing of existing classrooms/buildings should be reported in this category.

Standard sub-categories may include “permanent schools/classrooms (construction, furniture, and supplies),” “temporary schools/classrooms (construction, furniture, and supplies),” “school policies relating to safety,” and “research.”

Category 10. Grants, scholarships, and cash transfer

This category of expenditures intends to capture the amount and type of cash transfers to individual beneficiaries or families in support of education activity objectives, such as scholarships for attending an educational institution, grants to individuals to start business, or conditional cash transfers to families housing orphans or vulnerable children to support their schooling. This category captures both the actual amount of cash transfers as well as management and logistics costs associated with transferring the cash.

Standard sub-categories may include “scholarships,” “grants to individuals and households,” or “cash transfers.”

Category 11. Block Grants to Organizations

This category includes the amount of cash grants transferred to government and non-government organizations in support of specific purchases or activities that aim to contribute to achieving programmatic objectives. In cases when the implementing partner transfers money for direct education delivery (for instance, to pay teachers in refugee camps), such payments should also be captured in this category.

This category is not to be confused with sub-contracts or sub-awards to local partner organizations in which local organizations are to perform a set of specified activities and are subject to donor reporting requirements. Those will use standard expenditure categories and report following general guidance.

Standard sub-categories may include “block grants to government institutions,” “block grants to non-government institutes,” and “direct service delivery.”

In reporting expenditures in this category, the purpose should be included and, if possible, aligned with categories 2 through 9. For example, if a grant under contract is awarded to a local organization for procuring technology for a community center, in reporting on this grant the purpose should include a reference to Teaching and learning materials (Category 5).

Category 12. Other

Programs funded under multiple funding streams may have tasks in their scope that are different from the common education-related tasks. For example, a program may have a vaccination component, a school feeding component, or a family planning component. In such cases, additional thematic cost reporting categories could be created, or “Category 12. Other” could be used for all other expenditures.

Annex B. Reporting on Government and Non-Government Contributions

“Contribution” refers to inputs that are not accounted for in a donor’s budget, including cost share or funding from other donors.

In addition to routine reporting on the amount and structure of expenditures in each selected cost category, programs should report on **estimates** of contributions by the host government, specifically:

- Government-owned facilities donated for use by activity staff for the duration of the activity.
- Government staff time donated consistently and/or in significant amount, without which the activity component(s) could not take place. Examples may include teacher time outside of normal hours when teachers were trained, coaches or mentors hired by the government to support the activity, or government inspectors conducting learning assessments.
- Materials donated by the host government, such as vehicles, generators, etc.

Note: **If in the absence of the contribution, the activity would have to procure the contributed inputs, then such contributions should be reported.**

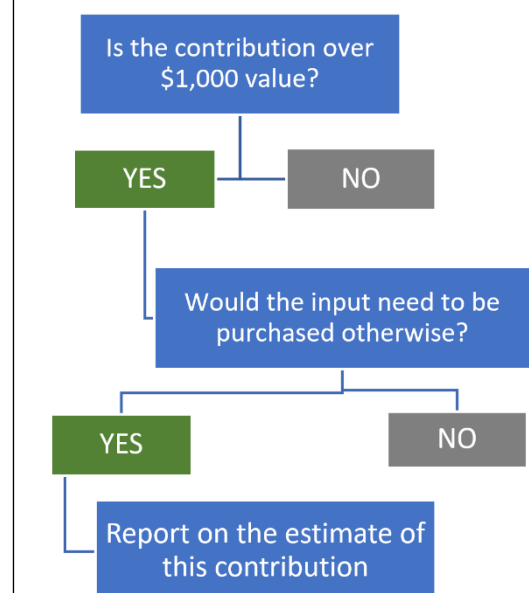
Similarly, it is important to document and report **estimates of non-government and private actors' contributions**, including, but not limited to, individual volunteer time and NGO and corporate donations (monetary or in-kind). Similar to capturing government contributions, only substantial and activity-critical contributions should be captured and reported. Examples include:

