WHITE PAPER SERIES

Centers for Excellence in Teacher Training (CETT) Program

Paper Five: Cost Effectiveness

FINAL REPORT

November 2011

This final report was prepared for the U.S. Agency for International Development (USAID), Bureau for Latin America and the Caribbean, Office of Regional Sustainable Development, Education and Human Resources Team, under the Evaluation and Technical Support to LAC/RSD/EHR Contract AFP-M-00-06-00047-00. It was prepared by the Aguirre Division of JBS International, Inc. Its primary authors are Pablo Javier Zardini and Ignacio Enrique Zardini.

CETT WHITE PAPER SERIES

This document is one in a series of white papers discussing the implementation and outcomes of the Centers for Excellence in Teacher Training (CETT) program. The CETT program was implemented by USAID's Bureau for Latin America and the Caribbean, Office of Regional Sustainable Development, Education and Human Resources Team from 2002–2009. CETT was based on a Presidential Initiative derived from commitments made by the U.S. Government at the Summit of the Americas in 2001 and operated in twenty-one countries in the regions of Central and South America, as well as the Caribbean.

The purpose of this CETT white paper series is to highlight the legacy of the initiative and to provide future program designers with some of the most important lessons learned and best practices developed within the long-term implementation of the CETT program.

The CETT white paper series includes five publications by theme:

Paper One: Regional Nature
This white paper discusses the challenges, successes, and lessons learned implementing a regional model for teacher training. The regional nature of CETT differentiated this program from other, strictly national, teacher professional development efforts undertaken by USAID. Three CETTs in the Caribbean, Central and South America underwent a significant process of compromise and cooperation to arrive at their regional models and this paper documents the initiatives taken.

Paper Two: Testing and Assessment
This white paper discusses the challenges and lessons learned in the process of creating a cross-country testing initiative. The three CETTs carried out testing initiatives to track student performance toward literacy benchmarks, with the goal of showing valid and reliable results. An extremely challenging endeavor, student assessment is further complicated when using tests across countries.

Paper Three: Sustainability
This white paper discusses the lessons learned while anticipating the challenges of sustaining the CETT program after the end of USAID funding. The CETTs worked closely with USAID to prepare for the continuation of the program at the regional, national, and local levels. The paper examines the political, financial, institutional, and social sustainability dimensions of these efforts.

Paper Four: Paradigm Shift
This white paper discusses the systemic change in the behaviors and attitudes of CETT stakeholder groups, including school administrators, teacher trainers, teachers, parents, and students. CETT's teacher training model stressed the inclusion of stakeholders at all levels to promote the importance of reading and writing. Achievement of the program's intended effects depended on the willingness of the institutions and individuals involved to change their behaviors. This paper highlights the lessons learned and best practices in promoting this change.

Paper Five: Cost Effectiveness
This white paper presents lessons learned from a cost effectiveness study linking financial inputs and CETT program outcomes. While the data was insufficient to provide a full cost effectiveness analysis, the paper identifies lessons learned and presents guidelines for future program design and cost effectiveness analysis.

WHITE PAPER SERIES

Centers for Excellence in Teacher Training (CETT) Program

Paper Five: Cost Effectiveness

Prepared by:

Pablo Javier Zardini
Ignacio Enrique Zardini

Edited by:

Mirka Tvaruzkova

The authors' views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

This white paper represents an attempt by USAID to conduct a cost effectiveness analysis of the Centers of Excellence for Teacher Training program. During the course of the study it became clear that due to significant data limitations, it was not possible to conduct such a study. This paper represents documentation of the process we went through to come to this conclusion. It is not intended to be a paper on how to conduct a cost effectiveness analysis.

- USAID Bureau for Latin America and the Caribbean, Office of Regional Sustainable Development

TABLE OF CONTENTS

Acronyms iii

Introduction and Methodology 1
- Purpose 2
- Cost Effectiveness Analysis: The Approach 2
- Research Questions 4
- Methodology 4
- Limitations of the Research Study 5

CETT Cost History 7
- Program Costs by Country 7
- Cost Variations by Component 12

Program Effectiveness: Teacher Performance 15
- Changes in Teacher Performance 15
- Limitations to Measuring Teacher Impact 18

Program Effectiveness: Student Performance 19
- C-CETT 19
- Centro Andino 20
- CETT CA-RD 22
- Limitations to Measuring Student Impact 22

Guide to Calculating Cost Effectiveness 25
- Researching Comparable Programs 25
- Comparing Program Costs 26
- Comparing Program Impacts 27
- Cost Effectiveness Ratios 28
- Sensitivity Analysis 29

Lessons Learned and Recommendations 31
- Key Findings and Lessons Learned 31
- Recommendations 32

Figures

Figure 1 Centro Andino – Teacher Costs (2006 – 2009) 8
Figure 2 Centro Andino – Student Costs (2006 – 2009) 9
Figure 3 CETT CA-RD – Teacher Costs (2006 – 2009) 9
Figure 4 CETT CA-RD – Student Costs (2006 – 2009) 10
Figure 5 Nicaragua – Teacher Costs 11
Figure 6 Nicaragua – CETT Teachers Trained 11
Figure 7 C-CETT Costs (2008 – 2009) 13
Figure 8 Average Costs per Beneficiary - C-CETT (2008 – 2009) 14
Figure 9 Centro Andino Student Performance Tests (2009) 27

Tables

Table 1 Cost Comparison of Alternative Programs (2009) 26
Table 2 Mean Scores of CETT and Comparison Schools (2009) 27

Acronyms

AOTR	Agreement Officer's Technical Representative
CA-RD	Centroamérica – Republica Dominicana
CARICOM	Caribbean Community
C-CETT	Caribbean CETT
CEA	Cost Effectiveness Analysis
CETT	Centers for Excellence in Teacher Training
ENRMA	Escuela Normal Ricardo Morales Avilés, Nicaragua
JBTE	Joint Board of Teacher Education, University of the West Indies
LAC	Latin America and the Caribbean
MOE	Ministry of Education
MOU	Memorandum of Understanding
PIU	Program Implementation Unit
UPCH	Universidad Peruana Cayetano Heredia, Peru
UPN	Universidad Pedagógica Nacional Francisco Morazán, Honduras
UVG	Universidad del Valle, Guatemala
USAID	United States Agency for International Development
UWI	University of the West Indies

Introduction and Methodology

The Centers for Excellence in Teacher Training (CETT) program was a Presidential Initiative to improve the pedagogical skills of teachers in the first, second, and third grades in economically disadvantaged communities of Latin America and the Caribbean (LAC). The hemisphere-wide program—announced in 2001 and implemented by the U.S. Agency for International Development (USAID)—created three regional[1] CETTs that began implementation in 13 countries, referred to in this study as:

1. C-CETT (beginning in the Caribbean countries of Jamaica, St. Vincent and the Grenadines, St. Lucia, Guyana, and Belize);[2]
2. Centro Andino (Ecuador, Peru, and Bolivia in South America); and
3. CETT CA-RD (in the Central American countries of El Salvador, Guatemala, Honduras, and Nicaragua; and in the Dominican Republic).

The Cooperative Agreements for USAID assistance to the CETT program ended in December 2009 after over seven years of technical support. (Two CETTs were issued a no-cost extension until early 2010.) As a result of the program, 35,095 teachers and administrators received training in interactive methods of literacy instruction. The program reached over 799,000 students in 21 countries.

CETT provided in-service training to teachers and administrators located in disadvantaged rural and urban areas that did not benefit from other donor programming. The program promoted the development of skills and adoption of active-learning strategies for teaching reading by aligning existing pedagogical practice with research-based best practices. The program had five core components:

1. **Teacher training** in effective reading methodologies and classroom management techniques
2. **Materials** for teachers to use to improve their reading instruction
3. **Diagnostic tools** to enable teachers to identify and address students' weaknesses and needs
4. **Applied research** to ensure the efficacy of the training, tools, and materials provided
5. **Information and communications technologies (ICTs)** to broaden access to the program

In addition, the CETTs also focused on sustainability efforts to ensure continuance of the program after the end of USAID funding. Within the parameters of these components, each CETT had the flexibility to manage and implement the program based on its regional context and needs. As a result, the CETTs developed with slight differences in each region.

CETT training *content* was related to seven literacy skills: reading comprehension, phonological awareness, phonics, fluency, oral expression, written expression, and vocabulary. Knowledge of these skills provided the foundation for integrated and effective reading instruction.

---

[1] In this study, "regional" refers to one of the three CETT areas: the Caribbean, South America, or Central America and the Dominican Republic. "Hemispheric" refers to all three CETTs as a single unit.

[2] By the end of the program in 2009, many more islands in the Caribbean had adopted CETT. Jamaica, St. Lucia, St. Vincent and the Grenadines, Belize, Guyana, and the Commonwealth of Dominica implemented CETT with USAID funding. After learning of the experiences and results of other countries, the governments of Trinidad and Tobago and Grenada approached C-CETT to join, fully financing their own implementation and purchasing C-CETT's technical support. In 2009, five additional countries signed Memoranda of Understanding (MOUs) to expand CETT implementation to St. Kitts and Nevis, Antigua and Barbuda, Anguilla, Montserrat, and the British Virgin Islands.

The CETT teacher training model introduced innovative *techniques* such as continuous teacher training throughout the school year and follow-up support in the classroom. Teacher trainers visited CETT classrooms where they observed teachers and provided feedback and recommendations. Teacher circles gave teachers the opportunity to share their experiences with peers. Each CETT also emphasized the role of parents and the greater community in embracing a "culture of literacy" to support the importance of reading in the early grades.

The program was implemented in two phases: Phase One (2002–2006) and Phase Two (2006–2009). Phase One launched the initial CETT program design and development. Lead implementing institutions in Jamaica, Honduras, and Peru signed Cooperative Agreements with USAID. Phase Two supported a continuation of the CETTs following USAID's emergent consensus that five years were not sufficient to fully implement the program and achieve the desired results.

Purpose

Because of the considerable limitations encountered to obtaining the cost data needed to complete a full cost effectiveness analysis, this paper is meant primarily to review the lessons learned and the information that will be needed to conduct this type of study in the future.

