



## 1. Project Data

<b>Project ID</b> P123151	<b>Project Name</b> PE Basic Education	
<b>Country</b> Peru	<b>Practice Area(Lead)</b> Education	
<b>L/C/TF Number(s)</b> IBRD-82260	<b>Closing Date (Original)</b> 31-Mar-2018	<b>Total Project Cost (USD)</b> 25,000,000.00
<b>Bank Approval Date</b> 17-Jan-2013	<b>Closing Date (Actual)</b> 31-Mar-2018	
	<b>IBRD/IDA (USD)</b>	<b>Grants (USD)</b>
Original Commitment	25,000,000.00	0.00
Revised Commitment	25,000,000.00	0.00
Actual	25,000,000.00	0.00

<b>Prepared by</b> Susan Ann Caceres	<b>Reviewed by</b> Robert Mark Lacey	<b>ICR Review Coordinator</b> Joy Behrens	<b>Group</b> IEGHC (Unit 2)
---	---	--	--------------------------------

## 2. Project Objectives and Components

### a. Objectives

According to the Loan Agreement (p. 5) and the Project Appraisal Document (PAD, p. 5), the objectives of the project were "to improve Ministry of Education (MINEDU)'s capacity to evaluate student learning, instructional practice, and school leadership in basic education".

The project was restructured on September 17, 2015 and June 7, 2017 (level 2 on both occasions), and key associated outcome targets were revised downwards. At the time of the first restructuring, nothing had disbursed and 70% had disbursed by the second restructuring.



The September 2015 restructuring revised an outcome target for an indicator (Number of school management positions staffed through the school principal evaluation management system) and added two additional indicators (instructional practices analyzed, cycles of student assessment analyzed). Additional intermediate outcome indicators were included (teacher performance evaluation system designed, implemented, and results reported; evaluation instruments to determine entrance to and promotions within teaching career designed). Three indicators revised the scale from yes/no to a number. These changes do not affect ratings before and after restructuring, and therefore a split evaluation is not undertaken.

**b. Were the project objectives/key associated outcome targets revised during implementation?**

Yes

**Did the Board approve the revised objectives/key associated outcome targets?**

No

**c. Will a split evaluation be undertaken?**

No

**d. Components**

The operation was a sector-wide approach (SWAP) that co-financed selected activities within eligible budget activities. The Government's budgetary framework was linked to the financing of areas expected to improve student learning. Disbursements were based on achievement of agreed upon targets from specified indicators. There were three components:

*Evaluating Student Learning* (appraisal US\$25.25 million; actual, US\$10.90 million) covered eligible budget activities related to: scale-up of second grade student assessment and additional grades in primary and secondary school; measurement of child development outcomes in preschool; and international assessment overseen by OECD and UNESCO.

*Evaluating Instructional Practice and School Leadership* (appraisal, US\$ 4.25 million; actual, US\$13.87 million) covered eligible budget activities related to the Ministry of Education's monitoring and evaluation of instructional practices at the classroom level and implementation of a competency-based system for selecting and training candidates to school management.

*Strengthening's the Ministry of Education's Implementation Capacity* (appraisal, US\$ 0.51 million; actual, US\$0.23 million) covered eligible budget activities related to the Ministry of Education's education management and monitoring capacity of the activities supported by the program and carrying out independent technical, financial, and procurement reviews.



Changes were made through project restructuring to better align with the Government's priorities such as: participation in an international assessment - International Civic and Citizenship Study (ICSS) - and implementation of the Teacher Reform Law (adopted in 2012), which revised teacher pay, promotion, and performance evaluation. Thus, the operation supported the design of the teacher performance evaluation system to bring merit into entrance and promotion in the teaching profession.