- Volunteer time spent to support activities. For example, volunteer community-based tutors working with children after school to help them with homework.
- NGO or corporation-owned facilities donated for use by activity staff for the duration of the activity.
- NGO or corporation staff time donated consistently and/or in significant amount without which the activity component(s) could not take place. For example, NGO or private company staff supporting after-school activities on a pro-bono basis, NGO supporting learning assessments without any cost to the activity, community volunteers engaging with children in reading clubs.
- Materials donated by the NGOs, corporations or individuals, such as books, instructional technology (including computers, tablets and mobile phones), or support for school renovation.

This annex includes detailed examples and a worksheet for tracking government, NGO, corporate, and individual contributions to the activity. Contributions refer to the following types of donations:

1. Labor
2. Time spent in training
3. Office space

Figure A-1. Reporting on Contributions Decision Tree



4. Venues
5. Materials/equipment/supplies
6. Transportation
7. Direct monetary contributions

It is recommended that the donors codify data collection on contributions through templates with clear directions for the partners.

1. Reporting on contribution of time by individuals

It is recommended to collect data on substantial and project-critical contributions of government actors, NGO staff, and volunteers that fall outside of the previously defined responsibilities of the individual. Examples include:

1. The Office of Vocational Training in the Ministry of Labor has assigned two vocational training curriculum specialists to support the revisions of the vocational training curricula. Specialists will work approximately 25 percent of their time on the project, working closely with project staff on the curriculum revisions and the development of teaching and learning materials. Twenty-five percent of two, full-time curriculum specialists' time should be reported for each month they spent working on the project.
2. The Ministry of Education is hiring 500 coaches to support the transition to the new model of teaching literacy designed by the project. The coaches are hired by the MoE in the beginning of the second year of the project implementation and will work full-time on coaching primary grade teachers on the new curriculum, instructional practices, and the use of learning materials developed and produced by the project. The project should report 100 percent full-time employee (FTE) for all 500 coaches for the duration of the project, beginning with the month when they are hired.
3. The project relies on a network of community volunteers to support behavior change communication campaigns. The volunteers are recruited via existing volunteer networks. Each volunteer is expected to spend approximately 3 to 5 hours every week working to promote education each month they spend working on the project. The project is to report 10% FTE for each volunteer (an average of 4 hours per week/40-hour week *100% = 10%), the number of volunteers, and the location where they work for the duration of their engagement.
4. An early grade reading project is working with a local NGO to establish a provision of after-school tutoring for struggling readers. NGO staff will tutor students 3 days a week and spend approximately 2.5 hours for each tutoring session. The project will report 18.75% FTE ($7.5 \text{ hrs}/40 * 100 = 18.75\%$) for each NGO staff who tutors, the number of staff, and the location where tutoring takes place.
5. A private company participates in a pilot of a workforce development project that requires the company's middle managers to spend about 2 hours each week providing feedback on the pilot activities for the duration of the pilot. The data from the feedback will be used to refine the intervention before the full rollout. The project is to report 5% FTE ($2 \text{ hrs}/40 * 100 = 5\%$), the number of managers who provide feedback, and the duration and location of the pilot.

Minor contributions of time of government staff and private individuals (e.g., a high level official dropping in on a project meeting), contributions of time that are not project-critical, or performing tasks previously defined in staff job descriptions should not be reported. Routine activities of government staff that were performed prior to the project but that the project benefits from are not to be reported since they are already included as fixed costs of the education system.