For the purposes of this paper, the research team focused on two beneficiaries of CETT: teachers and students. Using budget information available at the time of its preparation in June 2010, the study presents a comparative analysis of the changing program costs over six fiscal years of CETT (2003 – 2009), as well as the relative gains to teachers and students. Because the available testing and assessment data mainly referred to the last two years of CETT, the team did not have longitudinal data to measure the program's effect over time and to compare its effects to changes in costs. The paper presents an overview of the supplementary information that would be needed to improve the study in order to build a more robust analysis, as well as the lessons learned and recommendations to inform future program design.

This white paper is part of the CETT white paper series, a compilation of five research papers on key topics related to CETT: regional nature, testing and assessment, sustainability, paradigm shift, and cost effectiveness. Each of the white papers examines the three CETTs through a selection of lenses and analyzes the research findings to bring significant and specific lessons learned with respect to CETT activities into focus. This research gives form to the legacy of the Presidential Initiative, and provides future program designers with some of the most important lessons learned during the long-term implementation of the CETT program.

Cost Effectiveness Analysis: The Approach

Cost effectiveness analysis (CEA) is a decision-making tool used to compare several program alternatives to reach a given goal.[3] By comparing different interventions with similar goals, the method leads one to choose the option with the lowest cost effectiveness ratio. While CEA has been widely used in health interventions, its use in education has been less common.[4] Experts point to the valuable

[3] Boardman, A., Greenberg, D., Vining, A., & Weimer, D. (2006). *Cost-Benefit Analysis: Concepts and Practice.* New York, NY: Prentice Hall.
[4] Levin, H. M. (2005). Waiting for Godot: Cost-Effectiveness Analysis in Education. In R. Light (Ed.), *Evaluation Findings that Surprise*. San Francisco, CA: Jossey-Bass.

information that such analyses can provide, including informed decisions about resources allocated for education interventions, and the effects different levels of resources have on student outcomes.[5]

While there are different approaches to carrying out a CEA, for the purposes of this paper the research team considered the following steps:[6]

1. **Choose programs that are comparable.** The cost effectiveness analysis begins with a description of two or more programs and their relative costs and outcomes. It is important to choose programs that have a similar goal (i.e., the outcomes related to the beneficiary population that is being reached). Further, it is important to compare programs that identify the same, or similar, units of analysis. In this analysis, the research team focused on program outcomes per teacher and student. Many times education programs can be compared to similar programs being implemented by the Ministry of Education (MOE) or other international donors.

2. **Delineate costs of the programs over time.** Budget documents should be used to delineate the costs of each program at regular intervals based on annual or quarterly data. In CETT, annual program data was collected. In some instances, this program cost can be further sub-divided based on the specific use of the funding (e.g., materials or labor). The annual program cost divided by the number of program beneficiaries gives the average annual cost per beneficiary.

3. **Select measurement indicators and catalogue impacts.** Measuring the effectiveness of alternative programs for cost effectiveness analysis is largely the same as measuring effectiveness in a traditional evaluation. Standard statistical instruments are used and it is equally necessary to choose and define the correct measures for quantifying the effectiveness of the intervention. Data was analyzed from CETT monitoring and evaluation systems in each region to measure the program impact among teachers and students.

4. **Calculate the cost effectiveness ratio (C/E).** The cost effectiveness ratio of each program is the relationship of two values: the annual average cost per beneficiary and the measure of the program's effectiveness. The program with the lowest cost effectiveness ratio has the smallest fractional relationship between cost per beneficiary and measure of effectiveness.

5. **Produce recommendations and conduct sensitivity analysis**. The alternatives with the highest effectiveness (or lowest cost effectiveness ratio) have, in general, priority in decision-making. It is also important to carry out a "sensitivity analysis" that outlines the limits of the analysis.

The research team gathered the information that was available about the CETT program and went through the five steps of the CEA. It ultimately concluded that insufficient data was available to draw meaningful conclusions. Nonetheless, throughout this process, the team identified lessons that could be learned from the experience. This report outlines these lessons and the additional data that would be needed to make a full analysis possible. Given the importance of this type of analysis, it should be noted that the CEA presented here is more reflective than formative. It is nevertheless important to examine these results to shed light on the successes and challenges of the CETT intervention and to inform future programs.

---

[5] Hanushek, E. (1997). Assessing the Effects of School Resources on Student Performance: An Update. *Educational Evaluation and Policy Analysis*, *19*(2), 141–164.

[6] For further information about different approaches to cost effectiveness analysis, please see: Levin, H. M. & McEwan, P. (2001). *Cost-Effectiveness Analysis: Methods and Applications.* New York, NY: Sage Publications, Inc.

Research Questions

The research team compiled data on program costs and teacher and student outcomes from all three CETTs, as well as the opinions, experiences, and attitudes of CETT program stakeholders and beneficiaries. Consultants Pablo Javier Zardini and Ignacio Enrique Zardini led the research team. The team collected data for the white paper in order to answer the following research questions, which were outlined in the original Terms of Reference:

1. Was the CETT program cost-effective in achieving its program goals: (a) to improve teacher practices in the classroom, and (b) to improve early grade reading among students affected by the program?
2. Did the CETT program offer a cost-effective teacher professional development model? What components of the CETT teacher training model were effective, and how much did these components cost?
3. Did the CETT cost structure offer benefits to Ministries of Education for their teacher training efforts beyond the results from their existing, less expensive training?
4. What comparisons can be made, if any, between teacher and student outcomes and the implementation of the CETT program?

Once these research questions were drafted, the team collected data and conducted in-depth interviews with key stakeholders. A "lesson learned" in this process was identified by asking the relevant questions needed for the cost effectiveness analysis. The team found that some of these questions could not be answered given the cost and impact data that was available. As a result, the team shifted its perspective; given these research questions, the team noted what data and analysis *would be needed* to reasonably answer the questions posed.

Methodology

The research methodology included three stages. The initial phase involved compiling and reviewing program documentation, including cost documents and contextual reports about the CETTs, their impacts, and comparable programs in CETT member countries. The research team compiled cost data and reviewed program documents to understand the similarities and differences among the three CETTs and their organizational structures.

The second phase involved visits to each of the countries where the CETT lead implementing institution was located: Jamaica, Peru, and Honduras. Prior to each visit, the research team sent requests to CETT staff asking for detailed cost information covering all components of the program. Information was also gathered on the number of teachers trained, the number of students affected by the program, and the results of any teacher performance evaluations and student reading achievement tests. These requests for data were made to the CETT lead implementing institution. In C-CETT, the research team received regional cost figures. In Centro Andino and CETT CA-RD, the research team received country-specific information.

The research team conducted in-depth interviews with key CETT staff and stakeholders who had actively participated in the program. In each country, the sample consisted of individuals from the following stakeholder groups:

- CETT executive directors
- CETT national or regional coordinators

- CETT staff responsible for tracking and managing program finances (e.g., accountants or acquisition specialists)
- CETT testing and monitoring and evaluation staff
- Local ministry of education (MOE) staff[7]
- In the Caribbean, persons managing the virtual platform EDUCOMM[8]

In-person and telephone interviews took place during February, March, and May 2010. The research team collected necessary data and used this information in combination with an analysis of relevant documents such as annual and quarterly program reports. A semi-structured interview protocol, which was used in all interviews, included the main research questions.

The third phase of the study, which was not originally expected, took place when the research team and USAID concluded that a full cost effectiveness analysis was not possible given the cost and impact data available. The research team then shifted the purpose of the study to a lessons learned analysis, and worked with USAID and other experts to identify "lessons learned" from the CETT experience, and what one could learn from this type of study effort. During the third phase of the study, the research team reviewed all of the data again and followed the steps of the CEA to outline lessons learned and recommendations for future programs interested in this type of analysis.

Limitations of the Research Study

As noted, the research team identified several limitations of this study. Some of the limitations were known from the outset, while others arose during the data collection process.

- In many countries, USAID funding for the CETT program had already ended or was coming to an end during data collection. As a result, some staff members had already moved on to other positions or were transitioning out of the program. Because the program had been in place for over seven years when data collection began, it was impossible to conduct interviews or communicate with all of the current and former CETT staff.
- Information related to detailed costs of the CETT program, broken down by component, was a sensitive topic for program implementers. In some cases, respondents did not have authorization to share cost data or were not willing to share it with the research team.
- The research team could not verify the cost data that the CETTs submitted. In order to keep financial data secure, cost information was often tracked by a limited number of people in each CETT. In countries where the CETTs did not submit complete cost data, the research team had no other way of accurately measuring the costs incurred by the program. When triangulation was possible, the research team confirmed financial disbursements with USAID staff, specifically the Agreement Officer's Technical Representatives (AOTRs) who were responsible for USAID funding allocation.
- Each CETT developed its monitoring and evaluation system over the course of the program, and used different indicators to measure program impact (i.e., teacher and student progress). All

[7] The research team contacted Ministry of Education staff to gather information about other education programs, whether MOE-funded or not, related to early grade literacy.
[8] C-CETT developed EDUCOMM, a technological platform enabling virtual connections among countries and clusters participating in CETT. EDUCOMM's goal was to bring evolving technology into the classroom and to become a platform for distance learning and communication throughout the Caribbean. The research team considered the cost effects of developing and maintaining the EDUCOMM platform.

three CETTs tested at least a sample of students to measure changes in reading achievement. C-CETT used pre- and post-tests to track the progress of all students in the program. Centro Andino tested a sample of program and comparison third graders on their reading achievement during the beginning and at the end of the school year. In CETT CA-RD, a sample of students was tested against a comparison group in grades one, two, and three at the beginning and end of the school year. Most of these student assessment efforts were done in the last two years of the program.

CETT Cost History

One of the first steps in cost effectiveness analysis is developing an understanding of program costs over time. This section begins the process by documenting CETT costs in two ways: 1) cost variations over the two phases of CETT implementation by country, and 2) cost variations by CETT component.[9]

As noted in the Methodology section, the following sections of the paper discuss lessons learned that the research team noted while conducting the cost and impact analyses. Though cost data is presented, it is important to acknowledge the challenges of analyzing this data given the gaps in the data available. Nevertheless, the research team and USAID concluded that it would be useful to include this data in order to present the steps that the research team took to gather the necessary information for the CEA. The limitations of drawing significant conclusions from this data are noted.

USAID acknowledged regional and local differences and gave the CETTs, and in some cases, their implementing institutions, latitude to adapt the program to local conditions and needs. Over time, the costs of the CETT program varied both regionally and nationally. Rapid program expansion, from the initial 13 countries to 21 member countries by the end of the program, made it difficult to examine the cost effectiveness of each CETT implementing institution or country. As a result, the research team chose several countries to illustrate CETT program costs over time.

