DRAFT Guidelines for Reviewing World Bank Implementation Completion and Results Reports

A Manual for IEG ICR Reviewers

Last updated: July 25, 2017

Note: This is a living document under ongoing revision.

IEG
Independent Evaluation Group
WORLD BANK GROUP
World Bank • IFC • MIGA

Contents

Tables

Boxes

Figures

Abbreviations and Acronyms

APL	Adaptable Program Loan
BP	Bank Procedure (World Bank)
CAS	country assistance strategy
DCA	Development Credit Agreement
DGA	Development Grant Agreement
DPL	development policy lending
DPO	development policy operation
ERR	economic rate of return
FRR	financial rate of return
GEF	Global Environment Facility
GEO	Global Environmental Objective
ICR	Implementation Completion and Results Report
ICRR	Implementation Completion and Results Report Review
IBRD	International Bank for Reconstruction and Development (World Bank)
IDA	International Development Association
IEG	Independent Evaluation Group
IRR	internal rate of return
KPI	key performance indicator
LA	lending agreement
M&E	monitoring and evaluation
NCO	Note of Canceled Operation
NPV	net present value
OP	Operational Policy (World Bank)
OPCS	Operations Policy and Country Services Vice Presidency of the World Bank
OPSPQ	Operations Policy and Quality Department of the World Bank
PAD	project appraisal document
PD	program document
PDO	project development objective
PPAR	Project Performance Assessment Report
QEA	Quality-at-Entry Assessment
QSA	Quality of Supervision Assessment
TTL	task team leader

1. Introduction

What is an Implementation Completion and Results Report?

The Implementation Completion and Results Report (ICR) is one of the World Bank's main instruments for self-evaluation. It is prepared by the World Bank at the close of every project funded by the International Development Association (IDA) or the International Bank for Reconstruction and Development (IBRD), or, in the case of a series of programmatic policy operations, at the end of a series of projects. See the OPCS ICR Guidelines for additional information and exceptions.

According to the guidelines to World Bank staff for preparing ICRs, they are intended to:

- Provide a complete and systematic account of the performance and results of each project.
- Capture and disseminate experience from the design and implementation of a project in order to: (i) improve the selection of interventions to achieve the goals of the country partnership framework (CPF) (or, previously, the country assistance strategy [CAS]); (ii) improve the design and implementation of interventions through lessons learned; and (iii) help ensure greater development impact and sustainability of projects.
- Provide accountability and transparency at the level of individual projects with respect to the activities of the Bank, the borrower, and involved stakeholders.
- Provide a vehicle for realistic self-evaluation of performance by the Bank and borrowers.
- Contribute to databases for aggregation, analysis, and reporting, especially by the Independent Evaluation Group (IEG), on the effectiveness of lending projects in contributing to development strategies at the sector, country, and global levels.

ICRs are intended to contribute to accountability and learning with the audience for ICRs being both internal (the Board members and Bank managers and staff) and external (governments and their agencies, stakeholders, and beneficiaries in partner countries, as well as the general public). The final ICR is publicly disclosed when it is submitted to the Board unless exceptional circumstances argue against this.

What is an ICR Review?

The ICR Review, conducted by IEG, is an independent, desk-based, critical review of the evidence, results, and ratings of the ICR in relation to the project's design documents. Based on the evidence provided in the ICR and an interview with the last task team leader (TTL), IEG arrives at its own ratings for the project, based on the same evaluation criteria used by the Bank.[1] IEG reviews each ICR that is submitted to IEG.

Thus, in reviewing the findings and ratings in the ICR, IEG provides an independent view of the results and ratings, conditioned on the evidence presented in the ICR and from the last TTL for the project. However, IEG is not privy to evidence that was not included in the ICR. The ICR Review is thus an independent validation of the Bank's self-evaluation and ratings; it is not an independent evaluation of the project based on evidence collected outside the Bank's self-evaluation process.[2]

The ICR Review is intended to critically assess the evidence provided in the ICR, its quality, and the attribution of results to the activities or actions supported by the project under review. It is not simply a summary of what is in the ICR.

ICR Reviews serve as an independent validation of the results reported in the ICR and contribute to both learning and accountability. They also provide a systematic way for IEG to critically review the evolving portfolio as projects close and to summarize the projects' objectives and key results, in addition to the ratings. The write-ups are stored in a searchable database within IEG and, for all projects that closed from fiscal 2011 onward, are posted on IEG's external website.3 They are often useful as a starting point for IEG's ICR Reviewers as a quick way to identify projects of different types - with specific objectives or activities - in preparing to undertake larger country, sector, or thematic evaluations.

On what basis does IEG assess projects?

The World Bank and IEG share a common, objectives-based project evaluation methodology for World Bank projects to assess achievements against each project's stated objectives, as well as the relevance of the objectives and the efficiency of resource use in achieving the objectives. An advantage of this methodology is that it can accommodate country context in terms of setting objectives that are reasonable. The Bank and the governments are accountable for delivering results based on those objectives.

---

[1] When insufficient information is provided by the Bank for IEG to arrive at a clear rating, IEG downgrades the relevant ratings as warranted. This practice began on July 1, 2006.

[2] For a subset of operations—on the order of 20–25 percent—IEG conducts Project Performance Assessments in the field.

[3] ICR Reviews older than five years are declassified and disclosed on a quarterly basis.

What are the main ratings?

There are three main project ratings that IEG validates through the ICR Review, and one rating that is assigned by IEG only.[4] The three main project ratings are:

- **Outcome**: the extent to which the project's major relevant objectives were achieved, or are expected to be achieved, efficiently. Both the Bank and IEG rate Outcome. {should we mention the subratings, i.e. Relevance, Efficacy, Efficiency? Or too much detail for this location?}

- **Bank Performance:** the extent to which the services provided by the Bank ensured quality at entry of the project and supported effective implementation through appropriate supervision (including ensuring adequate transition arrangements for regular operation of supported activities after loan/credit closing), toward the achievement of development outcomes. Both the Bank and IEG rate Bank Performance. IEG also rates the two constituent elements – Quality at Entry and Quality of Supervision.

- **Monitoring and Evaluation Quality:** the quality of the design and implementation of the monitoring and evaluation arrangements of the project and the extent to which the results are used to improve performance. Both the Bank and IEG rate M&E Quality.[5]

In addition, IEG assigns one additional rating based on the material presented in the ICR:

- **ICR Quality:** the quality of the evidence and analysis in the ICR; the extent to which the lessons are based on evidence; the results-orientation of the ICR; and its conciseness, internal consistency, and the consistency with Bank guidelines. Only IEG rates ICR Quality.

Structure of this Manual

The Manual is organized into three parts, with appendixes.

- Sections 1 - 2 provide an overview and explain the responsibilities of the IEG ICR Reviewer, the materials to consult, and the ICR Review process. Line-by-line guidelines for completing the ICR Review form are in appendix A at the end of the Manual.

---

[4] For ICRs written before July 1, 2017, ICRs and IEG also rated Risk to Development Outcome (RDO) and Borrower Performance. RDO is the risk, at the time of evaluation, that development outcomes (or expected outcomes) will not be maintained (or realized). Borrower Performance is the extent to which the borrower (including the government and the implementing agency or agencies) ensured quality of preparation and implementation and complied with covenants and agreements, toward the achievement of development outcomes. Also, before July 1, 2017, IEG discussed and rated M&E Quality in the ICR Review, while the Bank did not do so in the ICR.

[5] For the Bank, rating M&E Quality is new beginning July 1, 2017.

- Sections 3–12 cover identification of the objectives and the criteria for the main ratings.

- Sections 13–16 are devoted to the definition and criteria for other issues covered in the ICR Review that are not rated—safeguards, fiduciary issues, unintended outcomes, lessons. There is also a section on assessing canceled projects, for which the Bank will issue a Note of Canceled Operation in lieu of an ICR. These Notes are also reviewed using the ICR Review form.

Three main appendixes present: ICR Review Checklist (appendix A—a work in progress); and the protocol for meeting the TTL of the project (appendix B).

Sections 1–11, 13, 14, and 16 apply equally as guidelines for Project Performance Assessment Reports (PPARs).

2. Procedures for the ICR Review

Responsibilities of the IEG ICR Reviewer

The IEG ICR Reviewer is responsible for:

- Correctly completing the ICR Review form in the online IEG ICRR Portal system, following the guidelines and procedures in this Manual and in the Onboarding Guidelines for IEG ICR Reviewers.

- Assigning ratings based on the evidence in the ICR and that gleaned by consulting the other key documents listed below.

- Meeting with the project's last TTL, recording a summary of the meeting, and updating the draft ICR Review and ratings to reflect any new and relevant information.

- Revising the ICR Review and ratings based on comments from
  - a member of the review panel (composed of senior staff and consultants in IEG),
  - the ICR Review coordinator, and/or
  - the IEG manager.

- Reviewing written comments from the Global Practice or Region that managed the project, incorporating any new and relevant information, correcting any inaccuracies, updating ratings if warranted, and drafting a response to the Global Practice to explain any updates.

The review process and the specific steps involved and expectations of the IEG ICR Reviewer at each stage are discussed below.

IEG's ICR review process

The ICR for a project arrives in IEG electronically after it has been approved in the Operations Portal and sent to the Board through the relevant system. The ICR Review Coordinator then assigns it to an IEG ICR Reviewer for review. A blank ICR Review form for the project is automatically created in the IEG ICR Review Portal, based on its project ID (see figure 2.1, a blank ICR Review form for an investment project). Certain fields in the basic data portion of the form will be automatically populated (the project name, project ID, and TTL).

ASSEMBLING THE KEY DOCUMENTS

In preparing the first draft of the ICR Review, the IEG ICR Reviewer assembles the key documents listed below. The IEG ICR Reviewer is not expected to go beyond these documents to look for additional evidence.

- The Financing agreement (loan, credit, or grant agreement) - primarily for use in verifying the project's original objectives and components. If the legal agreement was amended, the amended agreement(s) will also be provided.