2. Reporting on individual time spent in training

It is recommended to collect data on the amount of government staff, NGO, and private individual time spent in trainings provided by the project. The training may be delivered during scheduled work hours or outside of these hours. Only training of substantial duration (16 hours or more) should be captured. Examples include:

1. The project provides teacher training during summer break when ordinarily teachers would be off work. The project needs to report the number of teachers who participated in the training and the duration of the training.
2. The project assessment specialist provides a 1-day training to MoE staff involved in the assessment work on how to implement EGRA and analyze data followed by two 4-hour coaching sessions.
3. Project WFD specialist conducts a 3-day training of Ministry of Youth staff on using data from labor market assessments to customize implementation in different regions.
4. The project provides training for parent-teacher association members. The project needs to report the number of hours and the number of participants and the location and the purpose of the training.
5. The project provides a 1-day training to university professors on how to use online resources to improve relevance of their courses, followed by two 4-hour coaching sessions. The project needs to report the number of hours and number of participants and the location and purpose of the training.

3. Reporting on Donated Office Space

This category of contributions includes access to office space provided for free by the government, NGO, or a private company for continuous and/or substantial use by the project. Only an approximate square footage should be reported. Examples include:

1. The project is co-located with the Ministry of Education and is allowed to occupy approximately 30 percent of space in the three division offices with which the project actively collaborates, for the duration of the project activities. The project is to report approximate square footage and duration of co-location.
2. The project is co-located with a local NGO and is allowed to occupy approximately 10 percent of the space in its main office for the duration of the project activities. The project is to report approximate square footage and duration of co-location.
3. The project-funded coach is embedded in each of the participating higher education institutes and works from that office. The project is to report approximate square footage for each coach, the number of institutes, and the duration of support.

4. Reporting on Donated Workshop Venue

This category of contributions includes access to a venue provided for free by the government or a private company for an event organized by the project. Projects leaders should report the actual dollar/local currency value of renting equivalent space to enable activities that the donated space supports, if such information is available. If such information is not available, square footage, function, and the duration of use by the project should be reported instead. Examples include:

1. The government is providing space for bi-annual youth job fairs in regional centers.
2. A local youth association is providing space for bi-annual youth job fairs in regional centers.

The project is to report approximate value of the space or equivalent space should it be rented by the project instead of donated by the association.

5. Reporting on Donated Materials/ Equipment/ Supplies

This category of contributions includes all types of materials, equipment, and supplies provided for free by the government or private company for use by the project or to advance project objectives. Examples include:

1. Teaching and learning materials procured or cost-shared by a private company as a result of project activities that support objectives of the project.
2. Equipment for technical and vocational education and training (TVET) institutes procured or cost-shared by the government as a result of project activities.
3. Supplies for education activities within a refugee camp procured or cost-shared by the government as a result of project activities.
4. Mobile technology equipment procured or cost-shared by the government education institution as a result of project activities.

6. Reporting on Donated Transportation

Transportation donated by the government or a private company in support of project activities should be reported in dollar or local currency equivalent of what it would cost the project to fulfill the necessary function that the donated transportation fulfills. For example, if the project uses government trucks to distribute books to schools, the project should report the approximate monetary value of these services. The project should include a description of the purpose of the money transfer. For example, a workforce development project establishes a network of private “angel” investors and pairs them with participants of entrepreneurship training that the project conducts. The amount of direct investments generated through the angel investor network should be reported in this category.

The next section shows an illustrative reporting template (adapted from USAID’s Cost Reporting Guidance).

Table B-1. Worksheet for reporting on contributions

1. Donated time ¹	[# of individuals and their time ²]	[contributor ³]	[location ⁴]	[cost category ⁵]	[brief description of the purpose]
2. Time in training	[# of individuals and their time]	[contributor]	[location ³]	[cost category ⁴]	[brief description of the training]
3. Office Space	[# of individuals ⁶]	[contributor]	[location ³]	[cost category ⁴]	[brief description of the purpose]
4. Venue	[Capacity ⁷ or value estimate ⁸]	[contributor]	[location ³]	[cost category ⁴]	[brief description of the purpose]
5. Materials/Equipment/Supplies	[# of units or value estimate ⁶]	[contributor]	[location ³]	[cost category ⁴]	[brief description of the contribution and its purpose]
6. Transportation	[# of units or value estimate ⁶]	[contributor]	[location ³]	[cost category ⁴]	[brief description of the contribution and its purpose]
7. Direct monetary contributions	[amount ⁹]	[contributor]		[cost category ⁴]	[brief description of the purpose]
8. Other				[cost category ⁴]	[brief description of the purpose]
COMMENTS/NOTES					

¹ Please refer to the “Contribution definitions” section (Table A-1, pg. 45) for an explanation of each category.