The fiscal periods of the CETT program cycle mirrored the Cooperative Agreements signed in Phase One (2002 – 2006) and Phase Two (2006 – 2009).[10] The research team analyzed cost data for six annual fiscal periods:

1. FY1: October 2003 – September 2004
2. FY2: October 2004 – September 2005
3. FY3: October 2005 – September 2006
4. FY4: October 2006 – September 2007
5. FY5: October 2007 – September 2008
6. FY6: October 2008 – September 2009

Program Costs by Country

The research team calculated annual costs per trained teacher and per student benefitted from the CETT program given the data that was available. In Centro Andino and CETT CA-RD, the team obtained total cost figures for each CETT implementing country. The following sub-sections present an analysis of program costs for Centro Andino and CETT CA-RD, followed by an analysis of overall program costs by CETT component for C-CETT.

Centro Andino

The research team received data on program costs incurred by Centro Andino for fiscal years 2006 – 2009 (FY4 – FY6). The country data included the overall cost (the sum of all CETT components) for each of the three Centro Andino countries: Bolivia, Ecuador, and Peru. The Universidad Perúana

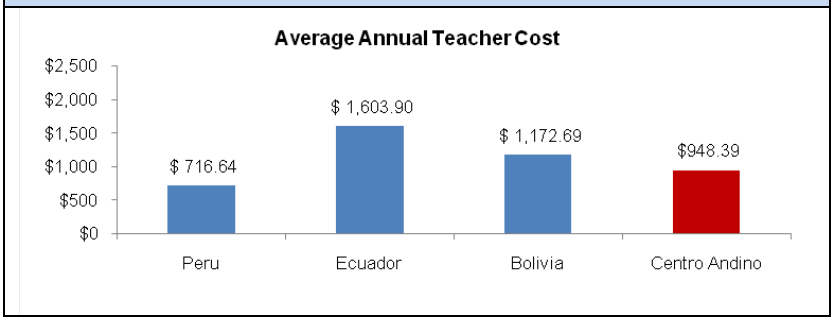
[9] The CETT program included five components: teacher training, materials, diagnostic tools, applied research, and ICTs. Sustainability was an additional component and a major focus of the program in Phase Two.

[10] The first Cooperative Agreements were signed with the CETT lead implementing institutions in 2002. Therefore, one year of program implementation (from October 2002 – September 2003) is not included in the analysis. Though teachers and administrators were trained during this first year, the research team was not able to obtain cost data related to the program.

Cayetano Heredia (UPCH), in Lima, signed the Cooperative Agreement with USAID and in turn signed sub-contracts with the two implementing institutions in Bolivia and Ecuador. USAID and the Program Implementation Unit (PIU), the governing body of Centro Andino located at the UPCH, jointly decided the allocation of program funds. The annual program costs differed by country, with UPCH receiving a larger share of the budget because it housed the PIU and implemented the program over a larger geographic area.

Average annual costs per teacher varied among the three Centro Andino countries, as shown in Figure 1. The total average teacher cost for Centro Andino used aggregated funding for Phase Two from all three countries, divided by the total number of teachers trained, as reported by Centro Andino in its annual reports. This average cost did not adjust for the differences in the national programs. The average cost per teacher in the Centro Andino was $948.39 during Phase Two.

Figure 1 Centro Andino – Teacher Costs (2006 – 2009)[11]



The average annual cost per teacher was lowest in Peru at $716.64 and highest in Ecuador at $1,603.90. Interviews with the UPCH in Peru and with implementing partners in Bolivia indicated that Bolivia and Ecuador needed significant program support and time to develop their teacher training programs. While Peru already had a cadre of teacher trainers who were experts in early grade literacy and training teachers in reading practices, Bolivia and Ecuador had smaller numbers of qualified teacher trainers and master trainers (individuals who specialize in training other trainers). As a result, these programs trained fewer teachers over time and incurred higher annual program costs per teacher.

Since the research team was given aggregate annual costs per country, it was not able to disaggregate program funds allocated specifically for each CETT component, i.e., teacher training, materials, etc. Though the interviews provided insight into possible differences between the country budgets, the team concluded that a more detailed cost breakdown for each country would have provided the necessary information to make reasonable conclusions.

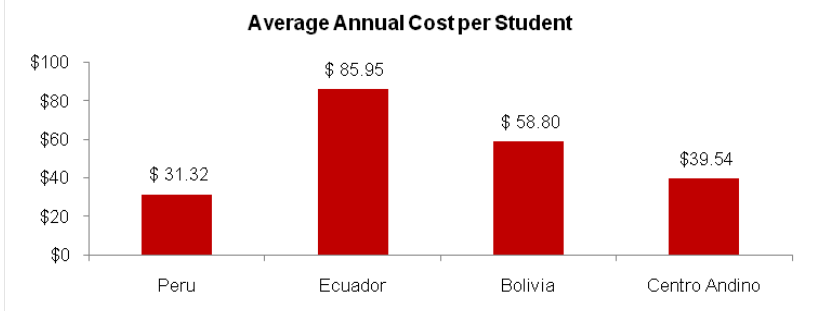
Lesson Learned

It is important to have a constant definition of "program cost" that program implementers are aware of and report regularly. In order to provide a detailed analysis, program cost should be delineated as much as possible by the various purposes of the program. In CETT's case, this would include a cost breakdown for each component: teacher training, materials, ICT, etc.

[11] Average annual program costs presented in this white paper indicate the actual costs of the program in each country during that year, thus presenting nominal rather than real rates. It was beyond the scope of this study to link nominal costs to changes in currency exchange rates or other effects on costs, such as inflation. It should be noted, however, that USAID coordinated each year to disburse CETT budgets based on the projected costs in each country given the country's tasks and objectives, and USAID permitted limited annual increases for inflation.

The average annual cost of Centro Andino per student was $39.54 from 2006 – 2009 (see Figure 2).

Figure 2 Centro Andino – Student Costs (2006 – 2009)

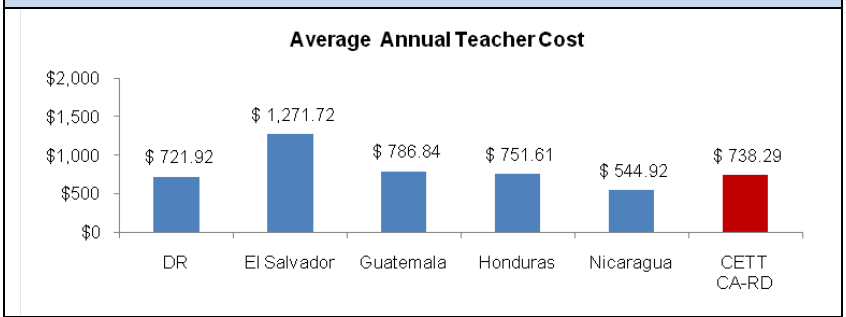


The annual cost per student was highest in Ecuador at $85.95 and lowest in Peru at $31.32. The total average annual cost per student is an aggregate measure that did not account for differences in the size of the national programs. The three countries had similar teacher/student ratios, indicating that class size did not have an effect on differences in average student costs.

CETT CA-RD

The research team obtained budget information for Phase Two of program implementation (2006 – 2009) for the five CETT CA-RD countries: the Dominican Republic, El Salvador, Guatemala, Honduras, and Nicaragua. The Universidad Pedagógica Nacional Francisco Morazán (UPN) in Honduras held the Cooperative Agreement with USAID, and signed sub-contracts with each of the implementing institutions in the other four countries. The budget statements received by the research team included totals for each of the subcontracts.[12] Figure 3 shows the average annual cost per teacher by country, followed by the average teacher cost for CETT CA-RD over the three program years.

Figure 3 CETT CA-RD – Teacher Costs (2006 – 2009)



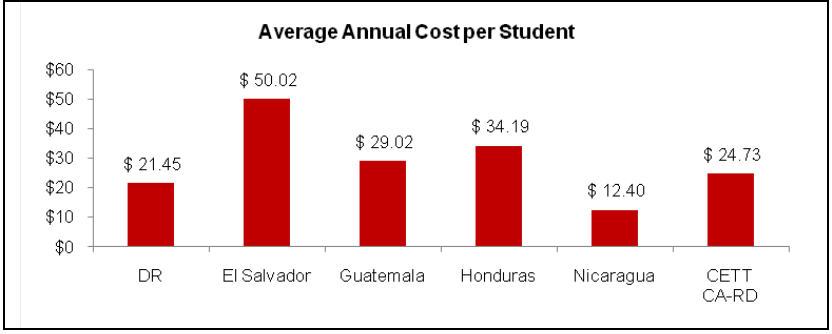
According to Figure 3, El Salvador had the highest average annual teacher cost at $1,271.72, while Nicaragua had the lowest average annual teacher cost at $544.92. However, these calculations are somewhat misleading; upon further analysis, the research team found that the calculated costs per teacher varied due to differences in the regional responsibilities included in each country's budget. In interviews with CETT CA-RD staff, the research team found that budget sums in the implementing institutions compared not only to the number of teachers trained, but also to the CETT components

[12] The total for all of sub-contracts included any and all costs associated with the program, including those not usually associated with the CETT components, such as program design and startup costs.

that each institution led.[13] In the last year of CETT, for example, the Universidad del Valle in Guatemala (UVG) had a larger budget due to the testing initiative it was implementing to collect and analyze student reading achievement tests. As a result, the teacher average annual costs varied mostly because the research team was not able to isolate costs directly associated with teacher training (see lesson learned above). The gross budgets provided lacked the necessary breakouts of regional and program development costs by each country, as well as the costs associated directly with teacher training.

The program in CETT CA-RD reached an average of 14,000 students per country each year from 2006 – 2009. Figure 4 presents the average annual cost per student and the total average annual student cost for the CETT. Average annual cost per student followed a similar pattern as teacher costs in the five CETT CA-RD countries mentioned above.

Figure 4 CETT CA-RD – Student Costs (2006 – 2009)



Nicaragua

An issue in all three CETTs was that the reporting mechanisms for program costs varied throughout the life of the program. The research team noted the importance of tracking costs similarly over time, in order to see a pattern of increasing or decreasing costs over time.

> **Lesson Learned**
>
> In order to analyze program changes in costs over time, it is important to have a set monitoring structure for reporting cost data from the beginning of the program. Changes in reporting costs will necessarily impact the comparisons that can be made across years of program implementation.