- The project appraisal document (PAD, for investment projects) - primarily for use in identifying the project's original objectives, components, planned amounts, cofinanciers, results framework, planned monitoring and evaluation (M&E) and the presence of baseline information, safeguard category (for investment projects), and other aspects of design. If the project has been restructured, there will also be a project paper.

- The country partnership framework (previously known as country assistance strategy) in effect at project closing (as well as the one in effect at approval if different) - primarily for use in assessing the project's relevance.

- Implementation Completion and Results Report (ICR) - the Bank's self-assessment of the project and the main document to review. The ICR includes information on: *revisions to the design* (such as restructuring and changes in objectives or components, allocations, cofinanciers, or expected counterpart contributions); the *implementation of project activities*; the implementing unit's assessment of the project's *outcomes*, the *relevance of the* project, the *achievement of its objectives,* the project's *efficiency*, and *safeguard and fiduciary compliance*; operational staff's self-*ratings (*on outcome, Bank performance, and M&E quality); and the lessons learned from the experience. These reports often include an assessment by the borrower as an appendix and, occasionally, an assessment by cofinanciers of the results, in addition to financial or economic analysis and the results of beneficiary surveys.[6] {update this paragraph and reference after OPCS guidance is complete} The ICR is expected to provide evidence for all statements made. Evidence may be provided in appendices to the ICR or in links referred to in the ICR. The ICR Reviewer should consult the evidence provided or in the

[6] For greater detail on ICR guidelines, see World Bank, OPCS 2011a.

ICR or the ICR Appendices. The ICR Reviewer should also consult evidence found in sources that are explained and linked to from the ICR or the ICR Appendices.

PREPARING THE INITIAL DRAFT

The IEG ICR Reviewer is expected to read all the above documents and use that information as indicated in this Manual and its Appendices, as well as in the OPCS ICR Guidelines, to draft the ICR Review. For instructions on use of the online IEG ICRR Portal, see the Administrative Manual for IEG ICR Reviewers. Before finalizing the draft, the IEG ICR Reviewer must contact the last TTL of the project to set up an interview.

INTERVIEWING THE LAST TASK TEAM LEADER

This interview, which is conducted before the draft ICR Review is finalized, provides an opportunity for the last TTL of the project to offer additional views or information to the IEG ICR Reviewer (beyond what is in the ICR) about the project experience. The interview also provides an opportunity for the IEG ICR Reviewer to pose any follow-up questions that arose in the course of reading the ICR, to improve the accuracy and quality of the ICR Review. The IEG ICR Reviewer should contact the TTL early in the process of drafting the ICR Review, to check availability, travel schedule, and to propose a meeting either in person or by audio. If the TTL is not responsive after three weeks to the request for a meeting, the ICR reviewer may proceed to finalize the draft. Note that the finalized draft ICR Review is later shared with the Global Practice for comments, and at that point the TTL also will have an opportunity to provide any additional information. If the IEG ICR Reviewer is a junior staff member, he/she is accompanied to the meeting by the ICR Review Coordinator. The IEG ICR Reviewer should not share the draft Review with the TTL, however, or share the proposed ratings.

Following the TTL interview, the IEG ICR Reviewer writes a summary of the meeting, uploads the summary to the Activity History for this ICR Review in the IEG ICRR Portal system, and shares it with the Panel reviewer when he/she is identified. The detailed protocol for the TTL interview is in appendix C.

SUBMITTING THE DRAFT FOR PANEL REVIEW

Following the interview with the TTL, the IEG ICR Reviewer updates the ICR Review with any new and relevant information from the TTL, if warranted, making sure that the source is identified as the project TTL (to differentiate it from evidence found in the ICR). After a final “Save” and “Exit”, the IEG ICR Reviewer uses the activity history pane to send the ICR Review back to the ICR Review Coordinator, who then identifies an appropriate Panel reviewer and assigns it to a Panel reviewer. (Detailed instructions on use of the system are to be added to the Administrative Manual for IEG ICR Reviewers.

Panel review

After the IEG ICR Reviewer sends the completed draft ICR Review to the ICR Review coordinator, the coordinator selects an appropriate Panel reviewer from the IEG Evaluation Panel, comprising senior evaluators. When possible and appropriate, the coordinator identifies the Panel reviewer in advance, so that the Panel reviewer and ICR Reviewer can agree on how to structure the Efficacy section – that is, what is the understanding of the PDO in terms of the objectives to be achieved. The tasks of the Panel reviewer are to:

- review the same documents as the IEG ICR Reviewer;
- read the ICR Review;
- ensure that the objectives have been properly identified;
- ensure that the guidelines have been properly applied;
- ensure that the ICR Review is complete, internally consistent, and sufficiently critical of the quality of the data and analysis; and
- comment on the ratings.

The Panel reviewer provides comments to the IEG ICR Reviewer by sending a message through the IEG ICRR Portal system. Many Panel reviewers find it convenient to download the draft ICR Review as a Word document, add comments to that document using track changes or comment boxes, and attach the Word document along with the message to the IEG ICR Reviewer. Comments from a Panel reviewer generally indicate areas of agreement, but also areas for improvement, areas of disagreement, and queries about the evidence.

The IEG ICR Reviewer then revises the ICR Review as needed and responds to the Panel reviewer as needed. The discussion can go back and forth several times. If the Panel reviewer and IEG ICR Reviewer cannot reach agreement, the ICR Review Coordinator may ask for a third opinion from the Thematic coordinator or another Panel member. When the Panel reviewer is satisfied that the ICR Review is ready and there is agreement, he/she clears the ICR Review in the system to send it to the ICR Review coordinator.

While it is tempting for the discussions between the Panel reviewer and the IEG ICR Reviewer to take place face-to-face or by regular e-mail, it is important to record these discussions in the IEG ICRR Portal system, because this is the only way that the discourse and any issues and their resolution can be retained in IEG's institutional memory and can be reviewed by the ICR Review coordinator and IEG manager. Recording interactions in the system also helps reinforce consistency across ICR Reviews.

Quality Check by ICR Review Coordinator and/or IEG Unit Manager

Following clearance by the Panel reviewer, the ICR Review coordinator and/or IEG manager reviews the draft and sends any questions or comments to the IEG ICR Reviewer and Panel reviewer, who address the questions and comments as needed.

GLOBAL PRACTICE REVIEW

Following clearance by the ICR Review coordinator and/or IEG manager, the ICR Review coordinator and/or IEG manager sends the draft ICR Review to the Global Practice senior director for comment, through the IEG ICRR Portal system.[7] The Global Practice senior director is responsible for forwarding it to the people on the country team most familiar with the project for comment, and for coordinating the response to IEG. The Global Practice senior director or country director has the option of inviting the borrower to comment. The standard review period for the Global Practice and the borrower is 2 weeks (10 business days).[8]

In its response, the Global Practice or borrower may point to factual corrections needed, suggest changes in the text, or indicate disagreement on the ratings. Often, certain information already presented in the ICR is repeated, but the Global Practice may also provide additional relevant and credible information concerning achievement of the objectives (or other aspects) not already in the ICR.

FINALIZING ICR REVIEW AND RESPONDING TO GLOBAL PRACTICE COMMENTS

The IEG ICR Reviewer should take on board any additional relevant information from Global Practice comments that is credible and will improve the accuracy of the assessment, but indicate in the text that this additional information was “provided by the Global Practice” (to distinguish it from the information in the ICR), and cite the source of the information, if known.

Since the IEG ICR Reviewer has already fully assessed the project with respect to all the information in the ICR, a response from the Global Practice that simply reiterates that same information would not be expected to result in changes in the ICR Review. However, new information, corrections, or new (and compelling) lines of argument could result in changes in the ICR Review text or ratings.

One rating that typically would not be affected by a response from the Global Practice is that for the quality of the ICR. An unsatisfactory ICR quality rating usually leads to substantial comments from the Global Practice, but in this case, there is no new evidence to influence the rating: the ICR itself is the only evidence required to assess its quality. In a case where the Global Practice comments include information additional to the ICR that was already available at the time of preparation of the ICR, however, it may be appropriate to assign a lower rating for quality of the ICR because it now becomes apparent the ICR itself contained incomplete information.

The comments from the Global Practice should be discussed with the Panel reviewer, and any proposed changes should be cleared by him/her. The ICR Review form is then modified and

---

[7] The draft Review is also usually copied to the country director, the Regional sector managers, the last TTL of the project, and key individuals in the Global Practice. Identification of the individuals to be copied is the responsibility of the ICR Review coordinator.

[8] The Global Practice senior director may ask for an extension if necessary (for example, if key staff are on annual leave or a mission and unavailable to reply).

saved, and the IEG ICR Reviewer drafts a response to the Global Practice, also cleared by the Panel reviewer, to be sent from the ICR Review coordinator and/or IEG manager back to the Global Practice. The response should begin by thanking the Region for their comments, followed by a succinct summary of any changes made based on the additional information and the reasons other information was not used (if this is the case), with the revised ICR Review attached.

The ICR Review coordinator and/or IEG manager sends the response and the final ICR Review to the Region and instructs a designated staff member (not the IEG ICR Reviewer) to post the ICR Review.

POSTING THE ICR REVIEW

Following IEG's disclosure policy, each finalized IEG ICR Review is publicly disclosed on Imagebank. IEG ICR Reviews are searchable there as well as on the IEG website.

REQUESTS FOR MEETINGS DURING AND AFTER THE REVIEW

In some instances, during the review period and occasionally after receiving the final ICR Review, the Global Practice will request a meeting with IEG. Normally, such a request would come with the written comments from the Global Practice. The IEG ICR Reviewer, the Panel reviewer, the IEG Thematic coordinator, and either the ICR Review coordinator or the IEG manager generally attend these meetings. The purpose of the meeting is primarily to listen to the concerns of the operational team about the draft ICR Review. The meeting also provides an opportunity for the IEG team to request clarification of specific points and seek additional information. Before agreeing to a meeting, it is strongly preferred for IEG to already have received the Global Practice's written comments on the ICR Review.
At such a meeting, the IEG team is not expected to explain what course of action or revisions it will implement as a result of the discussion. Further, it is expected that any additional information provided by the Global Practice will be provided in summary form, with the source noted. The Global Practice should not provide reams of documents for IEG to sift through.