² If different people contributed different amounts of time, the total can be averaged or described in detail. E.g., 2 people contributed 30% of their time over the reporting period, 5 people contributed 10%. The time can be reported either in hours over the reporting period, or in percent of full-time employment.

³ Contributor codes: 1= government staff; 2=private individual; 3=non-governmental organization (NGO); 4=private company; 5=other donor.

⁴ Location codes: 1=capital; 2=non-capital urban; 3=rural. For staff contributions in different categories of location, please indicate percentage breakdown or create separate rows. The template does not differentiate between seniority levels of contributions of staff time or staff time in training. The cost estimates will be derived based on the median salary in the private sector corresponding to the tasks performed in each of the three categories of location.

⁵ Cost category: the name of the activity’s cost reporting category associated with the contribution’s objectives.

For contributions in different cost categories, please indicate percentage breakdown or create separate rows.

⁶ Number of staff occupying donated office space should be reported based on the actual number of activity staff

who work the majority of their time at the donated office space.

⁷ Capacity of the venue can be reported in square feet or the number of people.

⁸ Value estimate, if available, can be reported in local currency or in USD.

⁹ Amount of direct monetary contribution can be reported in local currency or in USD.

Annex C. Reporting on Details of Outputs and Dosage

1. Why Report on Details of Outputs and Dosage

Standard Indicator Reporting

Every donor-funded project is typically required to report on a range of standard and/or custom performance indicators. Performance target and result data are collected against these indicators on an annual basis. Standard indicators allow for aggregation and reporting at the portfolio level, which enables sectors and the donor agency as a whole to tell a story about global results.

What Standard Indicator Reporting Misses

Standard indicator reporting allows us the opportunity to report on portfolio-level accomplishments. However, these data fail to give an accurate representation of the dosage of various components for each intervention, nor do they provide the level of detail necessary at the project level to compute unit costs and establish cost-effectiveness of different models or to promote and inform scale up and sustainment of effective intervention models by partner governments.

Capturing data beyond Standard Indicators

To compute unit costs or produce other cost analyses, it is essential that we capture more detailed data on the intervention dosage, outputs and outcomes for each activity, as well as the data on relevant contextual factors that moderate how the intervention is received by the beneficiaries. For example, the amount of curriculum-set instruction time will have an impact on project-level outcomes. In addition to financial data, cost reporting should incorporate data on the details of implementation, including information on beneficiary-level dosage of core intervention activities and products and outputs by cost category. Data on beneficiary-level outcomes also can be included, when available, or can be retrieved from the associated evaluation reports. Table C-1 summarizes intervention details collected as part of cost reporting and how they map out to cost categories and outcome indicators.

2. Output and Dosage Reporting Instructions

This section includes three reporting templates. Each template is designed to collect data on outputs and dosage to enable cost analyses of the activity.

Table C-1. Overview of output and dosage reporting templates

TEMPLATE	EDUCATION LEVELS	REPORTING FREQUENCY
C-1	Early childhood through secondary; accelerated and non-formal designed as a replacement of formal.	Annually at the end of the school year

C-2	Secondary TVETs, post-secondary TVETs, higher education institutions	Annually at the end of the academic year
C-3	Youth interventions not linked to education levels	Rolling basis following reporting frequency specified in the award