In order to compare program costs from the beginning of the program, the research team analyzed data for Nicaragua, the only country where the team had budget totals for Phases One and Two. The implementing institution in Nicaragua was the Escuela Normal Ricardo Morales Avilés with support from the Ministry of Education. Figure 5 presents the annual teacher cost of the CETT program in Nicaragua for the six fiscal years included in this study. The costs decreased over time, from $791.57 per teacher in the first year to $456.50 per teacher in the sixth year. The average annual cost over the six years was $544.92.

[13] In CETT CA-RD, each of the CETT member institutions was responsible for the development of one of the CETT components: (a) Honduras was responsible for teacher training, (b) Guatemala was responsible for research and evaluation, (c) El Salvador was responsible for sustainability and private sector involvement, (d) Nicaragua provided feedback from the field on the implementation of the CETT CA-RD materials and products, and (e) PUCMM in the Dominican Republic was responsible for developing materials. The development of the CETT components is discussed further in the first white paper on regional nature.

Figure 5 Nicaragua – Teacher Costs

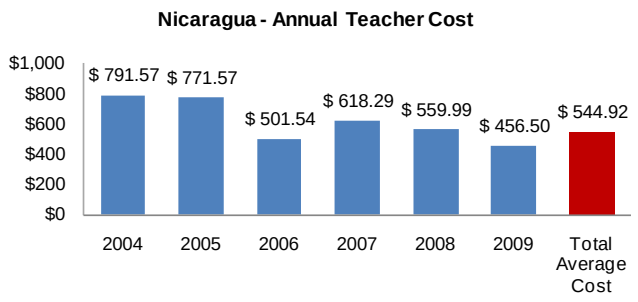
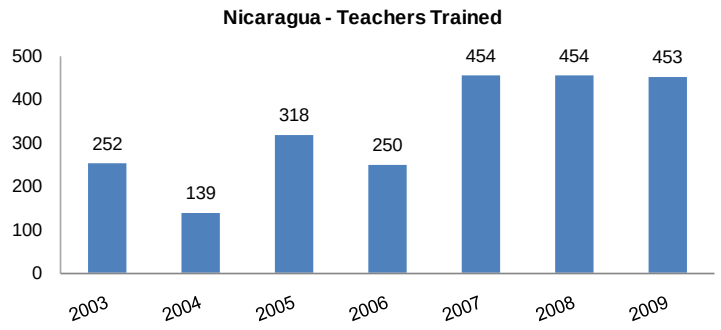


Figure 6 presents the annual number of teachers trained in Nicaragua from 2003 - 2009. There are several reasons for variations in numbers during the first four years (Phase One) of CETT implementation. The program in Nicaragua began in 2003, one year after the rest of the CETT CA-RD countries. As a result, there was an effort to train as many teachers as possible in the first year of the intervention (FY 2003). Further, in Nicaragua, as in other CETT countries, the number of teachers counted as "trained" initially included all teachers who had started training. Subsequently, reporting shifted to count as trained those teachers who were considered to have completed the training only, resulting in a decrease in the reported number of teachers in FY 2004 (because some had already been counted in FY 2003). A teacher was not considered "trained" until he or she had completed several requirements, such as attending several trainings over a period of time (rather than just once).

Figure 6 Nicaragua – CETT Teachers Trained



A final issue with the comparability of Phase One and Phase Two data is that, as in Centro Andino, "teachers trained" in Phase One included both teachers and school administrators involved in CETT training. In Phase Two of the program, a comparable number of teachers received training each year and the annual totals were more consistent. A lesson that emerged from Nicaragua's analysis is the need for a common reporting structure for program outcomes, in this case the number of teachers trained each year. As in Nicaragua, all three CETTs changed their reporting structures, indicating differences in the number of teachers trained.

Lesson Learned

Reporting mechanisms for program outcomes need to be standardized, and measurable definitions are needed for what those outcomes mean. If changes in reporting outcomes are noted, it is necessary to identify whether these changes should be applied throughout the program, and how these changes may affect impact analyses.

As noted above, analyses of Phase One program costs were less robust because data was not available. In order to measure program costs over time, it would be interesting to compare the changes in funds to number of teachers trained, for example, to show how implementing institutions adapted to the program over time. It is possible to hypothesize that average costs per teacher were higher during program start-up, due to extra start-up costs, and lower during subsequent years, due to the higher numbers of teachers being trained. Though the comparisons across countries in Centro Andino and CETT CA-RD were normalized and comparable, the costs per teacher and student may have been higher due to the inclusion of all costs in the calculations.

Cost Variations by Component

The Caribbean CETT (C-CETT) provided the research team with regional cost information for FY 6 (2008 – 2009) per CETT component. C-CETT had a different budget reporting structure than Centro Andino and CETT CA-RD. The differences between this CETT and the other two CETTs made it impossible to compare their costs and program outcomes. A standard reporting structure enforced from the beginning of the program would have made such comparisons possible, another lesson learned noted by the research team.

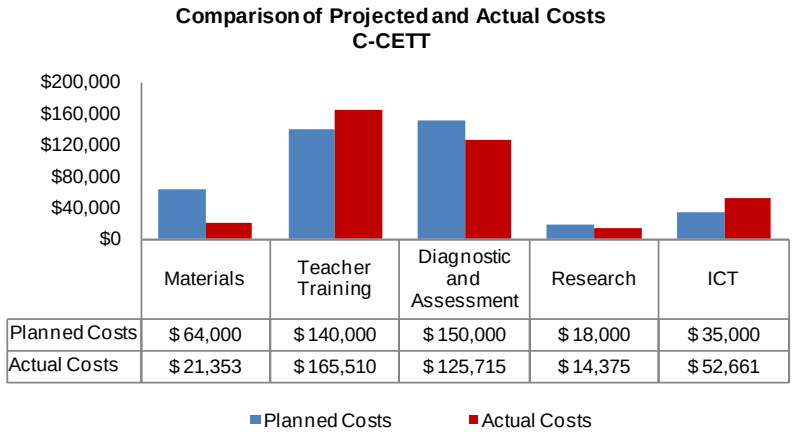
> **Lesson Learned**
>
> Ensure that a standard reporting structure is implemented throughout the program that includes cost breakdowns, as well as aggregate costs. If possible, standard costing worksheets can be part of the reporting requirements for each implementing partner.

C-CETT was implemented by one institution, the Joint Board of Teacher Education (JBTE) at the University of West Indies (UWI) in Jamaica, and the JBTE's budget was compounded for all countries in the Caribbean region (i.e., C-CETT did not provide cost figures by country). The team compared the projected and actual costs of each component, and the relative weight of the component on the overall cost. The five CETT components, as defined by C-CETT, included:

1. **Teacher Training:** In-person courses and/or workshops, professional development, independent study, training follow-up in the classroom, and other related expenses (trainers' fees, per diem allowances, trips, etc.)
2. **Materials:** Textbooks, guides, or materials created for teachers, principals, students, and other beneficiaries
3. **Diagnostics and Assessment:** Student and/or teacher testing design, sampling, data collection, and analysis
4. **Research:** Research and development component
5. **Information and Communication Technology (ICT):** Technology in the classroom, design and implementation, training, and follow-up in the classroom

Figure 7 provides a comparison of the projected and actual costs of the five components in C-CETT for 2008 – 2009.

Figure 7 C-CETT Costs (2008 – 2009)



	Materials	Teacher Training	Diagnostic and Assessment	Research	ICT
Planned Costs	$ 64,000	$ 140,000	$ 150,000	$ 18,000	$ 35,000
Actual Costs	$ 21,353	$ 165,510	$ 125,715	$ 14,375	$ 52,661

According to C-CETT's annual report the most expensive components, teacher training and diagnostics, were those that required the most labor time and resources. The C-CETT teacher training model included in-person training, follow-up support in the classroom, teacher circles, and other interactive techniques to support teacher professional development.[14] The relative cost of this component was also higher because teacher training was a major deliverable of the program, and each of the CETTs reported the number of teachers trained at the end of the year. The second most significant expense was testing and assessment, which became a focus in Phase Two as the CETTs developed testing initiatives to measure program impact of the CETT program among teachers and students.[15] Each CETT hired monitoring and evaluation staff and received technical assistance from testing specialists to create student diagnostics and reading achievement tests. The funding for testing and assessment increased significantly during Phase Two of CETT, supported by technical assistance provided by USAID.

While the actual cost of teacher training and ICT was slightly higher at the end of the year than projections, the actual cost of materials was $40,000 lower than projected. C-CETT's annual report revealed that successful partnerships with Scholastic Books and Pearson Books, and the two companies' donations of supplemental reading materials, were the basis for this saving.[16] In the last year of USAID funding Scholastic Books donated 25,000 books to first, second, and third grade classrooms and Pearson Books donated 10,000 books to first grade classrooms. These cost-sharing measures alleviated the costs needed for purchasing reading materials.

To compare the average overall program cost per teacher and per student, the actual total program cost was divided by the number of teachers and students affected by C-CETT. In all, 955 teachers were trained and 24,009 students benefitted from C-CETT in the last year of USAID funding. The average overall cost per teacher for the year was $397.50 and the average overall cost per student was $15.81. These totals were further broken out by the components that incurred the most costs. Figure 8

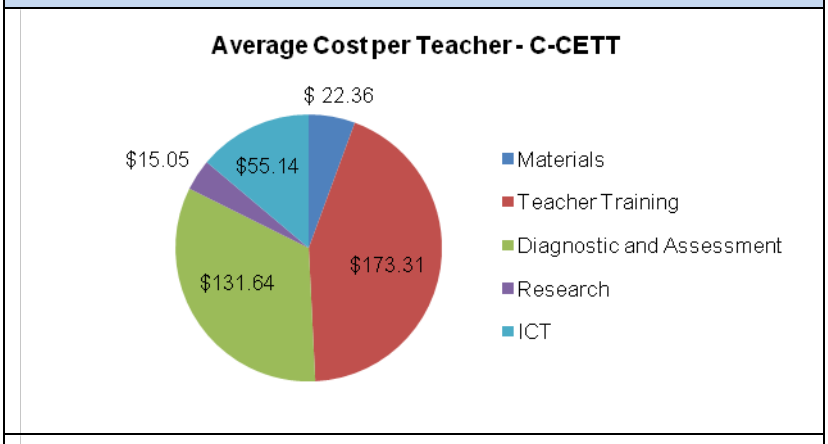
[14] The CETT model created in each of the three regions is described further in the first white paper on regional nature.