#### **e. Comments on Project Cost, Financing, Borrower Contribution, and Dates**

Project Costs: Project costs were US\$ 25.28 million, 16 percent less than the appraisal estimate of US\$30,01 million, . Substantially more was spent for Component 2 (evaluation of instructional practice and school leadership) than planned at appraisal (US\$ 13. 9 million compared with US\$ 4.3 million), due to the activities added during restructuring to implement the teacher performance evaluation system. Approximately half of the resources planned at appraisal were actually spent on Component one (Evaluation of Student Learning) and Component three (Strengthening Ministry of Education's Implementation Capacity). Disbursements were based on submission of actual budget execution of eligible budget activities and evidence of achievement of the disbursement linked indicators (DLIs). The DLIs were based on the proportion of project costs allocated to each component.

Financing: The IBRD loan of US\$ 25 million was fully disbursed. There were no other sources of external financing.

Borrower Contribution: The Borrower contributed US\$ 161 million, which was slightly higher than planned (US\$ 158 million).

Dates: The operation was approved by the Board on January 17, 2013 and became effective on May 20, 2014. The delay in effectiveness was due to the time taken to approve the Operational Manual (ICR p. 23), but this delay did not negatively impact implementation. The operation closed on schedule on March 31, 2018. In September 2015, there was a level two restructuring which made changes to the results framework, implementation schedule, component cost, institutional arrangements, disbursement arrangements, and legal covenants. At the Government's request, the scope of the project expanded in this restructuring to align with its priorities (ICR p. 5). A second level two restructuring was carried out in June 2017, which made some further changes in the results framework and disbursement and institutional arrangements.

### **3. Relevance of Objectives**

#### **Rationale**



The objectives are well aligned with country conditions, with the Government's vision and focus (currently implementing the third phase of its Education Strategy), and with the Bank's priorities throughout the life of the operation.

The Government of Peru has made efforts to promote more inclusive and equitable human capital accumulation to ensure broader poverty reduction and social and human well-being. Creating good quality basic education is critical (as the education system had already attained universal coverage in primary and high levels of pre-primary [80%]). Improving education quality is necessary for skill formation and productivity, particularly for marginalized groups. For example, PISA results in 2011 revealed high numbers of student completing primary level without mastery of basic mathematics and literacy skills. Skill deficits were even higher among marginalized groups such as rural or indigenous students. Gender differences were negligible.

The Government initiated a reform process known as the Learning Outcomes Education Program (Programa Educativo Logros de Aprendizaje, PELA). The first phase of this program (2009-2011) paid insufficient attention to rural students and bilingual education. The subsequent phase of the program (2013-2016) was refined to address these issues and focused on three aspects: improving school conditions and management; strengthening teaching and curriculum (including timely access to quality learning materials); and establishing a robust national student assessment system (to support the program). Prior to this operation, Peru's assessment system included: testing in grade two in mathematics and participation in PISA (every three years). The Ministry of Education needed information on student learning at the end of preschool and across primary and secondary grades to detect issues, inform teacher development programs, and provide feedback to schools. Simultaneously, the Ministry of Education passed a Teacher Reform Law that sought to create a performance-based career to raise the standards for entry into the profession, regulate promotions based on competencies, and implement a teacher evaluation system. Another aspect of the teacher reform was ensuring effective school leadership through competency-based selection process and specialized training. Thus, the objectives of strengthening the Ministry's capacity to evaluate student learning, instructional practices, and school leadership were highly relevant to the Government's overall strategy. Enhancing the Ministry's evaluation capacity underpinned these three aspects.

The objectives also aligned with the World Bank Group's Country Partnership Framework (CPF) for the fiscal years 2017-2021 that focuses on human capital in the workforce. The CPF describes the need for education quality improvements as part of a long-term process to improve the skills of workers and hence raise productivity and growth. The Country Partnership Strategy (FY12-16) at preparation focused on increasing access and improving the quality of services, particularly for the poor. It noted the need for the education system to become more results-oriented.

**Rating**  
High



## 4. Achievement of Objectives (Efficacy)

### Objective 1 Objective

Improve Ministry of Education's capacity to evaluate student learning in basic education.