Donor-funded projects should fill out a template for each intervention they develop and/or implement. “Intervention” refers to a specific set of activities designed and/or implemented by the donor-funded project to achieve a concrete objective and targets specific beneficiaries. It is not uncommon that a single activity is implementing multiple interventions. For example, an early grade reading activity might be operating in both government schools and in community schools. A youth project might be operating in secondary TVET institutes and in community-based centers. Even though the overarching objective of such interventions might be the same (“to improve learning outcomes in reading for primary grade students” in the first example, and “to improve employment outcomes for youth” in the second example), the interventions themselves could be quite different. Community schools might have fewer resources compared to government schools and therefore require more inputs from the donor-funded project. An intervention in secondary TVETs might focus more on strengthening the organizational capacity of TVETs to teach labor market-relevant skills, while a community-based youth intervention might have a different focus. In such cases, the interventions are distinctly different, and partners should fill out the reporting template for each intervention.

Each reporting template includes an intervention description section designed to provide background information on the intervention’s theory of change and the description of its beneficiaries. The “theory of change” is a description of the goal of the intervention and what is needed to achieve this goal. It also describes how and why a desired change is expected to happen in a particular context. In education programming, the goal is typically a change in learning outcomes in particular grades and content area or areas, or changes in youth development outcomes, including employability or employment outcomes, civic engagement outcomes, resilience to violent extremism outcomes, among others. The goal of the intervention must be clearly specified to allow for appropriate cost-effectiveness analyses. A description of the theory of change must have context and be beneficiary-specific. A description of targeted beneficiaries must accompany the description of the theory of change.

Reporting templates include sections to specify the reporting period. If it’s a school or academic year, dates at the beginning and end of the year should be reported. Whether schools or institutes where the intervention operates are state or non-state should be noted.

Notes for Completing Template C-1

Template C-1 is designed for primary and secondary education system strengthening activities that include government, community and other non-state schools, and non-formal/accelerated learning programs that operate on a standard school year.

Many primary grade interventions target specific subjects (e.g., reading or math). In such cases, the subject should be noted in the relevant section of the template, and the template should be completed for each subject. Only activity-relevant grade levels/grades should be included.

Field “# of minutes in a week of subject-specific instruction in the curriculum” should include the curriculum-intended duration of the instruction on a typical school week, regardless of the actual time on task and without factoring in absenteeism rate of teachers or students. If data on the actual time on task and on the absenteeism rate are available, partners should include references to the reports or summarize in the Notes field.

Field “# of hours of group training (e.g., cascade) received by each teacher” should include the average training time received by each teacher. For example, if the training model includes three seven-hour days of group training, then the field should have “21 hours.” The reporting timeframe is a school year, although the template can be filled out and submitted before the end of the school year (after the training has been completed). If the training occurs in the summer before the school year starts, it should be reported for the school year that it intends to affect. Please specify in the notes or description of intervention if the individual training is face-to-face, blended, or online/remote.

Field “# of hours of individualized training (e.g., coaching) received by each teacher” should include the individualized training time received by each teacher in the intervention, on average. For example, if the training model includes monthly coaching sessions that last 2 hours, and although the school year has 9 months, the coaching did not start until the third month, then the field should have “14 hours.” The reporting timeframe is a school year. The reported hours should reflect the actual hours if those are readily available; there is no requirement to conduct separate monitoring activities to document the actual hours beyond what the activity is already doing. Please specify in the notes or description of intervention if the individual training is face-to-face, blended, or online/remote.

Field “# of teachers trained in the reporting year” should reflect the actual number of teachers trained in the grade/subject in the reporting year. If one teacher was trained twice in different capacities in two separate training events (e.g., one for grade 1 and one for grade 2), such teacher should be reported twice. Therefore, the number of teachers trained in the reporting year might be larger than reported on corresponding standard indicators that capture unique individuals.

Field “# of TLMs received by each learner in the subject/grade in the reporting year” should report the number of TLMs received by each learner in the subject/grade. If different materials were provided, report that in different columns for different materials. If information and communication technology (ICT) plays a central role in the intervention, please report on ICT in a separate column from non-ICT materials.