Figure 2.1. ICR Review Form for Investment Projects

Independent Evaluation Group (IEG) **Implementation Completion Report (ICR) Review**
IN: Bihar Rural Livelihoods Project (P090764)

Report Number: ICRR0020734

1. Project Data

Project ID	Project Name
P090764	IN: Bihar Rural Livelihoods Project

Country	Practice Area(Lead)	Additional Financing
India	Agriculture	P130546

L/C/TF Number(s)	Closing Date (Original)	Total Project Cost (USD)
IDA-43230,IDA-51230	31-Oct-2012	70,000,000.00
Bank Approval Date	**Closing Date (Actual)**	
14-Jun-2007	31-Oct-2016	

	IBRD/IDA (USD)
Original Commitment	63,000,000.00
Revised Commitment	163,000,000.00
Actual	157,095,054.32

Prepared by	Reviewed by	ICR Review Coordinator	Group
--	---	--	--

2. Project Objectives and Components

a. **Objectives**

--

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

--

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

--

d. **Components**

--

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

--

3. Relevance of Objectives & Design

a. Relevance of Objectives

Rating

---

b. Relevance of Design **{After OPCS changes to the ICR in July 2017, this section (including text and rating) is expected to be dropped from the ICR, and will also be dropped from the ICR Review.}**

Rating

---

4. Achievement of Objectives (Efficacy)

{in this section, the ICR Reviewer creates a sub-section for each objective whose achievement is to be assessed. For purposes of illustration, one objective is shown below, but many ICR Reviews have more objectives.}

Objective 1: {title / wording of this specific objective}

Rationale

Rating

Rating of Overall Efficacy **{After OPCS changes to the ICR in July 2017, we expect to add this rating to ICR Review. Possible ratings will be as the other Efficacy ratings above.}**

{expect to add a set of options either here for characterizing the reasons for overall Efficacy rating of Modest or Negligible}

If this Efficacy rating is Modest or Negligible, is that rating associated with any of the following? {user should be able to select one or more}

☐ Low achievement of this objective, with evidence of the low achievement

☐ Insufficient evidence to show achievement of this objective

☐ External shock (e.g., natural disaster, economic crisis, or force majeure)

☐ Unintended negative effects associated with achievement of this objective (e.g., a road project that achieves improved access to markets but also results in deforestation)

5. Efficiency

Efficiency Rating

---

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		0	0 ☐Not Applicable
ICR Estimate		0	0 ☐Not Applicable

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

6. Outcome

---

a. **Outcome Rating**

---

7. Rationale for Risk to Development Outcome Rating {The title of this field is expected to change based on OPCS changes to the ICR in July 2017. New title will be: Risk to Development Outcome}

---

a. **Risk to Development Outcome Rating {This rating is expected to be removed based on OPCS changes to the ICR in July 2017. The text field is expected to remain.}**

---

8. Assessment of Bank Performance

a. **Quality-at-Entry**

Quality-at-Entry Rating

---

b. **Quality of supervision**

Quality of Supervision Rating

Overall Bank Performance Rating

---

9. Assessment of Borrower Performance {After OPCS changes to the ICR in July 2017, this section (including Government Performance text and rating, Implementing Agency Performance text and rating, and Overall Borrower Performance rating) is expected to be dropped from the ICR, and will also be dropped from the ICR Review.}

a. **Government Performance**

Government Performance Rating

---

b. **Implementing Agency Performance**

Implementing Agency Performance Rating

Overall Borrower Performance Rating

---

10. M&E Design, Implementation, & Utilization

a. **M&E Design**

---

b. **M&E Implementation**

---

c. **M&E Utilization**

M&E Quality Rating

---

11. Other Issues

a. Safeguards

---

b. Fiduciary Compliance

---

c. Unintended impacts (Positive or Negative)

---

d. Other

---

12. Ratings

Ratings	ICR	IEG	Reason for Disagreements/Comment
Outcome	--	---	---
Bank Performance	--	---	---
Quality of M&E	--	---	---
Quality of ICR	--	---	---

13. Lessons

--

14. Assessment Recommended?

---

15. Comments on Quality of ICR

---

a. Quality of ICR Rating

---

3. Project Data (ICRR Section 1)

{need to re-check whether the section below matches the current Form}

The IEG ICRR Portal system automatically pulls in project data from the Operations Portal record for the project. The IEG ICR Reviewer should note any discrepancies between the data appearing automatically and the information presented in the ICR document.

Appraisal amounts: The source of the project costs, loan/credit amounts, and cofinancing amounts at appraisal is the project appraisal document (for investment lending). The project costs include the contribution of the Bank, the government (counterpart funding), and any official cofinaciers.[9] Only the total project costs, the Bank's contribution, and the official cofinancing (as elaborated in Board documents) are mentioned here. The government's contribution to the project is not considered cofinancing and is not recorded in section 1. The costs should be recorded in millions of U.S. dollars.

Cofinancing refers to any arrangement under which Bank funds or guarantees are associated with funds provided by third parties for a particular project or program. The third parties may be official or private. There are two ways of channeling cofinancing:

- *Joint cofinancing:* a joint project in which expenditures from a common list of goods and services are jointly financed in agreed proportions by the Bank and the cofinancier.

- *Parallel cofinancing:* a project in which the Bank and the cofinancier finance different services, goods, or parts of the project.

Actual amounts: Actual total project costs, loan/credit amounts, and cofinancing should be copied from the ICR. If there was additional or supplemental financing, the amount actually disbursed should be included in the total actual project cost; it should not be added to the block on appraisal amounts.

Cofinanciers should include donors other than the Bank that provide *official cofinancing*, as mentioned in the PAD or program document/President's Report, but should not include donors or partners acting as project executors/implementors.

[9] The definition of cofinancing and the types are from OP 14.20 (1995, revised in April 2013) and annex A to OP 14.20. In joint cofinancing, procurement is carried out in accordance with the Bank's procurement and consultant guidelines. In parallel cofinancing, the Bank and cofinanciers finance their different components according to their own rules.

4. Identifying the Objectives (ICRR Section 2)

The World Bank's evaluation architecture - both self-evaluation (reflected in the ICR and in supervision reports) and independent evaluation (IEG's assessments) - is objectives-based. All the elements of the project outcome rating are linked to the objectives: the relevance of the objectives, whether the objectives were achieved (efficacy), and whether they were achieved efficiently. Accurately identifying the objectives is thus essential to the entire evaluation exercise and is critical for assuring accountability.

This section explains where the objectives can be found and the guidelines for interpreting them for the purposes of the ICR review.

The project's objectives include the statement of objectives, as articulated in the loan/credit/grant agreement, and key associated outcome targets, if any. The statement of objectives should be lifted directly from the lending agreement (for investment lending) - not from the ICR.

See also the discussion below on identifying the objective by parsing the PDO statement.

What constitutes a project's objectives?

A project's objective is a statement of what it intends to achieve, expressed in terms of an intermediate or final development outcome, as opposed to a financed deliverable (output). In its guidelines for the content of the PAD, Operations Policy and Country Services (OPCS) recommends that the project's development objective(s) should: "(a) be stated as concisely as possible; (b) indicate the primary target group(s) and the change/response expected from this primary target group as a result of project interventions; and (c) focus on outcomes for which the project can reasonably be held accountable. It should neither encompass higher-level objectives beyond the purview of the project, nor be a restatement of the project's components or outputs."[10] {This paragraph to be updated with the revised PAD guidelines}.

---

[10] World Bank, OPCS, n. d.; World Bank, OPSPQ 2013, p. 3. The OPCS/IEG Harmonized Criteria (World Bank, OPCS/IEG 2006) note that "For evaluation purposes, an operation's objectives encompass both the project development objectives (PDOs) stated in Board documents and key associated outcome targets. This means that whenever the PDOs stated in the Board documents are so broad and/or vaguely worded as to preclude any meaningful evaluation, intended objectives are inferred by the evaluator from key associated outcome targets (and/or the operation's design features as relevant)" (p. 1). However, it is only in the exceptional case of very poorly articulated objectives that the key outcome targets can be used to infer objectives, and in these instances the ICR Review coordinator should be consulted. (Key associated outcome targets refer to measurable or observable outcomes expected by completion [in terms of types of benefits or progress expected for primary target groups], as well as any indications of their scale and scope [which are normally captured in key indicators in the PAD/program document]).

Where to find the objectives?

The IEG ICR Reviewer should always assess the project development objective stated in the original and revised legal documents, and not take as given the PDO statement in the ICR, which is being assessed.

For **investment projects**, take the PDO as stated in the lending agreement, development credit agreement, or grant agreement (the legally binding document negotiated between the Bank and the government), in schedule 2, at the end of the agreement, entitled "Project Description."

If needed, additional information on the PDO can usually be found in the PAD. Such information may be found in the front matter/summary; in the section on "Project Development Objectives;" and in the technical appendix, "Detailed Project Description." However, if the wording of the PDO diverges from that in the legal agreement, it is important to take the wording as in the legal agreement. {need to add 'Series of Projects' and Multiphase Programmatic Approach}

Assessing Global Environmental Objectives

Projects wholly or partly financed by the Global Environment Facility (GEF) will likely include Global Environmental Objectives (GEO) in the project appraisal document, in addition to PDOs. Both the PDOs and the GEOs from the project appraisal document should be listed on the ICR Review form, in addition to the objectives noted in the grant/legal agreement. However, the project is assessed based on the wording in the grant/legal agreement. {Check with CN whether this is still relevant; if no longer relevant, delete this paragraph}

5. Project Components, Cost, and Dates (ICRR Section 2)

Components

In this section, provide a summary description of components (matching the PAD/lending agreement, not the ICR), with sufficient detail to make clear what activities were supported by project funds. No evaluation is needed here - only description - but any discrepancies in the description of components across the PAD/credit/loan agreement and ICR should be noted. List each component separately, followed (in parentheses) by both the appraisal and actual costs for that component in US$ millions. The estimated and actual component costs should add up to the estimated and actual total project costs, respectively.[11] If they do not, the reason for the difference should be noted - for example, if the estimates at appraisal exclude contingencies, but the actual costs do not.