### Rationale

The theory of change emphasized assessment of student learning as a tool to improve the education system and its quality. The Ministry (through UMC) previously conducted sample-based assessments of student learning. Expanding assessment- grades and subjects was needed to illustrate student performance for policy makers, teachers, principals, and the general public. The theory of change also emphasized international comparison as a means to motivate the actors in the education system by benchmarking its performance against other countries. The project reported student results at the national, regional, and school level. The assumption was that the availability of reliable, timely information on student learning outcomes, trends, and patterns across the school cycle would then be used to inform policy. The dissemination of the data would build accountability to improve education quality. However, as noted in the ICR (p. 46), assessments need to be coupled with incentives and/or sanctions to motivate teachers to teach better and schools to be more effective. Some, but not all, of these aspects were addressed in the other objectives. In IEG's learning engagement with the Education GP on student assessment, it was found that assessments do not automatically translate into utilization of the data by teachers, school leaders, parents, or Ministry personnel, as additional actions beyond dissemination are needed; these were missing from the project's theory of change. Even with a narrow objective, the theory of change could have considered steps that would be needed to translate the evaluation of student learning into utilization.

Key Performance Indicators (KPIs) included: number of additional grades at primary and secondary education in which sound national assessments are administered in at least two subject areas; annual results of national student assessment at the primary and secondary level are reported at national, and regional and school level. The objective was also measured through the following additional indicators: yearly student assessments for selected grades of primary and secondary education validated in a pilot study; number of surveys carried out to evaluate child development progress and quality of public preschool education; and participation in international assessments of student learning outcomes.

Four cycles of assessment were conducted at preschool, primary, and secondary grades (meeting the target of 4). The assessment in the new grades at primary and secondary levels were validated during a pilot. Three additional grades (4th, 6th, and 8th) - beyond the original grade (2nd) were developed for two subject areas. Results for the assessment were reported at the regional and school level (for 6th grade in 2015, for 8th grade in 2016, and 4th grade in 2017) and shared with parents. The Ministry disseminated assessment findings to the media, policy makers, education officials, schools, and parents (ICR p. 22). Census-based assessments were carried out in 2016 (and are planned for grades 4 and 8 in 2018).



The assessment data were analyzed, but the ICR does not discuss actions taken by the Ministry in response. The Bank team reported that the overall assessment trend and public school trend have improved between 2007 to 2016. Student learning assessment data were disaggregated (by gender, region, and urban/rural), but the data were not presented in the ICR, nor was there a discussion of how the data disaggregated by gender, spatial/geographical were used or what actions were proposed in response.

Two national surveys were carried out in 2014 and 2017 to evaluate child development and quality of public preschool services. The first used the Early Childhood Environment Rating Scale. The second used the Measuring Early Learning Quality and Outcomes. The report and data were expected by October 2018. The Ministry plans to refine curriculum and training, based on findings and analysis from the report.

Peru (and its students) participated in international assessments such as PISA ( in 2012, 2014, and a pilot in 2017), Tercer Estudio Regional Comparativo Explicativo (TERCE) in mathematics (grade 3 and 6), reading (grade 3 and 6), writing (grade 3 and 6), and natural sciences (grade 6 only ) in 2013, and ICCS (in 2015) that assessed citizen preparation.

The Ministry conducted a study of potential biases in tests administered to indigenous students. The ICR notes that the Ministry used this information to improve the tests (and their administration). For example, a radio campaign was conducted in four indigenous languages to inform parents of tests and their purpose.

**Rating**  
Substantial

## **Objective 2**

### **Objective**

Improve Ministry of Education's capacity to evaluate instructional practices.

### **Rationale**

The theory of change supported evaluation of instructional practices through classroom observational instruments and development of a teacher evaluation system as a means to improve teacher performance. The evaluation system can be a direct way of improving teachers' instructional practices and capacity, if the system is well designed and implemented. Additional steps are also needed that were not specified in the theory of change. For example, undertrained and low capacity teachers need structured teachers' guides, training in content and pedagogy (that models new teaching techniques) and follow-up support. (Villegas-Reimers 2003).



The objective was measured through the indicator: annual results of national studies of instructional practice at the primary and secondary level reported at the national level.