Field “# of TLMs received by each teacher in the subject/grade in the reporting year” should report the number of TLMs received by each teacher in the subject/grade. If different materials were provided, report that in different columns for different materials. If ICT plays a central role in the intervention, please report on ICT in a separate column from non-ICT materials and provide details.

Field “Total # of TLMs distributed in the reporting year per grade” should report the total number of TLMs received by learners and teachers in the subject/grade. Understanding the unit costs of different TLMs is important to the donor, so partners are encouraged to provide as much information as possible on both the expenditure side and the output side to allow for calculation/documentation of unit costs for different types of materials.

Field “# of learners reached by the intervention in the reporting year” should report an estimated number of learners in each grade/subject reached by the intervention, based on school records or other ways of calculating that the activity is using.

Notes for Completing Template C-2

Template C-2 is designed for secondary and higher education system strengthening activities that include secondary and post-secondary TVET institutes and higher education institutes that operate on a standard academic year.

Field “# of hours of group training (e.g., cascade) received by each instructor/faculty in the reporting year” should include the average training time received by each instructor/faculty. For example, if the training model includes two 8-hour days of group training, then the field should have “16 hours.” The reporting timeframe is an academic year, although the template can be filled out and submitted before the end of the year (after the training has been completed). If the training occurs in the summer before the academic year starts, it should be reported for the year that it intends to affect. Please specify in the notes or description of intervention if the individual training is face-to-face, blended, or online/remote.

Field “# of hours of individualized training (e.g., coaching) received by each instructor/faculty in the reporting year” should include the individualized training time received by each instructor/faculty in the intervention, on average. For example, if the training model includes monthly coaching sessions that last 2 hours, and although the year has 9 months, the coaching did not start until the third month, then the field should have “14 hours.” The reporting timeframe is an academic year. The reported hours should reflect the actual hours if those are readily available; there is no requirement to conduct separate monitoring activities to document the actual hours beyond what the activity is already doing. Please specify in the notes or description of intervention if the individual training is face-to-face, blended, or online/remote.

Field “# of instructor/faculty trained in the reporting year” should reflect the actual number of instructor/faculty trained in the grade/subject in the reporting year. If one instructor/faculty was trained twice in different capacities in two separate training events, such instructor/faculty should be reported twice. Therefore, the number trained teachers in the reporting year might be larger than reported on corresponding standard indicators that capture unique individuals.

Field “# of TLMs received by each student in the reporting year” should report the number of teaching and learning materials received by each learner in the area of training. If different materials were provided, report that in different columns for different materials. If ICT plays a central role in the intervention, please report on ICT in a separate column from non-ICT materials and provide details.

Field “# of TLMs received by each instructor/faculty in the reporting year” should report the number of TLMs received by each instructor/faculty in the area of training. If different materials were provided, report that in different columns for different materials. If ICT plays a central role in the intervention, please report on ICT in a separate column from non-ICT materials.

Field “# of hours in work-based learning per student” should report the number of hours of activity-supported practicum or work-based learning in the area of training, average per student, in one year. There is no requirement to monitor the actual hours. If the institution already provides

work-based learning to students and the activity does not support it in any way, then it should not be reported.

Field “# of students reached by the intervention in the reporting year” should report an estimated number of students reached by the intervention, based on institutions’ records or other ways of calculating that the activity is using.

Notes for Completing Template C-3

Template C-3 is designed for youth development activities that are not aligned with an academic year. This includes stand-alone workforce development activities, entrepreneurship, financial literacy, civics education, countering violent extremism programs, livelihoods programs, basic education programs for out-of-school youth, and hybrid programs that include a range of objectives, including HIV prevention and family planning education.

Field “# of hours received by each learner in group training” should report the dosage of training in a group or classroom setting at the learner level, as designed by the program. If monitoring data on actual hours are readily provided, the actual number of hours received by each learner, on average, can be reported.