Comments on Project Cost, Financing, Borrower Contribution, and Dates:

In this section, record the following information:

For all projects:

- The number of extensions to the project closing date and reasons for extensions
- In the case of formally restructured projects, the proportion that was disbursed before and after the revision in order to establish what weight to give to revised objectives/targets in the overall outcome rating (see section 5 {check section number}).

For investment projects:

- The reasons for changes in the share of component costs financed by the Bank and for reallocation of financing across components
- The share of beneficiary (government and community) contributions, both as estimated at appraisal and actual amounts.

6. Relevance of Objectives (ICRR Section 3)

Relevance of objectives, along with efficacy and efficiency, is one of the three criteria underpinning the outcome rating.

Relevance of objectives

DEFINITION

Relevance of objectives is the extent to which an operation's objectives are consistent with *current Bank country strategies* (expressed in CPFs). For the ICR, "current" refers to the time of project closing.

CRITERIA

Relevance of objectives is assessed with respect to the country strategy agreed between the Bank and the country. Assessment of relevance of objectives is based on the following criteria:

[11] The sum of the estimated costs of the components may differ from the total project cost because the component-costing often excludes contingencies. If there are two versions of component costs in the PAD, with and without contingencies, then the one with contingencies should be used. Otherwise, the IEG ICR Reviewer should note that the appraisal estimates for the components exclude contingencies.

- Alignment with strategy: Look for an explanation to answer the question, "How (and to what extent) does the objective aligns with the country strategy in effect at project closing. This explanation should provide details on the strategic nature of the alignment, that is,
  - what development problem the project sought to address,
  - where this development problem fit within area of activity where the Bank would contribute, and
  - how this problem was or is being specifically addressed through this project.
- Country context: Look for an explanation to answer the question, "Is the objective outcome-oriented and appropriately pitched for development status and capacity in the country as described in the CPF?" This judgment would incorporate government capacity, fragility issues, and what is a reasonable expectation for the objective given the possible wide range of constraints in the operational context.
- Previous sector experience: Look for an explanation to answer the question, "What is the historical experience of the Bank in the relevant country and sector?" The expectation is that a project occurring later in the Bank's engagement with the sector in that country would set more relatively more challenging objectives than a project early in the engagement. This means that second, third, fourth, fifth, etc., project within a sector in a country should include outcomes consistent with progress over time as compared with earlier projects.

GUIDELINES

Explanations of the areas mentioned above are likely to appear in the ICR section "Key Factors that Affected Implementation and Outcome" as well as in the sections on Relevance of PDOs and Efficacy, but any other section may also include relevant explanations.

The assessment of the relevance of objectives ensures, with respect to both accountability and lesson-learning, that the ICR takes into account whether the Bank's implementation assistance was responsive to changing needs and that the project remained important to achieving country, Bank, and global development objectives, all of which may change over time. If country circumstances have changed significantly during implementation, the ICR should explain whether and how these changes were accommodated (through changing of the objectives through formal restructuring or other means) to retain the relevance of the objectives.

Occasionally, the CPF at closing does not refer to the objectives of a particular project because the project achieved its objectives fully and the Bank's strategy shifted elsewhere. In these cases, the ICR Review needs to assess the extent to which the *outcomes* (that were embedded in the project's objectives and were achieved) are relevant to the CPF at project closing. For example, if a project aimed at eliminating marine pollution succeeded in fully achieving its objectives, and the CPF at project closing thus contained no reference to marine pollution, but instead emphasized land-based pollution, the project could still be rated favorably on relevance of objectives. The reasoning and evidence supporting this assessment must be provided in the ICR Review, however.

If the stated objectives are vague or not sufficiently monitorable, a relatively low rating can be appropriate for the Relevance of Objectives, as well as for Quality at Entry.

RATING OF RELEVANCE OF OBJECTIVES

Relevance of objectives is rated on a 4-point scale: *high, substantial, modest,* or *negligible:*

High	High relevance to country strategy: Full alignment between project objectives and country strategy
Substantial	Substantial relevance to country strategy: Almost full alignment between project objectives and country strategy, or minor misalignments in limited areas
Modest	Modest relevance to country strategy: Partial alignment between project objectives and country strategy
Negligible	Negligible relevance to country strategy: Very little alignment between project objectives and country strategy

7. Achievement of the Objectives (Efficacy, ICRR Section 4)

Definition

Efficacy is defined as the extent to which the project's objectives were achieved, or are expected to be achieved, taking into account their relative importance, and are attributable to the activities or actions supported by the project.
For the purposes of this section, the objectives refer to each of the key outcomes indicated in the statement of PDOs from the legal agreement (credit/lending/grant agreement) in the case of investment projects section.

Criteria

The achievement of each objective is assessed based on the level of achievement and the concept of "plausible causality." To establish this, for each objective, the IEG ICR Reviewer should:

(a) For each objective or outcome, assemble and succinctly present the evidence from the ICR for each part of the results chain or causal chain supported by the project—the inputs and outputs—and the observed intermediate outcomes or impacts for each objective. Explain your interpretation of the extent to which evidence presented in the ICR supports the conclusion that the causal relationships asserted in the results chain are true, or likely, or, at least, plausible. Also, assess the extent to which the ICR presents evidence that it was at least plausible that the outcomes achieved arose from the activities and outputs of the project, as distinct from other non-project factors – such as other interventions, policy changes unrelated to the project, natural events, or market factors. This type of analysis helps in making a reasonable assessment of plausible attribution of observed outcomes to the project's specific interventions. Note that If relevant outputs were not produced, or if relevant intermediate outcomes were not achieved, then any achieved outcomes cannot plausibly be attributed to the project.

(b) Discuss and (to the extent feasible) present evidence from the ICR of the contribution of other, non-project factors leading to these outcomes (**the counterfactual**), with the intent of examining whether the achieved outcomes can plausibly be attributed to the government program or project supported by the Bank.

General Principles for Assessment of Efficacy

Efficacy is to be assessed at the time of evaluation (that is, at the time of the ICR, the ICR Review, or the PPAR) (as are the other two constituent elements of outcome, i.e. Relevance of Objectives and Efficiency). For example, if a flood (or other natural disaster) had wiped out all project achievements *after* project closing such that at the time of evaluation there was nothing to be seen on the ground, Efficacy would not be rated favorably. (Bank performance could, of course, be rated in the satisfactory range depending on whether the Bank had done everything possible to avoid the unfavorable outcome.)

Each stated objective is to be rated, even if the objective is stated in output terms. If an objective is stated in output terms, however, this can be an example of "setting the bar" too low. If the IEG ICR Reviewer can make the case that a higher "bar", or promising intermediate or final outcomes (rather than outputs) would have been possible and desirable in the particular country circumstances, then this would provide support for a relatively low rating of Relevance of Objectives. {Question: if there are several objectives and some are at outcome level and some at output level, then would we not only rate the outcome objectives? Maybe we should clarify that output objectives are subject to rating only when no outcome-level objectives were identified?
On the other hand, recent practice has allowed 'access' to be rated separately from quality or e.g. health outcomes, so maybe that contradicts the above?}

Organizing the assessment of efficacy

In section 4 of the online ICR Review form, the IEG ICR Reviewer adds a section for each of the outcomes (objectives) that comprise the statement of the objectives being assessed. (The ICR Review form, by default, contains a section labeled "Objective 1" for the first objective. If a project has several objectives, the IEG ICR Reviewer can click on "Add Objective" and this will add a section labeled "Objective 2." Clicking "Add Objective" again will add "Objective 3" and so on.)

In the Objective header field, the IEG ICR Reviewer enters the text of the objective, as stated within the PDO.

In the Objective Rationale field, the IEG ICR Reviewer should:

(a) assemble and succinctly present the evidence from the ICR (or field work in the case of a Project Performance Assessment Report) that documents the realization of the complete results chain, from outputs to intermediate outcomes to final outcomes; and

(b) comment on the extent to which the outcomes can be attributed to the project or program in question. These two elements are discussed below, in turn. A rating will be assigned to each objective.

Clarify the specific objectives by parsing the PDO statement

Organize the assessment of PDOs around each objective or outcome captured in the statement of objectives. Compound PDOs with multiple outcomes should be "unpacked" and treated separately. An example of a compound objective is when several outcomes are linked together in a single sentence. Below are several examples of how to unpack the objectives/outcomes included in a single PDO statement.

Some PDO statements articulate objectives (i.e. expected outcomes) but also include components or activities or outputs contributing to those objectives, usually following statements such as "by means of," "through," or "by." In this case, the element whose achievement should be assessed is only the expected outcomes. Components, activities, and outputs should be factored into the results chain analysis – in other words, they may help to demonstrate the causal (or, at least, plausibly causal) relationship between the project's interventions and achieved outcomes. If the phrase "in order to" is used, however, then what comes <u>after</u> that phrase usually constitutes the main objective(s) whose achievement is to be assessed.

If the PDO statement expresses a goal to support a government's Program, the objectives of that Program (normally found in the PAD) should be used. If the project supports a subset of the government's Program objectives, the assessment of efficacy should include only the Program objectives specifically supported by the project.

PARSING THE PDO STATEMENT, EXAMPLE 1:

Here is the PDO Statement of a rural poverty reduction project:

> "To improve access to small-scale socio-economic infrastructure and services, raising incomes through investment in productive activities, and strengthening the capacity of Municipal Councils and Community Associations to raise funding and harmonize policies and institutional arrangements for delivery of public investments intended to benefit the rural poor."