The results of four studies of classroom instruction were reviewed (met target). The Stallings Classroom Snapshot was deployed to measure teachers' use of instruction time in primary (2012 and 2014) and secondary (2013 and 2015) levels. Findings were distributed to the Ministry's Regional Education Directorates (met target). The Government noted that the teams in the Ministry analyzed results, which were published in a series of research papers and briefs detailing analysis of pedagogical practices (ICR, p. 63). An additional report was published related to preschool teachers' pedagogical practices from the performance evaluation conducted by the Ministry in 2017. The ICR did not discuss the result (of the studies), nor planned actions.

A teacher performance evaluation system was designed and implemented with preschool teachers (the target specified implementation of the system but not the scale of implementation). 6,000 pre-school teachers were assessed according to the Bank team. New teachers and those who could be promoted were part of the initial phase of the evaluation. The recruitment and promotion of pre-school teachers is merit-based, which is initiated by a process of regular performance evaluation and feedback. Implementation of the evaluation system was delayed- due to teacher opposition (including a strike). Additional outreach measures were taken with the Teachers Union, including providing information and training to teachers, disseminating the rubrics and instruments, and meeting with representatives of the Teachers Union to ensure the process was clear. If a teacher does not perform well on the classroom observation, he/she received professional development assistance and could participate in a subsequent evaluation. The Bank team reported that the evaluation focused on classroom management, rather than content knowledge.

The promotion instrument was implemented in 2015, 2016, and 2017, resulting in promotion of 49,763 out of 410,429 teachers who were evaluated. However, it is difficult to interpret these numbers as the ICR does not describe the previous promotion trend. Access to leadership positions and salary increases resulted from the promotion instrument, according to the Bank team. This instrument and the evaluation may provide incentives for teachers to perform better.

The recruitment instrument was implemented in 2015 and 2017 and resulted in the recruitment of 19,069 teachers out of a total of 400,423 candidates for both years. However, the ICR does not provide prior recruitment data to identify demand and supply and assess how the changes facilitated a merit-basis to the selection and hiring.

The Ministry has made an effort to evaluate teachers in their mastery of indigenous languages in relevant classrooms.

**Rating**  
Substantial



### Objective 3 Objective

Improve the Ministry of Education's capacity to evaluate school leadership in basic education.

#### Rationale

The theory of change focused on selecting qualified principals and training school leaders and certifying them in school management. Instructional leaders are a crucial input to improve school management, education quality, and student outcomes. Effective principals are one of the important factors explaining education quality and student learning. The process to select principals includes: standardized assessment (reviewed centrally) and then the decentralized level applied an instrument to evaluate candidates. Training was implemented by selected universities in school management and pedagogical leadership.

The objective was measured through indicators such as the number of principals staffed through the evaluation system.

A principal evaluation system was designed and implemented. The first performance evaluation of school principals was initiated in 2014. The ICR does not discuss how many applicants applied to principal positions prior to 2014 to assess the relative attractiveness of the position, nor how this new process increased competition.

18,477 school management positions were staffed through the school principal evaluation system. For example, positions were filled from 51,563 candidates for 20,000 positions in 2014. 12,562 open positions selected from 13,421 evaluated candidates, resulting in 9,113 filled positions. The ICR noted that a smaller number were staffed by Local Education Management Units because of evaluation requirements, imbalance in demand for positions; and legal disputes; several positions were not filled. The competitive evaluation process reviewed 53,699 candidates for these positions. The target for the number of school management positions (staffed through the principal evaluation process) was revised from 25,000 to 18,000 due to factors beyond the control of the Ministry. The ICR broke down the indicator in two ways. The number of qualified candidates selected for open positions was 24,584 and that of qualified candidates accepting offers was 18,447. The Bank team clarified that some appointments were not confirmed locally because some candidates preferred to remain in the local area (as the offer was for a school outside their home community) and some principals filed protective measures to keep their positions. The evaluation involved an initial selection stage, where candidates were given an examination. The process also included an assessment of candidates by the decentralized level of the Ministry. These aspects were not described in the ICR.