Field “# of hours received by each learner in individualized training” should include the individualized training time received by each learner in the intervention, on average. For example, if the 6-month entrepreneurship training model includes monthly mentoring sessions that last 1 hour, then the field should have “6 hours.” The reporting timeframe is the duration of the program that should be noted in the corresponding field. The reported hours should reflect the actual hours if those are readily available; there is no requirement to conduct separate monitoring activities to document the actual hours beyond what the activity is already doing.

Field “# of TLMs received by each learner” should report the number of TLMs received by each learner in the area of training. If different materials were provided, report that in different columns for different materials. If ICT plays a central role in the intervention, please report on ICT in a separate column from non-ICT materials.

Field “# of hours in work-based learning per each learner” should report the number of hours of activity-supported practicum or work-based learning in the area of training, average per learner, in 1 year. There is no requirement to monitor the actual hours.

Field “Internship placement (paid or unpaid)” should code whether the activity provided paid or unpaid internships to learners (Internship placement codes: 1=paid internship placement; 2=unpaid internship placement; 3=other).

Field “Support services received by learners” should report whether or not the activity included support services for learners, such as childcare or transportation. If yes, please specify what support services.

Field “# of learners reached by the intervention in the reporting period” should report an estimated number of youth reached by the intervention, based on activity’s records.

Template C-1. Reporting on dosage and details of school-based interventions in one school year

Brief description of the intervention's¹ theory of change: _____

Brief description of beneficiaries (vulnerable/disadvantaged? selection criteria? grades? geography?) _____

Schools are state _____ or non-state _____ (specify); formal _____ or non-formal _____.

Reporting period is school year; begins on: _____ ends on: _____

Subject²: _____

Intervention is evaluated? Impact evaluation _____/performance evaluation _____. Evaluated outcomes: _____

GRADE LEVEL ³ (select applicable)	# of minutes in a week of subject-specific ⁴ instruction in the curriculum	# of hours of group training received by each teacher	# of hours of individualized training ⁵ (e.g., coaching) received by each teacher	# of teachers trained ⁶ in reporting year	# of TLMs received by each learner in a subject/grade ⁷ in the reporting year	# of TLMs ⁸ received by each teacher in a subject/grade in the reporting year	Total # of TLMs distributed in the reporting year per grade	# of learners reached by intervention in the reporting year
Early childhood								
Pre-primary								
Primary								
Secondary								
Accelerated/ non-formal education								
NOTES ⁹								

¹ By "intervention," we understand a set of project activities reaching a particular type of beneficiaries with specific outcomes in mind. When two concurrent interventions are being implemented, Template C-1 needs to be filled for each intervention. Similarly, if a sub-set of beneficiaries within one intervention receives different inputs, this needs to be reported separately.

² Indicate the subject based on what the project is focusing on. If the project works on multiple subjects, the table should be filled out for each subject.

³ Report only on the grades that are reached by the donor-funded intervention. Indicate which grades are reached. Split level category rows into individual grades if the amount of training or number of materials received by teachers or students is different across grades.

⁴ Report the number of minutes allocated for the subject-specific instruction in the current curriculum, regardless of the actual time spent on the task.

⁵ Specify in the notes or description of intervention if the individual training is face-to-face, blended, or online/remote.

⁶ The reporting period is one school year. Actual number of teachers trained should be reported. If the same teacher was trained more than once for different purposes (e.g., one training for reading, one for math), he or she should be reported more than once.

⁷ Specify which teaching and learning materials (TLMs) were used. Report the number received by each learner in each type of TLMs of the listed. Include details of TLMs (books, readers, workbooks).

⁸ If information and communication technology (ICT) plays a central role in the intervention, please report on ICT in a separate column from non-ICT materials. Include details of ICT in notes.

⁹ In the Notes field, please add information that could help interpret output data in the context of cost data and outcome data. For example, the timing of delivering intervention (beginning of the school year versus middle versus end of the school year), issues with implementation (e.g., disruption caused by crises or conflicts), specifics about the intervention that may influence results, and any other information useful for interpreting provided data.