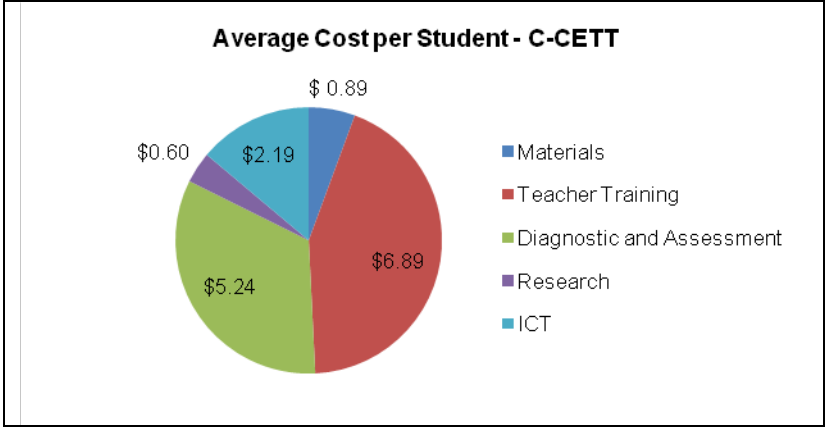
[15] The second white paper on testing and assessment outlines the development of student reading achievement tests in each of the three CETTs and the challenges of creating a cross-regional testing initiative. One of the main lessons learned from this paper is that testing and monitoring and evaluation systems should be in place (and considered in the program budget) from the beginnings of the program.

[16] The benefits of partnerships with book companies and other material or service providers are discussed further in the fourth white paper on sustainability.

provides a comparison of two charts that present average costs per teacher and student by component. As in Figure 7, a majority of the costs were allocated for teacher training and assessment.

Figure 8 Average Costs per Beneficiary - C-CETT (2008 – 2009)





It was important to not only analyze the costs of the component, but to understand which of the components were the most important and valuable to the beneficiaries. This information was obtained and included in the fourth white paper related to the paradigm shift. Since students were not interviewed for the purpose of that study, only the opinions of teachers were available. Most teachers felt that the CETT program was best presented as a "package" and that the strength of the program was in its comprehensive nature.

Program Effectiveness: Teacher Performance

In order to measure program effectiveness the research team analyzed available data related to CETT program impacts and indicators of program success. USAID gave the CETTs technical assistance in setting up testing and assessment systems to measure the impact of CETT program interventions on teachers and students. Each of the CETTs created mechanisms for monitoring program outcomes, such as the number of teachers trained and students affected. This section provides an analysis of changes in teacher effectiveness and the next section focuses on changes in student learning outcomes. The development of testing mechanisms to measure student performance is discussed in further detail in the second white paper of the series.

Changes in Teacher Performance

International literature consistently indicates that teacher knowledge and effective practices in the classroom have a positive effect on the quality of learning.[17] Thus, the goal of CETT was to improve the instructional skills of teachers in order to improve the learning opportunities for children. In an extensive review of best practices in successful teacher professional development programs, key features included: focusing on content knowledge; providing opportunities for active learning in training; fostering coherence of teacher learning and development; providing support through study groups or mentoring and coaching; and sustaining the training.[18] All of these features were incorporated into CETT teacher training, with technical assistance provided from local and international teacher training experts and literacy specialists. USAID provided technical assistance to the CETTs in developing their teacher training based on evidence-based practices. The CETTs also integrated other factors related to teacher quality into the teacher training, such as teacher's dedication, didactics, and pedagogy.[19]

During the beginning phases of the program, the emphasis was on training as many teachers and school administrators as possible. Thus, less emphasis was placed on measuring the program effects on teacher performance or student learning outcomes. As the CETT teacher training models were specified and refined, a need emerged to measure the impact of the intervention. Though efforts were made by CETT to measure these changes, it was not done regularly. Therefore, performance monitoring usually ended with the successful completion of the teacher training.

USAID commissioned several studies and qualitative assessments to analyze and measure the program impact of CETT on teacher performance. Each of these assessments is discussed further in the following sub-sections. In 2004 and 2006, two qualitative assessments were done in which teachers were observed in the classroom and interviewed about their changes in practice. Based on interviews conducted by the research team, Centro Andino also conducted a qualitative assessment of changes in teachers from the beginning of the training to the end of the school year in 2005. In 2008, USAID supported an impact study to test a sample of teachers participating in CETT for the first time. The study was repeated in 2009 to follow the same group of teachers through a second year of training.

---

[17] Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrics, 73*(2), 417-458.

[18] Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What Makes Professional Development Effective? Results from a National Sample of Teachers. *American Educational Research Journal, 38*(4), 915-945.

[19] Hunt, B. C. (2009). *Teacher Effectiveness: A Review of the International Literature and Its Relevance for Improving Education in Latin America*. Washington, DC : PREAL.

Teacher Training Qualitative Assessments (2004, 2006)

The first CETT systematic evaluation of teacher achievement was conducted in 2004.[20] A sample of 114 first-grade teachers (89 CETT and 25 in the comparison group) were observed and interviewed in eight countries: Belize, Guatemala, Honduras, Jamaica, Nicaragua, Bolivia, Ecuador, and Peru. A four-stage rubric of exemplary literacy instruction was used to assess program and comparison teachers. The evaluation team conducted classroom observations and teacher interviews, and recorded teachers' behavior according to the scale described in the text box below.

From the 2004 Qualitative Assessment

In observing teachers for the study, evaluators used a four-stage development continuum.

- *Initiating.* Teachers at the first stage implement reading, assessment, and instructional practices inefficiently or not at all. Reasons for this level of implementation may be lack of knowledge, difficulty in implementation, or resentment of change.
- *Becoming.* At the second stage, many teachers have become conversant with the new jargon and may try to implement some of the new practices. This stage has also been called *Form without Substance* because teachers do not yet have any depth understanding of new practices or new materials learned and are applying both incompletely and superficially.
- *Near Mastery.* With adequate training and coaching, many teachers will move to a third stage at which they understand the advantages of the new methods and have evidence of the efficacy of the practices, since their students are beginning to learn more. Even at this stage, however, many teachers have still not mastered the new practices and may require additional training.
- *Mastery.* At the fourth stage, teachers have fully assimilated the new practices, may be considered "master teachers," and can be excellent mentors or trainers for others.

Some specific outcomes of the 2004 study included the following:

- More than 80 percent of CETT teachers implemented practices to promote oral language development, writing skills, and an understanding of the functions of written text.
- CETT teachers were found to be significantly more skilled than comparison teachers in teaching phonological awareness; in offering opportunities for oral and written expression, vocabulary development, and comprehension; and in employing instructional skills.
- More CETT teachers (64%) than comparison (40%) were able to address student errors correctly.
- While CETT teachers [were] still at a low level of mastery of instruction strategies, their mastery was much higher than that of comparison teachers.

In 2006, evaluators visited some of the same teachers, and new teachers, and found the following:[21]

- The 2006 teachers (in comparison to 2004 teachers) had higher percentages of teachers at near mastery or mastery in every of the 21 performance dimensions in the classroom. These

[20] Chesterfield, R., Culver, K., Hunt, B. C., & Linan-Thompson, S. (2005). *A Reflective Study of Teacher Professional Development in the Latin American and Caribbean Regional Centers of Excellence for Teacher Training (April – November 2004).* Prepared by Aguirre International for USAID. Retrieved from http://pdf.usaid.gov/pdf_docs/PNADF177.pdf

[21] Kulver, C., Hunt, B. C., & Linan-Thompson, S. (2006). *Centers of Excellence for Teacher Training (CETT) Professional Development Review.* Prepared by Aguirre International for USAID. Washington, DC: USAID.

differences were statistically significant ($p<.05$).

- The new 2006 teachers (who had not been observed in 2004) were compared with all 2004 teachers. Comparisons of fluency and comprehension were significant ($p<.05$), and comparisons of questioning, differential instruction, use of resources, and working with others were highly significant ($p<.01$).
- Regarding differences between 2006 "new" CETT teachers and 2004 "old" CETT teachers (who were observed in 2004), the study found that the 2004 teachers had higher mastery in the understanding of written language (60%) and oral language (75%).

The 2004 and 2006 qualitative assessments found that mastery in the literacy and instruction components was significantly higher among CETT teachers in comparison with non-CETT teachers and had continued to improve from 2004 to 2006. The studies found, however, that several components were still lacking, such as differentiated instruction, provision of activities to help children develop fluency, and incorporation of all stages of the writing process.

Qualitative Assessment in Peru

As noted, in 2005 the UPCH in Centro Andino carried out a performance study of teaching literacy skills among a group of teachers who had been trained in 2004, another group of teachers who had been trained in early 2005, and a third group of teachers who never participated in CETT. This study, though indicative of Centro Andino interventions, was only implemented in Peru.

The study found that performance indicators for teachers who had received the CETT training in Centro Andino in 2004 were 18 percent higher than performance indicators for teachers who had not received CETT training. The study also found that the test scores of teachers who had been trained in 2004 were higher than those who had been trained in early 2005, suggesting that CETT teachers improved their performance in the classroom within one year following the training.

From a qualitative viewpoint, observers of teacher performance in the classroom noted that CETT teachers:

- Applied innovative and entertaining strategies in literacy instruction
- Turned the classroom into a stimulating learning environment
- Frequently applied diagnostic and process tests to guide their work
- Involved parents in their children's learning activities
- Promoted reading habits

Impact Study (2008 – 2009)

USAID commissioned a quantitative impact study of CETT teacher performance in 2008 and 2009. [22] The 2008 study included teachers from the Dominican Republic, Honduras, El Salvador, Guatemala, Nicaragua, Bolivia, Ecuador, Peru, and Jamaica. The 2009 follow-up study included teachers in their second year of training from Bolivia, Guatemala, Nicaragua, and El Salvador.

The two-year study found that CETT training had a significant impact on teaching practices in the classroom: Close to 70 percent of CETT teachers employed effective practices in their classrooms. However, the evaluation team also found that the level of teaching practice dropped significantly

[22] Aguirre Division of JBS International, Inc. (2009). *Centers for Excellence in Teacher Training (CETT): Two-year Impact Study Report (2008 – 2009).* Retrieved from http://pdf.usaid.gov/pdf_docs/PDACS248.pdf

between the end of the first year and the beginning of the second year. The training was shown to be most effective at the beginning of the school year, when the concepts were still fresh in the teachers' minds. The focus of training and resulting practices in the classroom were clearer in the first year of training than in the second year.