The ICR notes that merit-based promotion increased the share of female teachers being promoted to management positions. There was an increase of 26 percent for school leaders and 107 percent for deputy director positions. Evidence of school leaders' performance; however, has not yet been generated.



A principal training course was designed and carried out by training institutes. The training program consisted of a one-year certification and a one-semester complementary second degree. The Bank team indicated that the training was originally planned to last four months, but was extended for 18 months. 13,536 principals were certified in school management. (target 10,000) The Bank team indicated that the training program was evaluated. While no data were reported, the Bank team said beneficiaries found the training's focus on management and soft skills useful. 12,000 trained principals were able to evaluate the performance of their teachers- which reduced the need to hire an outside evaluator to conduct the teacher performance evaluation. Outside evaluators were needed when trained principals were not present.

The Ministry has also developed a certification to demonstrate proficiency in indigenous languages, which is a prerequisite for appointing school principals that serve indigenous communities.

The Ministry has launched a third process to staff 9,000 additional school management positions, which should be complete in 2018, showing sustainability of the implementation of the system. 50,000 teachers have registered for these positions.

**Rating**  
Substantial

## Rationale

The ICR reports the planned indicators which are output-orientated measures demonstrating implementation of additional student assessment, evaluation of instructional practices and school leadership. It does not provide evidence of the rigor or quality of these evaluations, which would demonstrate improved capacity of the Ministry. The extent to which capacity has been enhanced is unclear. The project, through the implementation of additional student assessments, and of expanded appraisal of teaching performance and training of additional school managers, has laid the foundation for later outcomes. Therefore, efficacy is rated substantial, which is indicative of minor shortcomings in the attainment of the operation's objectives.

**Overall Efficacy Rating**  
Substantial

## 5. Efficiency



An economic analysis prepared at appraisal calculated an internal rate of return of 10.1% and net present value of US\$ 417 million. These were estimated only from investment in the student learning evaluation system, which may underestimate the overall benefit. The economic analysis was updated at closure and found an internal rate of return between 10.7 percent and 13.8 percent (the lower bound being based on conservative estimates of the impact of scaled-up student evaluations alone, and the upper bound including that of better institutional practice and school management) . The project was expected to increase future earnings of project beneficiaries- those exposed to the new student assessment system during their school years and who complete secondary school (ICR. p. 50). This illustrates the value of the investment, as well as education investments broadly.

The ICR (p. 17) discusses the potential impact on sector efficiency and service delivery of the quality improvements financed by the operation, including improved data for design of education policies and increased accountability for school performance, although it acknowledges that it is still too early to identify concrete results." Thus, the ICR is not able to document improvements in efficiency of the service delivery.

Other qualitative aspects of efficiency include: low overhead costs (1.04% of total project cost), which is half the anticipated amount at preparation. The low rate is the result of implementation by units within the Ministry supported by consultants, consistent with on-budget financing by the Bank. There was punctual implementation of contracts and payment, which established a good reputation among bidders and providers, thereby attracting more competition and better candidates (ICR p. 61). The ICR states that this resulted in reducing the cost of services that were contracted, but no data were provided to illustrate this. The project closed without an extension. More activities were completed than originally planned, for example extended specialized training for selected school principals, development of a new teacher evaluation system, and participation in an additional international assessment.

There was delay in project effectiveness due to the time needed to approve the Operational Manual. The Ministry implemented the project prior to effectiveness with its budget funding so that the delayed effectiveness did not negatively affect implementation. The Bank paid the first disbursement in 2015. There was a delay in activities at the beginning of each calendar year, as the project had to "prepare a plan twice a year with incomplete or dated information" (ICR p.28) The ICR noted that "the project was challenged by the different schedules of the Government’s budget cycle (preparation in April–May) and the preparation of the Project’s Annual Operating Plan (AOP, November–January)".

Considering all of these aspects, project efficiency is rated substantial, consistent with minor shortcomings.