In this example, assess efficacy separately for each of the project's three objectives (or outcomes):

> Objective 1: Improving access to small-scale socio-economic infrastructure and services;
>
> Objective 2: Raising incomes through investments in productive activities;
>
> Objective 3: Strengthening the capacity of Municipal Councils and Community Associations to raise funding and harmonize policies and institutional arrangements for delivery of public investments intended to benefit the rural poor.

PARSING THE PDO STATEMENT, EXAMPLE 2:

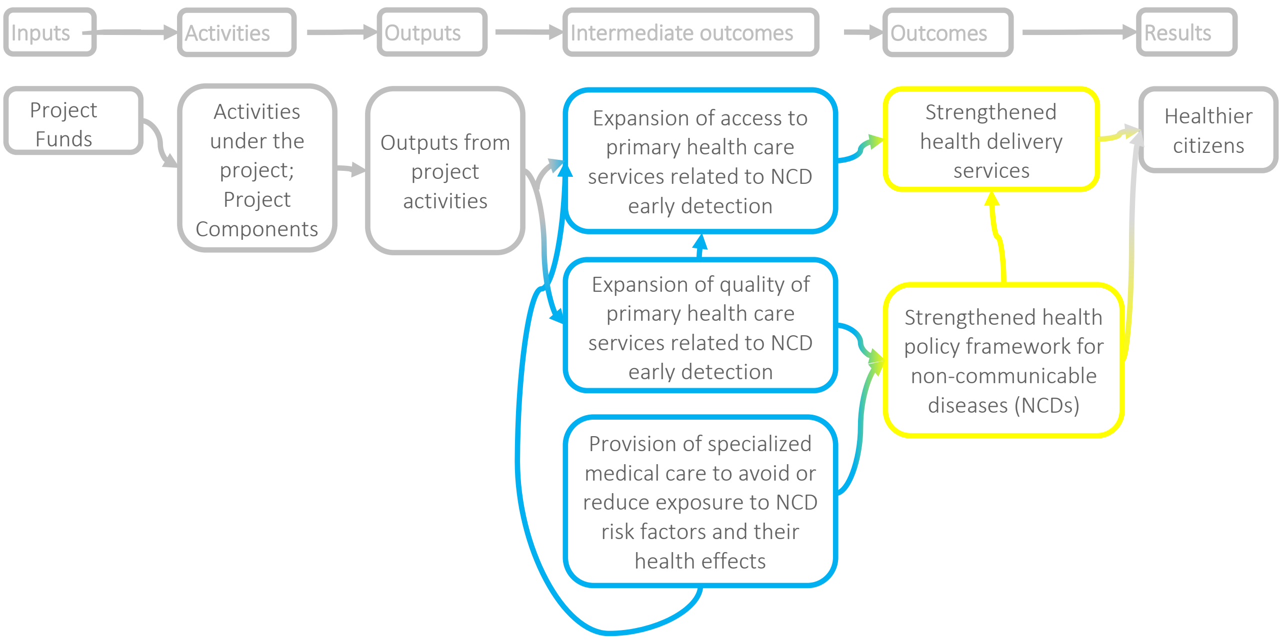
Here is the PDO statement of a health services project:

> "To support the borrower's efforts to further strengthen its health delivery services and the current health policy framework for NCDs [non-communicable diseases] through
> - The expansion of access and the quality of primary health care services related to NCD early detection; and
> - The provision of specialized medical care to avoid or reduce exposure to NCD risk factors and their health effects"

In this example, assess efficacy separately for the two elements highlighted in yellow (which come before the "through" phrase). These constitute the main objectives to rate in the Efficacy section (highlighted in yellow):

> Objective 1: Strengthened health delivery services
>
> Objective 2: Strengthened health policy framework for NCDs

The elements highlighted in blue above are considered intermediate outcomes. Below is a sketch of the results chain suggested by the PDO statement above:



PARSING THE PDO STATEMENT, EXAMPLE 3:

Here is the PDO statement of the India Third Technical Education Project:

> The objective of the project is to assist the industrially and economically underdeveloped, and geographically remote states of the northeastern region (Arunachal Pradesh, Meghalaya, Mizoram, Nagaland, Sikkim, Tripura), Jammu & Kashmir, and the Union Territory of Andaman & Nicobar Islands to expand capacity and improve the quality and efficiency of technician (polytechnic) education to meet the specific economic needs of each state. The project also aims at increasing access of some disadvantaged sections of society (women, scheduled tribes and rural youth) to technician education and training.

The IEG ICR Reviewer then noted that she would organize the discussion of achievement of the objectives in section 4 of the ICR Review around the following technical education outcomes:

> Objective 1: expand the capacity of technical education
>
> Objective 2: improve the quality of technical education
>
> Objective 3: improve the efficiency of technical education
>
> Objective 4: meet the economic needs of the industrially and economically underdeveloped, and geographically remote states of the northeastern region (Arunachal Pradesh, Meghalaya, Mizoram, Nagaland, Sikkim, Tripura), Jammu & Kashmir, and the Union Territory of Andaman & Nicobar Islands
>
> Objective 5: increase access of some disadvantaged sections of society (women, scheduled tribes and rural youth) to technician education and training.

Components, activities, and project outputs should not be included as objectives, even if they appear in the same sentence as the outcome-level objectives. The PDO statement of the Kyrgyz Agricultural Support Services Project incorporates both multiple outcomes and a listing of components in a single objective statement: The statement of objectives was:

> To improve the incentive framework for, and productivity, profitability, and sustainability of Kyrgyz agriculture by means of: assisting the government in implementing land and agrarian reforms; providing emerging private farms with advisory and development services; developing the seed industry; establishing a legal framework, organizations, and procedures for crop protection and plant quarantine; establishing an agricultural market information system; and enhancing institutional capacity of the Ministry of Agriculture and Water Resources. [emphasis added]

Up to the phrase "by means of," there are four outcomes that the project seeks to affect, listed below. These four are, in effect, the main outcomes that the project sought to achieve and would be the main headings in section 4 of the ICR Review form, on achievement of objectives (see below). All of the activities *after* the words "by means of" are the project's components, which are activities and outputs. Thus, they should not be considered as objectives, but rather as outputs within the project's results chain leading to the three main outcomes.

> Objective 1: Improve incentive framework for Kyrgyz agriculture
>
> Objective 2: Improve productivity of Kyrgyz agriculture
>
> Objective 3: Improve profitability of Kyrgyz agriculture
>
> Objective 4: Improve sustainability of Kyrgyz agriculture

Establishing a results chain or theory of change that links to each objective

The results chain or theory of change for each objective is the logic, expressed in project design documents and in the ICR, that links the project's inputs and outputs to the desired outcome. A useful conceptual definition of a results chain is found in IEG 2012:

> *A results [chain] is an explicit articulation (graphic display, matrix, or summary) of the different levels, or chains, of results expected from a particular intervention—project, program, or development strategy. The results specified typically comprise the longer-term objectives (often referred to as "outcomes" or "impact") and the intermediate outcomes and outputs that precede, and lead to, those desired longer-term objectives. Although the World Bank has used the term "results [chain], similar conceptual tools, also designed to organize information regarding intended outcomes and results, are used across different agencies: logical frameworks, logic models, theories of change, results chains, and outcome mapping. Thus, the results [chain] captures the essential*

elements of the logical and expected cause-effect relationships among inputs, outputs, intermediate results or outcomes, and impact…

Defining cause-effect linkages for one or more interventions lays the groundwork for a results [chain]. Thus, the development of a good results framework requires clarity with respect to the theory of change—the reasons why the project, program, or strategy will lead to the outputs; why those outputs are likely to lead to the immediate or intermediate outcomes; and how those outcomes are (at least hypothetically) linked with longer-term outcomes or impact. The theory of change also requires knowing or estimating how long it will take to achieve each stage of the program and how much of the outcome is likely to be achieved. Thus, defining cause-effect linkages for one or more development interventions lays the groundwork for a results framework….

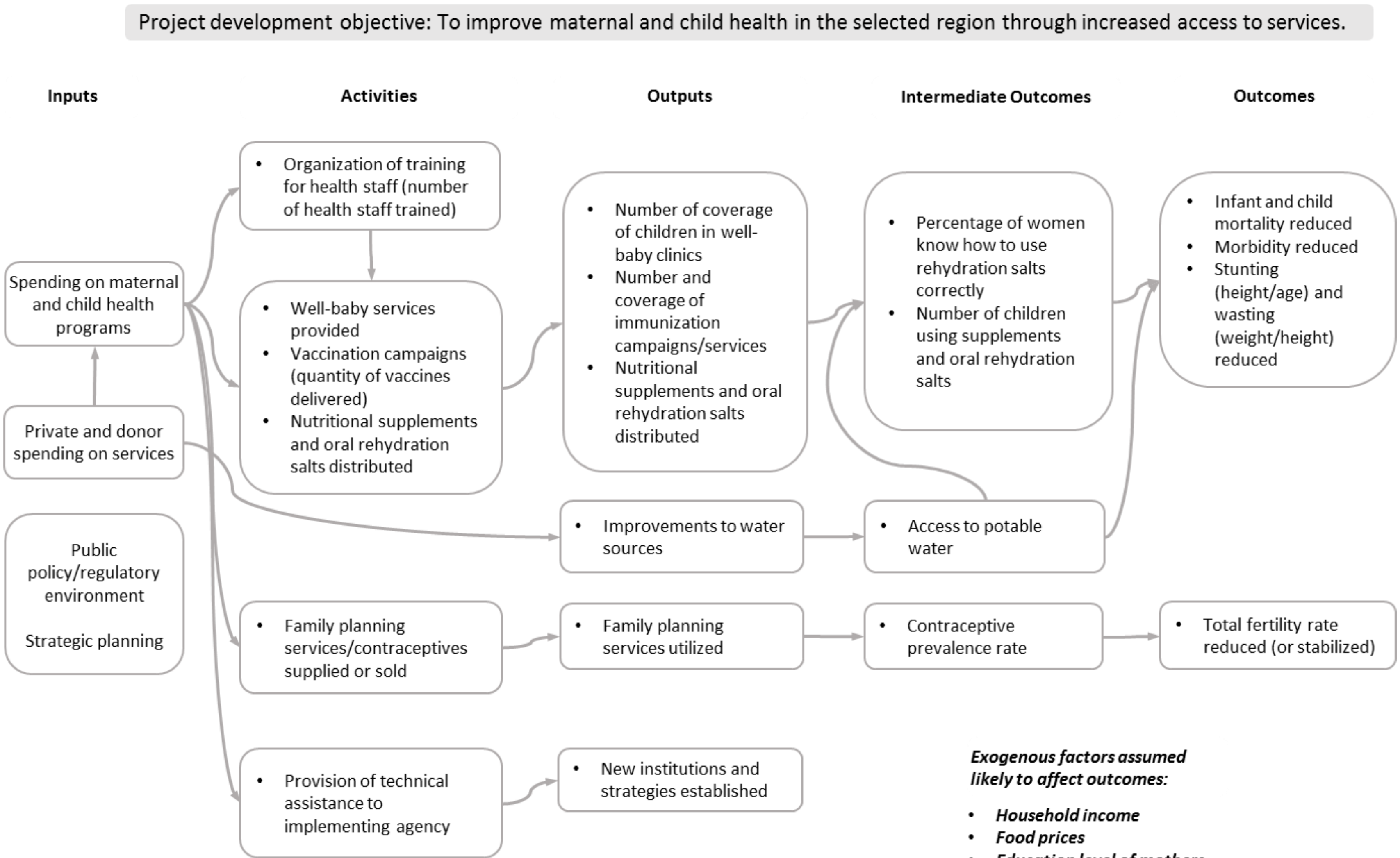
A results framework also often identifies any underlying critical assumptions that must be in place for the intervention to be successful, that is, to lead to achieving the targeted outcomes and impacts.