Knowledge of teaching reading was an indicating factor for the use of effective teaching practices in the classroom as well; however, the findings of the impact study were mixed. In 2008 and 2009, there was a significant correlation between final knowledge at the end of the school year and effective practice observed during the year. However, the impact of CETT on changes in teacher knowledge was unclear. Although 2008 scores on the knowledge survey significantly increased from the beginning to the end of the school year, the teacher knowledge scores remained extremely low. On average, CETT teachers correctly answered only one-third of the questions on the knowledge test. In 2009, knowledge scores from 2008 had decreased by slightly more than one point for teachers in their second year of training. These teachers exhibited lower scores in procedural and declarative knowledge related to teaching reading.

The impact study team concluded that, while improving effective practices in the classroom and especially the practices of rural and female teachers, the CETT program did not show definitive results in terms of increased knowledge in teaching reading. The team recommended that future programs focus more effort on teacher knowledge and provide a clearer structure for teacher training during the second, and subsequent, years of training.

Limitations to Measuring Teacher Impact

Improvements in teacher performance are affected by multiple variables beyond just the training intervention. Rigorous assessments that consider a range of factors can more accurately reflect the actual program impact on teacher effectiveness. It is important to consider any factors that might affect teachers' performance by increasing or decreasing their opportunities to implement what they learn, including external factors such as local capacity and the support of the school administration and community. Although studies of teacher performance in CETT considered factors such as teacher age, gender, prior background, and location, no studies examined the impact of external factors such as local capacity, administration support for CETT practices, or the availability of teacher resources such as in-class coaching.

Another limitation of measuring CETT teacher impact was that monitoring and evaluation related to teacher performance was not developed at the beginning of the program (no baseline was established), and was not continuous (there was no longitudinal data). The only cross-regional impact study of knowledge and effective teacher practices was conducted in the last two years of the program. As a result, the average cost per teacher over the life of the program could not be compared with teacher performance outcomes. A better indicator of change would be continuous monitoring of teachers to document changes in knowledge and performance over time.

Lesson Learned

It is important to launch cost data collection and impact measurement at the same time, so program implementers can measure whether the impact observed was generated by a given level of costs.

Program Effectiveness: Student Performance

The CETT program, as many other teacher training programs, was designed to ultimately improve student learning opportunities, and consequently, to increase student reading outcomes. As noted in the previous section, the link between improving teacher quality and increasing student performance is strong. Studies have found that teacher training in content-based pedagogical knowledge is positively associated with students' achievement. Although many studies focused on student outcomes related to mathematics, more recent U.S. studies found effects of content-based pedagogical knowledge in reading and language achievement. The goal of CETT was to have students in the first grades of primary school significantly improve their literacy skills as a result of being taught by CETT teachers.

In all three CETTs, initiatives for testing and assessment of students were designed primarily in Phase Two of the program, as outlined in the second white paper. Student performance was measured at the beginning and end of the year, by testing samples of students who studied with CETT teachers (treatment) and who studied with non-CETT teachers (comparison). The student testing initiative was a prominent part of the CETT methodology and was reflected in the costs allocated to diagnostic tests and assessment in Phase Two (e.g., see C-CETT costs in 2008–2009). The next section summarizes the results obtained in each CETT based on these student performance tests. It should be noted that the format of the tests was different in each CETT, and thus comparisons between the regions were not possible.

C-CETT

In the Caribbean, C-CETT designed student reading achievement tests that distinguished three levels of reading mastery: students at risk (i.e., reading below grade level), students with average acquisition, and students with an acceptable reading mastery.[23] The reading achievement tests were only prepared for students who completed grades 1–3 with CETT teachers in CETT intervention schools; no comparison group was chosen. C-CETT administered the reading achievement tests in 2006, 2007, and 2008.

The lowest scale for students at risk was 35 percent in 2006 and decreased to 20 percent in 2008. The highest level (acceptable mastery level) increased from 38 percent in 2006 to 55 percent in 2008. The results revealed the program effectiveness for students who studied with the CETT program for three years. These tests were only conducted in CETT intervention schools, however, and the program effect could not reflect control for general improvements in student literacy over time or external factors that may have affected student performance, such as improvements in teaching curriculum or MOE initiatives.

In the 2007–2008 school year, students in Grades 1–3 were given the reading achievement pre-test in November 2007 and the post-test in June 2008. The results showed a general improvement in performance of students. The percentage of students at risk decreased from 27.3 percent to 10.6 percent and the percentage of students at mastery (i.e., reading at or above grade level) increased from 42 percent to 70 percent.

[23] The results of the reading achievement tests were taken from C-CETT's annual report (October 2008 to September 2009). The research team interviewed C-CETT staff and requested testing data, but was not able to obtain it. Thus, any conclusions about student performance increases were reported based on the graphs in the annual report.

Centro Andino

The reading performance of CETT students in Centro Andino was measured from 2005 to 2009 using standard reading and writing tests.[24] Centro Andino's technical testing team designed the student tests with representatives from each of the three implementing institutions. USAID also provided technical assistance with the support of two international testing specialists. One of the greatest challenges in designing these tests was accommodating the differences in reading levels and writing knowledge across the three countries. Centro Andino tried to address this challenge. Though it used various versions of its student performance tests, respondents in interviews noted that the challenge had still not been completely overcome by the end of the program.

2005 - 2007

In the first year of student testing, the reading and writing tests were administered in grades1–3 at the end of the school year. The sample was divided into CETT treatment schools (classrooms with CETT teachers) and comparison schools that did not implement CETT. The outcomes for all grade levels showed that Ecuadorian and Peruvian children generally performed at an adequate level on the test, while Bolivian students performed poorly. However, there were no significant differences between the CETT and comparison groups. In Bolivia and Peru, the treatment group did better than the comparison group, but the differences were not statistically significant.

Centro Andino concluded that more than one year would be required before significant differences between treatment and comparison groups would be evident. To allow sufficient time for program impacts to take place, the 2006 reading and writing tests were administered in third grade classrooms only. The 2006 test was still a post-test, but in this application "treatment" third graders were considered those who had been involved in CETT (or had a CETT teacher) for all three years of their schooling. The 2006 application was also the first time that Centro Andino had used a four-level rubric (based on the Rasch model) to measure student skills acquisition:

- *Optimal:* Demonstrates a complete mastery of reading and writing skills expected for the third grade.
- *Sufficient:* Demonstrates a near-full skill level, but with some weaknesses. Manages to provide answers that demonstrate understanding of criteria.
- *Basic:* Demonstrates a basic mastery of the skills expected for the grade level. Achieves a good level of response to inferential comprehension questions of medium difficulty and compositions written respond to specific (but very basic) tasks.
- *Previous level (the base):* Shows a mastery of skills that are in lower grades and not those of the third grade. Achieves a good level of response to inferential comprehension questions of low difficulty and questions that involve extracting information directly from the texts. Written compositions are inconsistent.

The results of the 2006 reading and writing tests again indicated that third-graders in the treatment groups performed better than those in the comparison groups, though these differences were still not statistically significant. In Phase Two, Centro Andino decided to administer pre-and post-tests to measure differences between treatment and comparison students at the beginning and the end of the school year. Two external evaluations were planned for 2007 and 2009. The initial application in 2007 faced difficulties and was repeated in the 2008 school year.

[24] Data obtained from Centro Andino's "External Evaluation of Results – Consolidated Report," submitted to USAID in January 2010.

2007

A 2007 external evaluation applied a test to treatment (CETT) students and comparison (non-CETT) students at the beginning and end of the school year in Peru, Ecuador, and Bolivia. These tests were only administered in the third grade and 3,375 students were included in the study. Two criteria were used for inclusion in the treatment group:

1. Students had to have been in the CETT program for all three of their primary school years (grades 1–3).
2. Teachers had to have been trained in CETT practices prior to 2007 (teachers "newly" trained in 2007 were not to be included).

Centro Andino used the same levels of skills acquisition as in 2006. In the pre-test, the treatment schools surpassed the comparison schools in Bolivia and Peru; in Ecuador, the performance of the two groups was almost the same. In the same tests run at the end of the school year, the CETT students outperformed the comparison students in all three countries.

More than 90 percent of CETT children in Peru were placed in the two highest levels of competency compared to 81.7 percent in the comparison group. In Bolivia, almost 57 percent of CETT students tested at these levels, compared to 37.9 percent in the comparison group. In Ecuador, CETT students placed higher than the comparison group at the optimum level (37.3% compared to 29.7%). While these differences were positive, not all were statistically significant.

Interviews with Centro Andino and USAID Peru staff indicated that several challenges were encountered while implementing the 2007 tests. The indications, as in previous years, were that the Peruvian students were doing much better than the Bolivian and Ecuadorian students, and that their skill levels were higher at the beginning of the year, indicating less change throughout the school year. While some of this was due to inherent differences between the countries, gaps in data sampling were also noted. In particular, the consideration was that few schools could be found in low socio-economic areas that had not been exposed to *some* aid or intervention (whether from donors or the MOE). Since CETT was implemented only in low socio-economic areas, this presented a challenge of finding a valid counterfactual for the comparison group.

2008 – 2009

In 2008 and 2009, the tests were repeated only in treatment schools in Bolivia and Ecuador; tests were repeated in both treatment and comparison schools in Peru. In both test applications the criteria for inclusion were the same as the previous years (third grade only), and the tests were administered at the beginning and end of the school year. Test results in Bolivia and Ecuador indicate significant changes in test scores in reading and writing from the beginning to the end of the school year ($p<.01$). In 2009, reading comprehension scores among CETT students improved by 55 points in Bolivia and 47 points in Ecuador. Reading scores increased by 32 points among CETT students in Peru, though this change was not statistically significant. Increases were smaller compared to Bolivia and Ecuador, but Peru started with a higher pre-test score than did the other two countries.

On average, CETT students scored higher on both the pre- and post-test than did the comparison groups. CETT students scored 14 points higher on the pre-test and 16 points higher on the post-test. However, the differences in these results between the treatment and comparison groups were not statistically significant.

CETT CA-RD

Regional student tests were first administered in Phase Two in CETT CA-RD. The regional achievement tests were prepared by testing specialists from the UVG in Guatemala.[25] The tests were prepared in two formats for grades 1-3, at the beginning and end of the school year, for CETT (treatment) and non-CETT schools (comparison).