**Efficiency Rating**

Substantial

a. If available, enter the Economic Rate of Return (ERR) and/or Financial Rate of Return (FRR) at appraisal and the re-estimated value at evaluation:

Rate Available?	Point value (%)	*Coverage/Scope (%)
-----------------	-----------------	---------------------



Appraisal	✓	10.00	80.00 <input type="checkbox"/> Not Applicable
ICR Estimate	✓	14.00	100.00 <input type="checkbox"/> Not Applicable

\* Refers to percent of total project cost for which ERR/FRR was calculated.

## 6. Outcome

Relevance of objectives is assessed across the entire project and is rated High. Efficacy is rated Substantial overall (and for each of the objectives). Efficiency is also rated Substantial given the results of the quantitative analysis and smooth implementation with minor delays. Outcome is therefore rated Satisfactory, indicative of minor shortcomings in the operation’s achievement of its objectives.

### a. Outcome Rating

Satisfactory

## 7. Risk to Development Outcome

The project has likely enhanced the capacity of the Ministry of Education in relation to assessment of student learning, instructional practices, and school leadership that are sustained. The Ministry continues to implement the Teacher Reform law for recruitment, evaluation, and promotion. Lessons learned over the course of the operation have remained with staff in the Ministry. The teacher evaluation system was implemented with preschool teachers and will likely be expanded. The Ministry supported training two cohorts of school leaders. It carried out student learning assessment in additional grades and subjects. It plans to continue international assessment in 2018. The Ministry has plans to design and deploy an assessment for the last year of secondary education (11th grade) (ICR p. 14-15).

The Bank continues to provide financial and technical resources to reform basic and secondary education. The Government is participating in the Bank- financed Human Capital Project that should raise political awareness of the importance in investing in education and improving education quality. The Bank is also targeting improvements in human capital with a Deferred Drawdown Option employing indicators related to teachers, school principals, and measurement of school learning.

However, political risks to the development outcome remain. These were already evident during implementation - a teachers strike from June to September 2017 impeded the teacher evaluation and led to cancelation of the student assessment scheduled for November 2017).



## **8. Assessment of Bank Performance**

### **a. Quality-at-Entry**

This operation built upon lessons from the previous one. For example, there was insufficient attention paid in the latter to rural and bilingual education, and hence regressive education policies remained unaddressed. Additional lessons learned included: the need to focus on activities central to the Government's strategy; emphasizing institutional capacity building and the use of data; and building public awareness and confidence. Integrating disbursement-linked indicators into the financing was appropriate and consistent with the Ministry of Finance's approach that prioritized sector resources through results-based budgeting (ICR p. 21).

To complement the financing in this operation the Bank also supported evaluation of student learning, instructional practices and school leadership with technical assistance, trust funds, and learning resources to assist staff in the Ministry of Education. Design also provided for non-lending support to increase the impact of student assessment data to inform education policy, and to build awareness and confidence in the results. The project supported the development of performance evaluation systems for teachers and school leaders.

Overall, preparation was sound. A clear and straightforward method of compliance with disbursement requirements was indicated. Design was simple, and the scope and number of activities aligned well with the implementation period. There was in-built flexibility, enabling the additional activities subsequently requested by the Borrower to be accommodated. The Bank brought considerable technical expertise and provided technical assistance as needed. The project was ready for implementation upon approval (as already noted, the Ministry was able to begin implementing with its own resources, while the required approval of the Operational Manual delayed effectiveness).

There were some minor weaknesses, several of them pointed out by the Borrower. It would have been better to have selected a unit in the Ministry to coordinate the work that was hierarchically above those implementing the activities (ICR p. 58). The arrangement was later changed. The Government believed that market studies should have been part of preparation (ICR p. 61). M&E design could have been stronger, since indicators were output-oriented (see Section 9a below).

**Quality-at-Entry Rating**  
Satisfactory

### **b. Quality of supervision**



The Bank was responsive to requests made by the Government. As a result, there was a change in governmental supervisory responsibility, which facilitated better coordination of project activities. As noted in the previous section, the new unit was hierarchically above all the directorates implementing activities, and was thus better placed to oversee and coordinate the activities. When the Minister prioritized implementation of the Teacher Reform Law, the Bank responded by financing the design of the new evaluation instruments and system for merit-based recruitment and promotion. The Bank supported the efforts of the Government in expanding its student assessment experience by facilitating international advisers with experience in student assessment and teacher evaluation.