Source: World Bank Independent Evaluation Group, 2012, "Designing a Results Framework for Achieving Results: A How-To Guide"

Figure 7.1 below illustrates the links in the results chain for a program to improve child health among the poor. Examples of evidence presented in the ICR can include monitoring data that compares observations made in regions or populations where the project was active with observations made in similar regions or populations where the project was not active; data that compares observations made before the project (baseline) with observations made later (a pre-post design), and so on. Exogenous factors (such as, in this case, household income, food prices, and education level of the mother) can also be important determinants of project outcomes and should be taken into account in the discussion of the counterfactual. An ICR can also usefully present impact evaluation and related research to verify the links in the results chain underlying the logic of the project design.

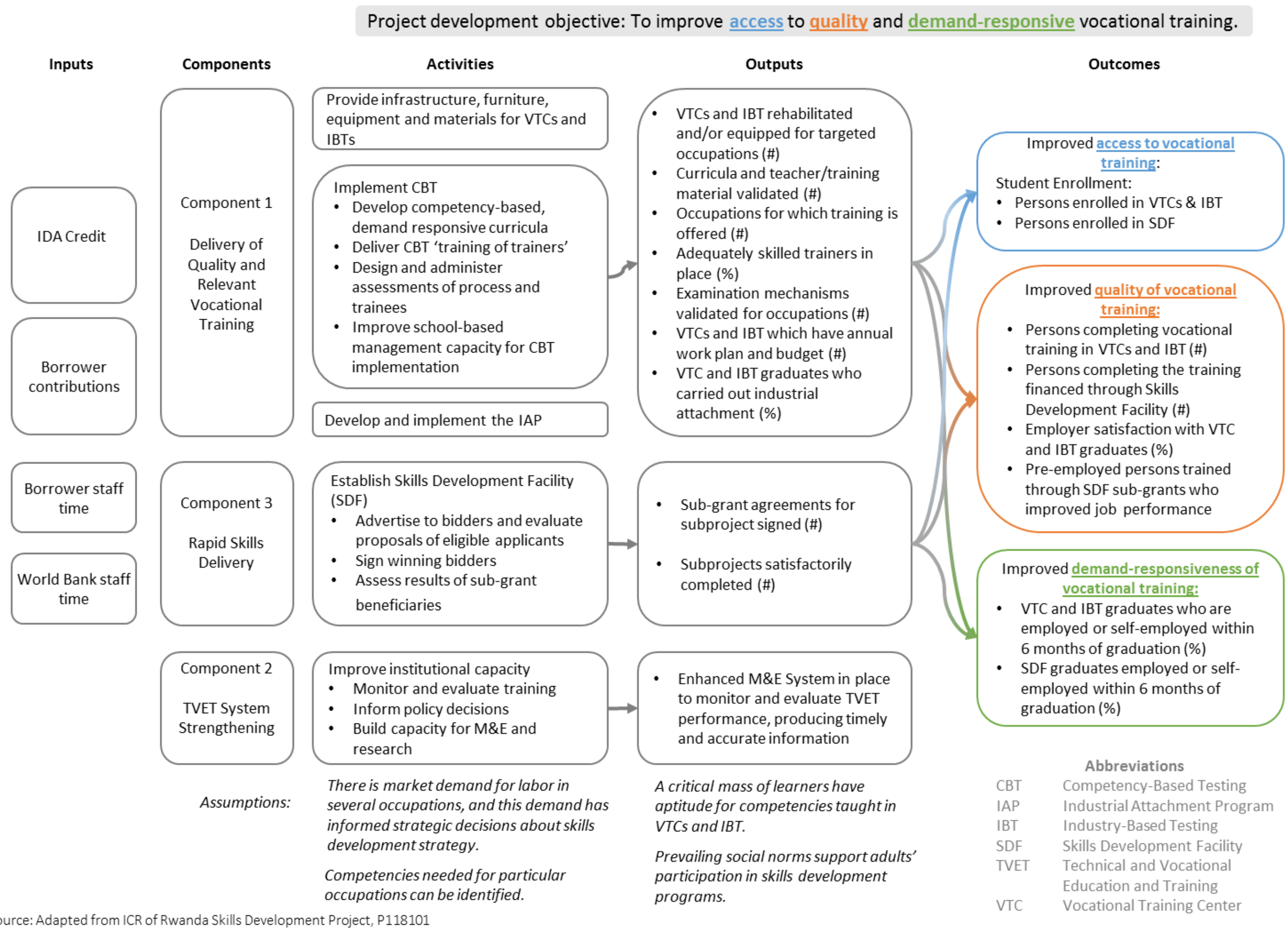
Figure 7.2 below illustrates a results chain for a skills development project. This example, illustrate that the main objectives as defined in the PDO statement are placed at the level of outcomes in the results chain.

Figure 7.1. Results Chain for a Maternal and Child Health Project



Source: Adapted from IEG 2004

Figure 7.2. Results Chain for a Skills Development Project



Source: Adapted from ICR of Rwanda Skills Development Project, P118101.

Discussion of validity of indicators

As part of the discussion of efficacy, the ICR should contain an assessment of the validity of the PDO-level indicators in the results framework, and the ICR Review should comment on the validity of the PDO-level indicators used to show the extent of achievement of objectives. If the indicators defined in results framework had shortcomings (for example, in validity, measurability, timeliness, etc.), check whether the ICR included other sources of information that would speak to achievement of project objectives or outcomes.

Even in cases where indicators defined in the results framework were excellent for assessing the outcomes, the ICR may include additional data and evidence that speaks to achievements. Use of multiple sources of information helps with “triangulation” – when different sources point to the same types of achievement, this convergence provides a richer, more accurate assessment of achievement of objectives, and, if it is well done, the ICR Review may highlight this as a commendable practice. An important element of complementary data and evidence is the perspective of beneficiaries.

Discussion of attribution and the counterfactual

The **counterfactual** is defined as what would have happened in the absence of the government intervention, project, or program supported by the Bank. Establishing the evidence for the elements of the results chain for each outcome is a necessary but not sufficient condition for attributing the outcomes to the project. In most cases, other factors beyond the scope of the project also affect these outcomes, contributing or detracting from them. These factors might include the influence of weather or rainfall, economic crises, natural disasters, favorable or unfavorable international prices for farmers’ production, other government policies outside the project, or the activities of other donors.

Programs supported by the Bank are rarely subjected to an evaluation design, such as a randomized experiment, capable of contrasting results “with” and “without” an intervention or program. In some cases, rigorous impact evaluations can be conducted on specific parts or interventions within a program; the results from these studies can be useful in understanding the counterfactual, at least in determining what parts of a program “worked.”

However, most of the projects or programs supported by the Bank involve large-scale and multifaceted interventions or country or sector-wide policies for which it would be difficult or impossible to establish an airtight counterfactual as the basis for attributing outcomes to the project.

To assess efficacy, for each objective, the IEG ICR Reviewer should nevertheless identify and discuss the key factors outside the project that might have contributed to or detracted from the outcomes, as well as any evidence (from the ICR) for the actual influence of these factors.[12]

The following types of information—in addition to evidence from the results chain—have been found useful in assessing evidence on the extent to which the outcomes achieved can plausibly be attributed to the project or program:

- A timeline of key events, showing the relationship between project activities, events beyond the project, and changes in outcomes flagged by the objectives.
- Evidence of trends in the outcomes before, during, and after the project or program.
- Evidence of trends in outcomes in project and non-project areas, taking into account the ways the two areas may differ in baseline characteristics and other factors that may be affecting the two areas.
- Trends in other factors that plausibly could have influenced the outcomes independently of the project—such as weather, natural disasters, economic trends, other government policies, the activities of other donors.