2007 – 2009

In 2007, there were no statistically significant differences in pre- and post-test scores between CETT and comparison schools in El Salvador, Guatemala, and Honduras. The situation was similar in Nicaragua with the exception of a first grade post-test, which indicated a significant improvement between the two groups. In 2008, all five countries in CETT CA-RD had higher post-test reading scores than those in comparison schools, but in no country was the difference statistically significant. Improvements from the beginning to the end of the year also were not statistically significant.

In-depth analysis of the 2008 scores for Guatemala showed the following. At all grade levels, test scores for CETT and comparison schools in Guatemala statistically significantly improved from the beginning to the end of the year ($p<.01$). In first grade, CETT improvements in grade one were statistically significantly higher than improvements in comparison schools ($p<.05$) over the course of the school year. CETT pre-test scores in grade two were statistically significantly higher than comparison school scores on two of the three forms of the reading test ($p<.05$). CETT improvements from the beginning to the end of the school year were statistically significantly *higher* than improvements in comparison schools on one of the forms of the reading test, but not the other two ($p<.01$). Finally, CETT improvements in grade three were statistically significantly higher than improvements in comparison schools on two out of three forms of the reading test ($p<.05$).

In 2009, USAID Guatemala contracted an independent evaluation of the CETT program in Guatemala.[26] Though the data received from USAID was preliminary, it showed significant differences between CETT and non-CETT schools. CETT classrooms were more student-centered (the student initiated teacher-student interaction more), focused more on reading and writing, and included more group work. The results also indicated a greater gender balance in terms of girl/boy initiation of student/teacher interaction in the classroom.

The change in test scores in Guatemala for 2007-2008 suggested that significant score improvements may take a longer time to manifest. Indeed, in interviews with the UVG, the national coordinator indicated that the 2009 test application, which was taking place during the white paper data collection, was indicating preliminary significant results. Unfortunately, due to the late start in applying regional tests, CETT CA-RD did not have tests that followed student performance from the beginning of the program, and this led to truncated results for the second phase only. The CETT staff in Guatemala noted that being unable to show significant progress was one of the unfortunate outcomes of CETT.

Limitations to Measuring Student Impact

While USAID stressed the goal of improving capacity building among the CETTs in measuring student achievement and demonstrating valid and reliable results, some CETT staff and practitioners shared their

[25] The conclusions were obtained from the presentations sent by Ileana Cofiño (UVG) and the information submitted by Roger Rasnake.
[26] Preliminary results were sent to the research team by USAID staff in Guatemala.

concerns about the validity of assessing teachers' performance through student test scores.[27] In fact, during interviews for this white paper some CETT staff voiced their concerns about the causal link between CETT teacher training and student improvements in reading. However, noting the literature that supports the causal relationship between the two, the research team concluded that the assessment of student reading results was a valid measure of program success. Further, these tests were the only constant indicator of measurement, and thus they signified, to a certain extent, the degree to which student knowledge changed (and in some cases, improved) as a result of the program.

A limitation of this study, and any other hemisphere-wide comparison, was that reading achievement tests were designed and conducted by different evaluators in each CETT; thus, the test results among the three regions were not comparable. If tests had been designed in a similar manner, the results could also be scaled based on local contexts.[28] These comparisons would have also provided a more robust measurement of outcomes to program costs in the cost effectiveness analysis.

Lesson Learned

When implementing programs that are compared with other programs to measure relative impact, it is necessary to monitor at least one common indicator across the program interventions. Define these measures at the outset of the program, so all evaluations can incorporate them.

Another limitation of the study was the fact that only some CETTs tested the program intervention against a counterfactual group. In order to measure program impact using inferential analysis, it is necessary to incorporate comparisons of the program intervention to the counterfactual, i.e., a group that was not included in the program. The research team was not able to make these comparisons across the CETTs.

Lesson Learned

When program implementers would like to measure program impact over time, it is important to have a valid counterfactual group. If there is uncertainty about the impact of programs (especially if different programs are evaluated with different methods), then one will not know whether cost effectiveness is being driven by relative costs to impacts, or biased impact estimates.

[27] Hunt, B. C. (2009). *Teacher Effectiveness: A Review of the International Literature and Its Relevance for Improving Education in Latin America.*

[28] Options and challenges for the design, implementation, and analysis of testing data in cross-regional initiatives is discussed in white paper two on testing and assessment.

Guide to Calculating Cost Effectiveness

While overall measurements of teacher and student impact underscored the possible intrinsic value of the program, which was the behavioral change in the stakeholders and the improvements among CETT students in literacy skills, limitations of the cost data prevented the research team from fully analyzing the cost effectiveness ratio of the CETT program in comparison to other program interventions. Learning from this experience, the current section outlines the necessary components of a cost effectiveness analysis. The intent is to inform future program designers of the necessary components in monitoring and comparing program costs and impacts, and in sustaining these measurements over time to understand cost effectiveness throughout the program cycle.

Researching Comparable Programs

The first consideration in calculating cost effectiveness ratios is to find two programs that have a common program goal and that measure program impact using the same units of analysis. In CETT's case, the regional data available was not prone to comparison, since no similar programs that were regional in nature were being implemented. In C-CETT, for example, costs and program outcomes were reported regionally, indicating difficulties in comparing the program to a national intervention. In Centro Andino and CETT CA-RD, the data provided was based on a country-level analysis. After visiting the CETT in Peru, the research team obtained information on the costs of two other development programs aimed at improving teacher training in low socio-economic areas. The research team used these two examples to map out a potential cost effectiveness analysis.

The description of the two programs, a national teacher training program supported by the MOE and the AprenDes program supported by USAID, were both initially considered as programs comparable to CETT. However, upon further review the research team concluded that the scope of the AprenDes program was quite different from CETT. The AprenDes program aimed to improve primary education in multigrade classrooms, which was not a focus of CETT. Further, the AprenDes program was a four-year comprehensive program to empower communities in rural areas only, while Centro Andino in Peru was a two-year teacher training program in peri-urban areas. The national teacher training program could be considered an alternative to the CETT program, since it aimed to improve primary education through teacher training. However, the MOE program is implemented nationwide in selected districts, usually carried out through agreements with universities. The research team was not able to obtain data on how many teachers were trained and in what districts, which led to further complications in making comparisons.

Step one: Choose programs that are comparable

- A cost effectiveness analysis begins with a comparison of two or more programs that share a similar program goal and that measure program outcomes using similar units of analysis.
- The program goal indicates the results desired for the intended population of interest. In an MOE program, the population of interest may be the whole student or teacher population. In a donor program, the population of interest may be specified to the schools or students included in the intervention. The two programs are comparable if they identify the same unit of analysis, e.g., the teacher or the student.

Comparing Program Costs

As noted in the Introduction, the second step of CEA is to outline and compare the costs per beneficiary of each program that is being compared. In the Peruvian CETT, the three programs differed in their program objective and thus a comparison was not possible. Instead, the research team hypothesized a comparison of two programs, one more expensive and one less expensive, to indicate differences in program costs. It should be stressed that these calculations were *strictly fictional*, and did not reflect actual costs of CETT or any other program. Table 1 provides a unit cost breakdown per teacher and student for Centro Andino in Peru (based on calculations done previously) and two alternative programs for FY 2009. Program A is more expensive than CETT, while Program B is less expensive. The assumption in these analyses is that the program objective of Programs A and B are comparable to that of Centro Andino in Peru.

Table 1 Cost Comparison of Alternative Programs (2009)

Teacher Training Program	Unit Cost per Teacher ($ per annum)	Unit Cost per Student ($ per annum)
Centro Andino	$837.90	$30.67
Program A*	$1000.00	$40.00
Program B*	$600.00	$25.00

** Programs A and B are fictitious programs whose average unit costs are presented in order to calculate cost ratios.*

Using Centro Andino as the program of comparison, the research team calculated the cost ratios for 2009:

1) Cost/teacher ratio:
    1. Program A is 16.2 percent more expensive than CETT per teacher trained.
    2. Program B is 28.4 percent less expensive than CETT per teacher trained.

2) Cost/student ratio:
    1. Program A is 30.4 percent more expensive than CETT per benefitting student.
    2. Program B is 18.5 percent less expensive than CETT per benefitting student.

A comparison of Centro Andino and Programs A and B provides a first indication of the cost differences between the programs. While Program A appears to be more costly than Centro Andino, Program B is less expensive per teacher trained and benefitting student. Though these cost ratios are indicative of overall cost, a further analysis of costs per component would be preferred. As noted previously, the funding for Centro Andino included all program costs, including program design, component development, monitoring and evaluation, and in Peru's case, the regional management of Centro Andino. If further information was provided on annual costs for teacher training only, the unit costs for Centro Andino would likely decrease and indicate a more valid value of actual teacher and student costs per training.

Step two: Delineate program costs

- The second step of CEA is to delineate costs of two or more programs over a given time period. Comparable programs should have the same unit beneficiary.
- The average unit costs are calculated over a given period of time (e.g., one year) and for a specific beneficiary (e.g., a student).

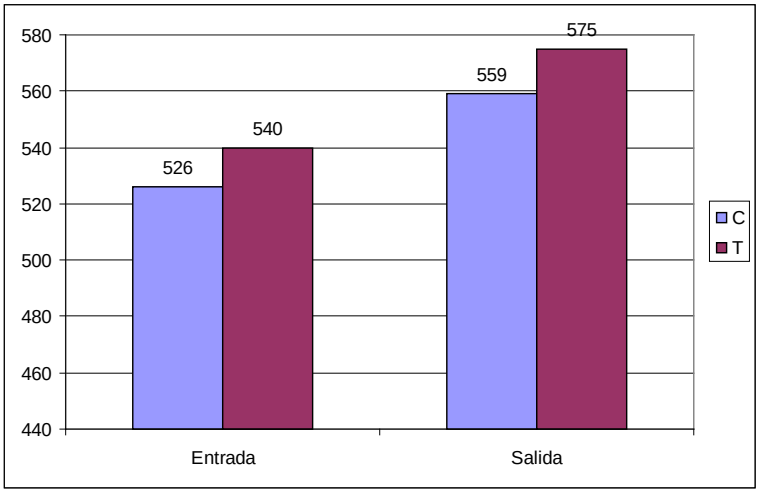
- The program costs of two comparable programs must include the same delineation of funding, whether total program cost or cost per specific program component.