There was continuity in the supervision team, including the task team leader who was based in Peru. These factors aided in developing a good working relationship with the teams in the Ministry. The Bank was proactive in resolving potential problems. For example, the Bank's team worked with the Government to identify expenditures that could be considered eligible (ICR p. 26). The midterm review, held in June 2016, resulted in additional interventions to ensure Indigenous People benefited from project activities.

The team implemented two level two restructurings to align implementation to the strategies of the Government and its actual implementation. During the September 2015 restructuring: (i) the implementation schedule was modified to allow time to complete additional activities; (ii) the legal covenant was clarified to specify hiring of an independent verification agent, rather than "Agency"; (iii) disbursement arrangements and amounts were brought into line with revised cost estimates; (iv) the definitions of eligible budget activities, and of basic education were changed ; and the results framework was updated (while the ICR argues that there were improvements in the outcome indicators, the changes consisted essentially of revised targets to reflect actual implementation).

However, the reimbursement process did not match up with the annual activity planning and budget preparation process. Thus, the majority of the activities in the project's annual activity plan did not have a budget allocated to them at the beginning of the year, so that budget modifications were needed to begin the procurement process (ICR p. 59).

### **Quality of Supervision Rating**

Satisfactory

### **Overall Bank Performance Rating**

Satisfactory

## **9. M&E Design, Implementation, & Utilization**

### **a. M&E Design**

The ICR suggests that the DLIs were the right ones to measure impact, stating "the Results Framework and DLIs provided a well-defined yet flexible framework with which to monitor implementation progress and



disburse” (ICR p. 24). However, indicators were output-oriented and could monitor only implementation of activities; they did not provide evidence of the quality or rigor of the outputs (see Section 4 above). The unit in charge of coordinating project activities within the Ministry was responsible for supervision of monitoring and evaluation.

## **b. M&E Implementation**

The unit within the Ministry (Strategic Planning Secretariat) that coordinated overall project activities implemented the M&E responsibilities well and provided timely reporting of the indicators. Given that the loan disbursed annually based on eligible budget activities and disbursement-linked indicators, an independent verification agency carried out a review of progress of disbursement-linked indicators.

At the restructuring in 2015, changes were made in responsibility for data collection, better to reflect the input of the line units which were responsible for project implementation. The restructuring also revised targets to be non-cumulative to be consistent with the disbursement-linked indicators. The number of intermediate results indicators was increased from five to ten to measure steps to track performance. Three indicators were revised from a yes/no scale to a number. Some targets were lowered to reflect actual implementation. According to the Government, this was justified since targets for some indicators were too ambitious (ICR p. 59). The timeline for data collection was expanded to five years, given the additional activities and initial delay in effectiveness. The original time frame planned on data collection for only four years even though the length of the loan was five years.

## **c. M&E Utilization**

The ICR notes that data from the evaluation systems were used by the Ministry, but how they were used is not explained. The ICR states that dissemination of student assessment results brought public awareness and buy-in for the reforms. For example, public support increased for the teacher evaluation system, despite the teacher strike.

### **M&E Quality Rating**

Substantial

## **10. Other Issues**

### **a. Safeguards**



The project was classified as Category C for environmental assessment purposes, as it did not include construction or infrastructure rehabilitation. The Indigenous Peoples safeguard (OP 4.10) was triggered. A Social Assessment and Indigenous People Planning Framework were prepared and disclosed ( ICR p. 25). The Social Assessment recommended strategies to minimize bias against indigenous students in standardized tests- due to cultural and language differences, which were implemented. For example, it recommended that indigenous families and communities should have timely access to culturally appropriate reports of results, and that cultural and linguistic profiles of principal candidates should be considered when selecting school principals in relevant schools.