The importance of attention to the idea of the counterfactual is illustrated by this scenario:

> A project aimed to improve access to health services in 3 rural provinces. It planned to assess achievement by measuring the number and percentage of people in those provinces who had received essential health, nutrition, and population (HNP) services (disaggregated, of course, by gender and poverty status), and collected baseline information accordingly. At the end, instead of observing the expected increase (in the number and percentage of people in the targeted provinces who had received essential HNP services), the project observed a decrease. In the absence of any information about the counterfactual, a likely conclusion based on this observation would be that the project failed to achieve the objective.
>
> Imagine, however, that in fact a shift in policy (or another external factor) had led to a country-wide decrease in access to HNP services. If project designers paid attention to the counterfactual, they would have collected baseline data not only for the 3 rural provinces targeted, but for all provinces. If project implementers and the ICR author had paid attention to the counterfactual, they would have looked for data at the end of the project not only for the 3 targeted provinces, but also for the remainder of provinces. Imagine, further, that comparison of the data showed that although a decline in access to health services occurred in the targeted provinces, the decline in provinces not served by the project was significantly larger. This observation would then provide support for the assessment that, although the planned objectives were not achieved, it

[12] In conducting Project Performance Assessments, the evaluator should plan to search for evidence of the influence of non-project factors, such as economic trends, other government policies, other donor support, and exogenous factors that may also be affecting the anticipated outcomes.

is plausible that the project had a positive effect, i.e. that project activities mitigated against the adverse effects of the external factor that drove access down more broadly.

For all ratings, the IEG ICR Reviewer needs to assess whether the outcomes achieved are attributable to the project. When the expected outcome is achieved but the evidence is not presented to show how the activities and outputs of the project have led to the outcomes, then the rating of efficacy should be lower than would be the case with the same level of expected outcome accompanied by evidence of how the project activities and outputs led to the outcomes. Similarly, when evidence indicates that expected outcomes achieved occurred due to factors other than project activities and outputs, the rating of efficacy should be lower than would be the case with similar outcomes accompanied by evidence of how the project activities and outputs led to the outcomes.

It is important to note that the efficacy rating reflects the incremental contribution of the project or program to observed outcomes, regardless of whether the observed outcomes moved in the "right" or "wrong" direction. For example:

- If the expected outcome target was met or exceeded, but there is evidence that the change was due mainly (or solely) to external factors, an efficacy rating of modest (or negligible) may be warranted.

- If the outcome deteriorated, falling short of the target, but there is evidence that the decline would have been even worse in the absence of the project, a rating of substantial (or high) could be warranted.

To justify these judgments, a high standard of evidence is expected. For example, it is insufficient for the ICR to claim that the project fell short of achieving its objective because of macroeconomic conditions without strong evidence that these conditions were responsible for the trend in the outcome indicator. The burden of proof is on the ICR itself to show that improved outcomes were the result of the project, and that declining outcomes were not.

How to treat overarching objectives and objectives across projects

{this paragraph needs to be revised based on guidance for SOP and Multiphase Program Approach; also check with Soniya & ICRR Coordinators} When there are both overarching and specific objectives, the ICR should report both. This is very likely to be the case for a series of projects, which constitute a series of investment projects (phases) offered over a medium to long-term period, with objectives for each phase and for the overall program. The ICR, and the ICR Review, should assess achievement of the overarching objectives of the overall program, as well as the achievement of the specific objectives of each completed phase and their contribution to the overall program outcome. Regarding the overarching objective, the ICR Review should also comment on the likelihood that the overarching objective will be achieved in the future, and should cite the reasoning and information on which this assessment is based. The overarching development objective should not be rated, however. (In the electronic ICR Review form, open an Objective section for the overarching objective, and use the Rationale field to comment on it. Then, in the Rating field, select "Not Rated/Not Applicable.")

How to assess Efficacy in the case of a pilot project

In assessing pilot projects, M&E assumes particular significance. The nature of a pilot project means that it tests a new approach on a small scale, and that results of that test inform the decision of whether to implement the new approach on a larger scale or not. Testing an approach implies producing information about whether and how it works. The special importance of performance information for pilot projects thus means that pilot projects are held accountable for rigorously monitoring and evaluating the approach they were testing and distilling lessons. In such cases, the ICR Review should emphasize any weaknesses in M&E (weighting them more heavily than might be the case in non-pilot projects), and should factor M&E weaknesses into both the Outcome rating and the Bank Performance rating. If the PDO of the pilot was phrased in terms of learning from the findings of the project, and if pilot project rigorously tested the approach, the pilot project may be rated favorably on efficacy (in achieving "piloting"), even if the approach being tested did not yield positive outcomes. {Possibly to be revised based on further discussion with OPCS?}

Other considerations in assessment of Efficacy

For evaluation purposes, a project's objectives encompass both the PDOs stated in Board documents and key associated outcome targets. This means that whenever the PDOs stated in the Board documents are so broad and/or vaguely worded as to preclude any meaningful evaluation, intended objectives are inferred by the IEG ICR Reviewer from key associated outcome targets. If the project lacks key associated outcome targets for some or all aspects of the objectives, then the description of the project's intent provided in the paras that follow the objectives statement can also be used.

If objectives are not yet achieved at the time of project closing, the ICR and/or the ICR review can still make a case that they are "likely to be achieved," but the IEG ICR Reviewer would need to look for, and present, convincing evidence of the likelihood of such achievement through a strong results chain. If sufficient evidence does not exist, no leaps of faith should be made. {MK flags this as needing further explanation; what kind of evidence?}

A downgrade in the Efficacy rating is warranted in both of the following cases: (i) when there is insufficient evidence of impact; or (ii) there is evidence of insufficient impact.

Rating of Efficacy

In the ICR Review, the efficacy of each objective (intended outcome) is rated on a 4-point scale: *high, substantial, modest, negligible*. The ICR should include a discussion of efficacy of each objective, but is not required to contain a rating of efficacy for each objective. Both in the ICR Review and the ICR, an overall efficacy rating is given. The ratings are defined as follows:

- **High:** The project fully achieved or exceeded its objectives (or intended outcomes), or is likely to do so.
- **Substantial**: The project almost fully achieved its objective (or intended outcomes), or is likely to do so.
- **Modest**: The project partly achieved its objectives (or intended outcomes), or is likely to do so.
- **Negligible**: The project barely achieved its objectives (or intended outcomes), or is likely to do so.

Arriving at the Overall Efficacy Rating in Projects with Multiple Objectives.

In projects with multiple objectives, IEG's ICR Review rates the achievement or likely achievement of each individual objective, and also provides a single overall Efficacy rating covering all objectives. The Bank's ICR should include a discussion of achievement of each individual objective, but will provide only a single overall Efficacy rating covering all the objectives. Both the World Bank's ICR and IEG's ICR Review use the same approach for arriving at the overall Efficacy rating, and also use a harmonized rating scale: High, Substantial, Modest, and Negligible.

If the Legal Agreement indicates differences in relative importance (or "weight") of objectives (for example, stating a "primary" and "secondary" objective), the World Bank's ICR and IEG's ICR Review should apply the Legal Agreement's relative importance (or "weights") in assessing overall Efficacy. If the relative importance or weights of the different objectives were not stated in the Legal Agreement, but were explicitly stated in the PAD (e.g., under the Project Description, Theory of Change, or Results Chain), those relative weights could be factored in at project completion in arriving at the overall Efficacy rating. {to be clarified further: should IEG ICR Reviewer expect to see this in the ICR? Or PAD?}
If neither the Legal Agreement nor the PAD indicates indicate the relative importance or weight of objectives, then the World Bank's ICR and IEG's ICR Review assume equal importance (or equal weight) for each of the objectives.

In any case, the amount of resources allocated to an objective (even if it were possible to determine this amount) should not be used to impute the relative importance (or weight) of that objective.

As a rule of thumb, a High or Substantial Efficacy rating is warranted when all three of the following are true:

(i) there is sufficient evidence of outcomes/impact;
(ii) there is evidence of sufficient outcomes/impact; and
(iii) the observed outcomes/impact can be attributed to the project interventions/activities.

If a majority of the project objectives are rated High or Substantial, the Overall Efficacy rating will tend to be High or Substantial. If a majority of the project objectives are rated Modest or Negligible, the Overall Efficacy rating will tend to be Modest or Negligible.

If half the objectives are rated High or Substantial and half Modest or Negligible, then there is the possibility to either round-up or round-down based on the strength of the evidence. Rounding-up would be justified in the case of 'strong' Highs or Substantials (i.e., they are at the top of their range, so that, for example, the Substantial is nearly High). Rounding-down would be justified in the case of 'weak' Highs or Substantials (i.e., they are at the bottom of their range, so that, for example, the Substantial is barely so). {add further explanation of the idea of a range, and rounding up or down}

The overall Efficacy rating derived based on the guidance above should then be cross-checked against 'best judgment,' in other words, stepping back and asking the questions "To what extent did the project achieve the objectives promised?" and "Are the shortcomings in achievement of the objective absent, minor, moderate, significant, severe, or major?" The idea of 'Best Judgment' is intended to lead to iteration back to the specific objectives, evidence, and achievement demonstrated, to bring out information that supports (or does not support) the characterization of achievement as a whole as High, Substantial, Modest, or Negligible. 'Best judgment' is used to amplify and provide support for rating Efficacy within the context of the guidelines. Note that 'best judgment' in itself cannot be used as a reason to diverge from the guidelines. In every case, the ICR should include a specific and logical explanation of the rationale for the rating and specific presentation of the evidence underpinning the rating, particularly if any rounding-up or rounding-down is done. {add further explanation or definition of rounding up or rounding down}

Both of the above approaches should be used to guide and arrive at the final Overall Efficacy rating.

8. Efficiency (ICRR Section 5)

Definition

Efficiency is a measure of how economically resources and inputs are converted to results (OECD/DAC 2002). For a development project, the central question is whether the costs involved in achieving project objectives were reasonable in comparison with both the benefits and with recognized norms (value for money) (World Bank, OPCS 2011a, p. 24). Was the project implemented at least cost?