Comparing Program Impacts

Program impacts were monitored and noted by all three CETTs, though comparisons of these program impacts to comparison groups were only available in Centro Andino and CETT CA-RD. In order to compare cost effectiveness ratios, the research team would need comparable performance indicator data for other programs implemented in these countries, which unfortunately was not available. *Nevertheless, an illustrative analysis follows of using comparison data from Peru as performance data for fictitious Programs A and B.*

In 2009, Peru reading achievement tests indicated differences between treatment (CETT) and comparison groups over the course of the school year. Figure 9 presents the test scores. The results shown in this table refer to 2009 comparisons of treatment and comparison groups.

Figure 9 Centro Andino Student Performance Tests (2009)



NOTE: The table describes performance levels for "Entrada" (pre-test) and "Salida" (post-test). The legend refers to "C" comparison group and "T" treatment (CETT) group.

For comparison purposes, the post-test score of the two groups was used to compare impacts. Centro Andino in Peru confirmed the challenge of finding comparison schools that had not had any intervention by the MOE or donor programming, and thus this score may be indicative of comparative impacts of similar programs in Peru. The following table of program results was constructed to indicate mean scores at the beginning and end of the 2009 school year in CETT and Program A or Program B schools.

Table 2 Mean Scores of CETT and Comparison Schools (2009)

Group		Mean	Standard Deviation	Error	Significance (p-value)
Pre-test	CETT	540.48	82.28	3.24	.507
	Comparison (Program A or B)	525.92	79.63	6.93	
Post-test	CETT	575.05	78.12	3.08	.862
	Comparison (Program A or B)	558.80	75.87	6.60	

An analysis of the pre- and post-tests in Peru indicate, as noted, that there were no statistically significant differences between Centro Andino and the comparison group in 2009. While both groups increased their means throughout the school year, the differences in these means were not significantly different.

At this point, the analysis would move to measuring cost effectiveness ratios. If the comparison group was composed of students from Program A, the more expensive program, the cost effectiveness ratio could be calculated. Even if the differences between the two groups are marginal, the differences in cost may indicate that Centro Andino is still more cost-effective. However, if the comparison group was composed of students from Program B, the cost effectiveness ratio would not need to be calculated and the analysis would stop here. Since Centro Andino does not have a significant program impact that is greater than the comparison group, and because Program B was less expensive than Centro Andino, the analysis would conclude that the CETT program is not justified or cost-efficient.

Step three: Catalogue program impacts

- The third step of CEA is to measure the program impacts of two or more programs, and to identify statistically significant differences of the program interventions. Program impact may be measured against a baseline or between two programs directly if they have the same measures of impact.
- If there are significant differences between programs, a cost effectiveness ratio is necessary to calculate cost effectiveness.
- If there are no significant differences between two programs, the cost effectiveness ratio may still be useful to compare program costs and outcomes. However, if a program does not demonstrate impact and is more expensive than other comparable programs, then it is not cost-efficient and one should consider whether it would be feasible.

Cost Effectiveness Ratios

With information available on program costs and outcomes, the fourth step of CEA is to calculate the cost effectiveness ratios. For the purposes of this analysis, the cost effectiveness ratio is a calculation of the program costs divided by the program outputs. In order to continue with the information that was presented for Centro Andino in Peru and fictitious Program A in 2009, the research team was able to identify the costs (C) and effects (E) to calculate the cost effectiveness ratios.

- C_1 = $30.67 (cost of Centro Andino program per student, 2009)
- C_2 = $40.00 (cost of Program A per student, 2009)
- E_1 = 576.05 (post-test score for Centro Andino, 2009)
- E_2 = 558.80 (post-test score for Program A, 2009)

The C/E ratio for Centro Andino in Peru, where [$C/E = C_1/E_1$] is **0.053.**

Interpretation: The average student cost of Centro Andino per each test score increment is $.053.

The C/E ratio for Program A, where [$C/E = C_2/E_2$] is **0.072**.

Interpretation: The average student cost of Program A per each test score increment is $.072.

In this case, and based on the test scores provided in 2009, Centro Andino *is* more cost-effective than Program A, as indicated by a lower cost effectiveness ratio. In this exercise, calculating the cost effectiveness ratio was valuable, because it indicated that the lower costs of Centro Andino make it a more cost effective initiative. The savings in cost offset the smaller, and insignificant, differences between Centro Andino and Program A.

Step four: Calculate the cost effectiveness ratio

- The fourth step of CEA is to calculate the cost effectiveness ratio: C/E. An alternative calculation if two programs are compared to a baseline is the calculation: $C/E = (C_2 – C_1)/ (E_2 – E_1)$.
- Program costs are calculated in the numerator of the cost effectiveness ratio. Thus, the ratio increases with costs if effects stay the same, indicating a less efficient program. The effects are in the denominator of the cost effectiveness ratio. If effects increase given a constant cost, the ratio will decrease and the program will be more cost-effective.

Sensitivity Analysis

Though CEA provides a quantitative measure of program cost effectiveness, it is critical to remember the limitations of such calculations and to specify these limitations in a sensitivity analysis. The measurement of cost effectiveness has a clear utilitarian purpose: to suggest strategies that maximize benefits in relationship to the costs. The interests of beneficiary groups are not considered, nor are long-term program investments, which may be more costly in the short-term. Nor does cost effectiveness analysis change or increase the validity of its inputs. [29] If there is uncertainty about the program costs or how they are measured or about the effectiveness of the program as measured by the tests conducted, there will be uncertainty about the CEA. Since both of these were of great concern in the CETT program, it is important to note that the CEA does not eliminate these limitations.

Step five: Sensitivity analysis

- A sensitivity analysis should outline the limitations of measuring cost effectiveness between programs. It is important and necessary to understand the limitations of the analysis, and to remember that cost effectiveness ratios do not eliminate limitations related to the program data. Data that is not comparable and invalid will likely yield invalid results.

[29] Castaño Yepes, R.A. (1997). Análisis de costo efectividad: una herramienta para la toma de decisiones de política del sector salud. (In English *Cost-Effectiveness Analysis: A decision-making tool in the health sector.) CES Medicina*, *11*(2). Retrieved from http://bdigital.ces.edu.co/ojs/index.php/medicina/article/view/1066

Lessons Learned and Recommendations

This white paper analyzed the average unit costs of the CETTs over time, as well as the program impacts in teacher and student performance. Due to the limitations of the data collected, the research team was not able to complete a full cost effectiveness analysis of the CETT implementing institutions. Nevertheless, the conclusion of the research team was that valuable lessons were learned in the efforts of the CETTs to design and implement a cost-effective program. The CETT program provided a comprehensive package of teacher training that was evidence-based and innovative, including such techniques as continuous professional development and follow-up support in the classroom. Program impact, as indicated by teacher and student changes in performance, further indicate the value of the intervention. This section summarizes the key findings, lessons learned from the analysis, and recommendations for future cost effectiveness analyses of this type of program.

Key Findings and Lessons Learned

The research team documented several key findings in its research concerning the costs and impacts of the CETT program:

- Several factors made it extremely hard to compare program costs and benefits across the regional CETTs. Differences in program size, scope, and the cost data reported made it impossible to calculate cost effectiveness comparisons among the CETT regions and member countries.

> **Lesson Learned**
>
> Program cost data should be monitored and collected annually in a standardized way during the program life cycle. To provide a more precise measure of cost effectiveness, detailed cost data should include both total costs and costs delineated by program goals. Program managers of cross-national or regional programs should attempt to collect identical cost data from each country in order to make valid comparisons.

- Though the qualitative evaluations in 2004 and 2006 gave some indication of the impact of the program on teacher effectiveness, some of these indicators should have been monitored on an annual basis. A continuous teacher performance monitoring system would have made it possible to study teacher impact over the program cycle. The 2008–2009 impact study finding that teacher knowledge was particularly low suggests the value of a greater emphasis on testing teachers to more quickly reveal other important challenges.

> **Lesson Learned**
>
> Teacher training programs should continuously gather information on teacher performance appropriate to the needs of each intervention country and region. This kind of data continuity allows future cost effectiveness specialists to measure differences in teacher behavior over time and to compare these differences with changes in average teacher costs.

- In Centro Andino and C-CETT, continuous student assessments were used to monitor the success of the program over time. In CETT CA-RD, however, student assessment data was only available for the last three years of the program. This prevented both CETT and the research team from mapping program successes and linking them to costs.

Lesson Learned

Programs that measure student achievement should ensure that their tests are (a) designed with the consent of all the parties involved, (b) contextualized to different interventions, (c) standardized across various beneficiary populations, and (d) administered from the beginning of the program.

Recommendations

The key components of a cost effectiveness analysis comprise the steps outlined at the beginning of this white paper. Programs similar to CETT that wish to employ this type of analysis in the future will want to consider the following recommendations:

1. **Choose several alternatives to assess.** The alternatives may include: (i) examining a similar program in schools in grades 1-3; (ii) comparing teacher training programs such as those funded by other donors, similar to the subject program, to be applied to the comparable population and with similar objectives; (iii) continuing with traditional teacher training instead of implementing a special program.

2. **Calculate the costs of alternative programs using appropriate methods.** In this case, having access to data on the costs of national teacher training plans and other alternative plans is essential. The average "unit" cost calculation is straightforward: dividing the total annual cost of each alternative program by the number of teachers involved yields an average annual cost per teacher. The average annual cost per student can also be a relevant point of comparison.

3. **Collect beneficiary performance data to check the impact of the program.** Cost effectiveness can only be determined when program cost data is analyzed in the context of program outcome data. When outcomes create the context for determining effectiveness, it is vital to measure the skill level of teachers at the beginning and end of training in order to gauge the increase in skill level that is the intended outcome of the training. While many factors can influence changes in the learning and skills exhibited by students (the performance of the ultimate beneficiaries, in this case), students can also be tested to assess program impact.

4. **Measure the cost effectiveness ratio.** The cost effectiveness ratio can be calculated using the values resulting from steps two and three. The lowest cost for the increase in score demonstrated by the teacher (or student) is, in principle, the best model for maximizing the use of program resources.

Effective program assessment necessarily involves determining outcomes relevant to beneficiary populations and considering a range of policy options. In assessing programs as a whole, a well-conceived cost effectiveness analysis is a useful decision-making tool because it offers indicative information that objectifies reality. While indicative analysis is a highly useful component of program assessment, the task remains for policy makers to decide how and where resources will ultimately be allocated.