### **b. Fiduciary Compliance**

Fiduciary management was the responsibility of the Administrative Coordination Team within the Ministry experienced in handling prior World Bank operations. This team also handled projects financed by the Inter American Development Bank, thus making it well placed for fiduciary management. Compliance with financial management was reported to be satisfactory (ICR p. 25-26). Mitigation measure proposed at preparation such as specific arrangements, recruiting additional members to the technical team were adhered to by the Implementing Agency. The Legal Agreement required the Ministry to have an independent agency carry out technical reviews of agreed results during implementation. This was complied with.

The Administrative Coordination Team hired procurement specialists to support the operation. Bank procurement policies were satisfactorily complied with (ICR p. 26). The ICR nonetheless notes that the Bank's procurement policies and procedures placed a heavy burden on the implementing unit. Three ex post procurement reviews were performed (in 2015, 2016, and 2018), which, together with the Bank's supervision missions, provided recommendations to improve procurement management; these recommendations were implemented. Procurement involved some learning, as the Bank's procedures expanded the pool of providers to make the process more competitive; this different from national procedures which defined service requirements more narrowly. The Government noted ongoing communication between the technical units (implementing activities) and the unit in charge of procurement.

### **c. Unintended impacts (Positive or Negative)**

---

### **d. Other**

---



## 11. Ratings

Ratings	ICR	IEG	Reason for Disagreements/Comment
Outcome	Highly Satisfactory	Satisfactory	A rating of Satisfactory is indicative of minor shortcomings in the design, implementation, and attainment of the objectives. The project has laid the foundation for potential outcomes, but evidence of improved capacity was lacking in the ICR.
Bank Performance	Satisfactory	Satisfactory	---
Quality of M&E	High	Substantial	M&E design provide evidence of implementation of activities, but did not permit monitoring of impact.
Quality of ICR		Substantial	---

## 12. Lessons

The following lessons are taken from the ICR with some adaptation of language:

**Government commitment is a necessary condition, particularly when implementing politically sensitive reforms, but it is also necessary to build awareness for the changes and of the reasons for them.** The Government requested adding implementation of the Teacher Reform Law to the operation. It knew that implementation of the Law would be politically difficult. Following a teacher strike and ministerial resignations, a plan was put in place to address the concerns expressed by teachers, and the Ministry moved forward with the performance evaluation of teachers. The plan involved consultation and outreach with teachers, unions, and local authorities. It also involved training principals to be able to evaluate teachers in their schools. This reinforced public awareness of the need for change and increased political support for reform.

**Timely project restructuring can facilitate implementation and have impact on results.** The restructurings were important to increase the project's relevance to the borrower's priorities. They also helped to increase the scope of the project and thus increase its impact.

**Conditions of effectiveness should match the instrument and should be reserved for critical issues.** In this case, the Bank was financing eligible budget activities through a Sector-Wide Approach, which meant that individual activities were not required to begin implementation. Yet, effectiveness was held up for the approval of the Operational Manual, which required multiple levels of review. This process was made more difficult, as



there were changing administrations. Thus, effectiveness conditions should focus on the most critical dimensions.

**Implementing an effective teacher evaluation system needs to balance multiple aspects.** Initially, the evaluation of teachers was lenient with standards so that the majority could pass. However, an evaluation system should not focus on removal of teachers. Rather, there is a need to ensure a critical mass of teachers in the system. The evaluation aimed to identify training needs. Thus, when designing an evaluation system, it is important to understand its goal. An evaluation system requires transparency to ensure teachers understand the metric and process. A communication campaign was critical to build public awareness and gain their support.

### 13. Assessment Recommended?

Yes

Please explain

A field-based assessment would help to bring out details not adequately discussed in ICR. Moreover, the passing of time will help identify actual achievements realized by the operation.

### 14. Comments on Quality of ICR

The ICR is concise and provides a thorough overview project preparation and implementation, including the factors that facilitated and hindered implementation. It contained an updated economic analysis. It was consistent with the guidelines. Yet, there was limited discussion and evidence related to efficacy. By focusing only on output-oriented indicators without discussion of their quality and rigor, evidence of improved capacity of the Ministry is obscured. The Borrower's completion report was excellent and provided helpful information.

#### a. Quality of ICR Rating

Substantial