Template C-2: Reporting on dosage and details of vocational and higher education institution interventions in one academic year

Brief description of the intervention's theory of change and content area: _____

Brief description of beneficiaries (vulnerable/disadvantaged? age range? selection criteria? geography?) _____

Institutions are secondary TVET _____, post-secondary TVET _____, or higher education institution¹ _____

Institutions reached by the intervention are state _____ or private _____.

Intervention is evaluated? Impact evaluation _____/performance evaluation _____. Evaluated outcomes: _____

Reporting period is academic year; begins on: _____ ends on: _____

TRAINING AREAS (select applicable)	# of hours of group training (e.g., cascade) received by each instructor/faculty in the reporting year	# of hours of individualized training (e.g., coaching) received by each instructor/ faculty ² in the reporting year	# of instructor/ faculty trained ³ in the reporting year	# of TLM received by each student ⁴ in the reporting year	# of TLM ⁵ received by each instructor/faculty in the reporting year	# of hours in work-based learning per student	# of students reached by the intervention in the reporting year
Technical content							
Pedagogy							
Soft skills							
Other (specify)							
NOTES							

¹ Higher education institutions may include public or private universities, colleges, community colleges, academically affiliated research institutes, and training institutes, including teacher training institutes. Please specify which type the activity is working with.

² Please specify in the notes or description of intervention if the individual training is face-to-face, blended, or online/remote.

³ The reporting period is one academic year. Actual number of teachers trained should be reported.

⁴ Only TLMs issued to enrollees of TVET/higher education institutions are considered "students." Report on books (text books, readers), not workbooks, procured by donor-funded activity and delivered to students in the reporting year. Report on the number received by each student in each of the training areas.

⁵ If information and communication technology (ICT) plays a central role in the intervention, please report on ICT in a separate column from non-ICT materials. Include details of ICT in notes.

Template C-3: Reporting on dosage and details of youth-focused interventions

Brief description of the intervention's theory of change: _____ .

Brief description of the intervention's beneficiaries (vulnerable/disadvantaged? age range? selection criteria? in-school or out-of-school? geography?) _____ .

Reporting period is an intervention cycle (e.g., cohort). Start date ____ end date ____.

Intervention is evaluated? Impact evaluation _____/performance evaluation _____. Evaluated outcomes: _____

WORKFORCE DEVELOPMENT INTERVENTIONS							
TYPE OF TRAINING (select applicable)	# of hours received by each learner in group training	# of hours received by each learner in individualized training	# of TLMs ^{xxiv} received by each learner	# of hours in work- based learning per learner	Internship placement ^{xxv} (paid or unpaid)	Support services received by learners	# of learners reached by the intervention in the reporting period
WFD/vocational training ^{xxvi}							
WFD/literacy							
WFD/numeracy							
WFD/ICT and digital skills							
WFD/soft skills							
WFD/entrepreneurship/livelihood							
WFD/other (specify)							
NOTES ^{xxvii}							

NON-WORKFORCE DEVELOPMENT INTERVENTIONS

TYPE OF TRAINING (select applicable)	# of hours of group training received by each learner in	# of hours of individual training received by each learner in	# of TLMs ^{xxviii} received by each learner	Support services received by learners	# of learners reached by the intervention in the reporting period
Civics/leadership education					
CVE education					
Health/family planning					
Other training (specify)					
NOTES					

^{xxiv} If ICT plays a central role in the intervention, please report on ICT in a separate column from non-ICT materials.

^{xxv} Internship placement codes: 1=paid internship placement; 2=unpaid internship placement; 3=other (specify)

^{xxvi} May be vocational skills or necessary foundational skills

^{xxvii} In Notes field, please add information that could help interpret output data in the context of cost data and outcome data. For example, issues with implementation (e.g., disruption caused by crises or conflicts), specifics about the intervention that may influence results, and any other information useful for interpreting provided data.

^{xxviii} If information and communication technology (ICT) plays a central role in the intervention, please report on ICT in a separate column from non-ICT materials. Include details of ICT in notes.