Guidelines

This section should report on all available measures of efficiency both ex ante and ex post and highlight any data gaps and methodological weaknesses in the Bank's assessment of efficiency. The analysis should discuss both the traditional measures of efficiency (as applicable and practical)—for example, net present value, economic rate of return, cost effectiveness, unit rate norms, service standards, least cost analysis and comparisons, and financial rate of return—and aspects of design and implementation that either contributed to or reduced efficiency. The ICR should also analyze the project's efficiency using any other appropriate cost-effectiveness criteria to determine whether the project represented the expected least-cost solution to attain identified and measurable benefits by an analysis of either cost per unit of input or cost per unit of output.

Examples of implementation aspects that may reduce efficiency include delays in implementation, financial management issues, procurement issues (including procurement delays), and cost overruns.[13] {check} The ICR should also indicate the components and the percentage of total project costs covered by any such analyses (noting any differences from the analyses at appraisal).

Implementation efficiency may be affected by:

- The complexity of the project/program and its organizational arrangements;
- The commitment demonstrated by government, its agencies, and other participants to the objectives of the project;
- Whether risks were identified and their mitigation adequate
- The adequacy of participatory processes
- Unforeseen security and natural events

[13] World Bank, OPCS 2011a, p. 25 (emphasis added). Using this guideline, the efficiency of design would be included in the efficiency rating. The Bank's ICR Guidelines note that "the ICR analyses the project's efficiency using any other appropriate cost-effectiveness criteria to determine whether the project represented the expected least-cost solution to attain identified and measurable benefits by either an analysis of cost per unit of input or cost per unit of output" (p. 53).

Possible causes of implementation delays may include:

- Incomplete preparation of the project
- Unrealistic implementation schedule, including failure to schedule sufficient time for start-up activities and mobilization
- Unforeseen technical difficulties
- Changes in project/program scope
- Quality of management, including financial management
- Delays in selecting staff/contractors/consultants
- Delays in receiving funds, whether counterpart, or funds from the Bank or cofinanciers
- Inefficient procurement or disbursement, and
- Security problems and natural disasters

If an economic rate of return(ERR) or internal rate of return(IRR) has been calculated, the assumptions should be fully explained and transparent in the ICR. The reasons for selection of samples or examples should be explained, including rationale for extrapolating findings based on a selection of data to a wider group. Any data other gaps and methodological strengths or weaknesses in the Bank's assessment of efficiency should be noted. Normally, if an ERR or IRR or NPV (or similar) was estimated at project appraisal, the ICR should repeat the calculation based on information available at the time of closing. The ICR normally would indicate what the ERR or IRR or NPV was at project appraisal, what it is when re-estimated at completion, and on what percent of total project costs the original and revised estimates were based. Even if the ICR is prepared early, with a short period during which actual benefits have been observed, a re-estimate of ERR or IRR or NPV at least provides actual latest cost figures and an updated projection of benefits, reflecting any changes made during implementation. Underlying assumptions about costs and benefits, and any other information supporting the analysis (e.g. output volumes, major cost items, or prices) should be presented.

The ICR reviewer should include all available indicators of efficiency, including efficient use of project funds, in the assessment. Shortcomings in efficiency may have to do with the extent to which the project fails to achieve (or is not expected to achieve) a return higher than the opportunity cost of capital, and is not the least cost alternative (World Bank, OPCS 2011a; also see World Bank/IEG 2006).

In the Efficiency fields in the electronic ICR Review form, "Coverage/scope" refers to the percentage of total project cost for which the ERR/IRR/financial rate of return (FRR) was calculated. A comment should be included, where possible, on the reliability of the ERR/IRR/FRR calculation(s) presented in the ICR.

When a project objective is to improve the efficiency of a sector, avoid confusing efficiency of the project with efficiency of a sector

In some projects, one of the objectives is to improve the efficiency of a sector (e.g. the health sector, or the agricultural sector) within the country, or to improve the efficiency of a government program being supported. In such cases, achievement of improved efficiency of the sector or the government program represents achievement of a project objective, and therefore should be assessed under Efficacy, not under Efficiency. The rating of Efficiency in the ICR is intended to cover the extent to which project resources were used wisely and achieved good value for the money.

Especially in such cases, it is important to avoid confusing the efficiency of the project with achievement of improved efficiency of the sector or program being supported. The latter is an outcome and would be included in the assessment of efficacy. For example, in an education system, repetition and dropout rates might decline as the result of an education investment, which would suggest improved internal efficiency of the education system. It would not necessarily indicate that project resources were used efficiently (that is, that the project was implemented cost-effectively or at least cost). Likewise, efficiency is about the cost-effectiveness of *project* resources, not the use of *World Bank* budgetary resources.

Rating of Efficiency

Efficiency should be assigned an overall rating, based on a 4-point scale: negligible, modest, substantial, or high.
{To do: definitions for the different levels of efficiency rating .}

9. Project Outcome (ICRR Section 6)

Definition

The project outcome is defined as "the extent to which the project's major relevant objectives were achieved, or are expected to be achieved, efficiently."

Thus, the outcome rating is derived from the prior assessment of the relevance of objectives, efficacy in achieving each objective, and efficiency.

To assure consistency across IEG ICR Reviewers, IEG has developed guidelines for deriving the project outcome rating from the subratings on relevance, efficacy, and efficiency in the previous sections.

Guidance

The IEG/OPCS Harmonized Evaluation Criteria (World Bank, OPCS 2011a, p. 31) provide the following guidance in assigning an outcome rating, based on relevance, efficacy, and efficiency:
As the Bank is an objectives-based institution, achievements against the project development objectives (PDOs) are paramount. Thus, the benchmark for evaluation is the project's own stated objectives – not any absolute standard or someone else's conception of what good performance is. {to discuss with SC and MG – this and below}
What if there are unintended positive outcomes? If a project achieves positive outcomes that were not part of the objectives statement, then credit is not given for those positive outcomes

in the Outcome rating. The rationale for this is that neither the Bank nor IEG want to encourage 'serendipity' – the fortuitous chancing upon favorable outcomes without realizing that they were materializing. True, development is an unpredictable business, but that's why the Bank allows project restructuring. If positive outcomes that were unforeseen during project design become visible during implementation (and are the result of project actions), project teams have the option of restructuring the project, changing the objectives statement to internalize those positive outcomes (and also change the project design to maximize those outcomes), monitoring and evaluating them, and then getting credit for them in the Outcome rating. Typically, marshalling evidence for such unintended positive outcomes is difficult since their "unintended" nature means that they have likely not been well-monitored.

What if there are unintended negative outcomes? If the negative outcomes are the result of project actions, they will adversely affect the Outcome rating (e.g., if the Bank's road project in NE Brazil achieves its objectives in a highly efficient manner and was also highly relevant to the Bank's strategy in that country, but if it ended-up deforesting half of NE Brazil in the process of building the roads, the Outcome cannot be rated Highly Satisfactory).

Rating of Outcome

Highly satisfactory	There were <u>no</u> shortcomings in the project's achievement of its objectives, in its efficiency, or in its relevance.
Satisfactory	There were <u>minor</u> shortcomings in the project's achievement of its objectives, in its efficiency, or in its relevance.
Moderately satisfactory	There were <u>moderate</u> shortcomings in the project's achievement of its objectives, in its efficiency, or in its relevance.
Moderately unsatisfactory	There were <u>significant</u> shortcomings in the project's achievement of its objectives, in its efficiency, or in its relevance.
Unsatisfactory	There were <u>major</u> shortcomings in the project's achievement of its objectives, in its efficiency, or in its relevance.
Highly unsatisfactory	There were <u>severe</u> shortcomings in the project's achievement of its objectives, in its efficiency, or in its relevance.

Deriving the Outcome rating from IEG's subratings of Relevance, Efficacy, and Efficiency

The Outcome rating is derived from the assessment of the Relevance of Objectives, Efficacy in achieving each objective, and Efficiency. As WB and IEG have adopted an objectives-based evaluation methodology, achievements are assessed against the Project Development Objectives (PDOs).

The assessment of Outcome should first assess if the PDOs were relevant, second, if the PDOs were achieved or are expected to be achieved, and third, if the PDOs were achieved efficiently. If Relevance is High or Substantial, then the Outcome rating may or may not be above the line depending on the Efficacy and Efficiency ratings. If Relevance is Modest or Negligible, then the Outcome rating will tend to be below the line regardless of the Efficacy and Efficiency ratings.

For consistency, the following rules-of-thumb will the derivation of the project Outcome rating from the sub-ratings for Relevance, Efficacy, and Efficiency (table 9.1):

- In order to receive an Outcome rating of Moderately Satisfactory or higher, a project must be rated High or Substantial on Efficacy.

- A rating of Modest on Efficacy produces an Outcome rating that is, at best, Moderately Unsatisfactory.

- A rating of Negligible on any one of Relevance, Efficacy, and Efficacy produces an Outcome rating that is at best Unsatisfactory.

- In order for the Outcome to be rated Highly Satisfactory, a project must be rated High on Efficacy, High on any one of the other sub-ratings, and Substantial on the third.

- Tables 9.1 and 9.2 provides guidance on other possible scenarios.

The Outcome rating indicated by the guidance above and that contained in Table 1, should, in every case, be cross-checked against 'best judgment,' stepping back and asking the question "Were the shortcomings in Outcome non-existent, minor, moderate, significant, severe, or major? (leading respectively to a Highly Satisfactory, Satisfactory, Moderately Satisfactory, Moderately Unsatisfactory, Unsatisfactory, or Highly Unsatisfactory Outcome rating).

Table 9.1. Deriving the Overall Outcome Rating for a Project, Tree View

Sub-Ratings			Outcome Rating
Relevance	**Efficacy**	**Efficiency**	
High	High	High or Substantial	**Highly Satisfactory**
		Modest	**Moderately Satisfactory**
		Negligible	**Unsatisfactory**
	Substantial	High or Substantial	**Satisfactory**
		Modest	**Moderately Satisfactory**
		Negligible	**Unsatisfactory**
	Modest	High or Substantial or Modest	**Moderately Unsatisfactory**
		Negligible	**Unsatisfactory**
	Negligible	High or Substantial or Modest	**Unsatisfactory**
		Negligible	**Highly Unsatisfactory**
Substantial	High	High	**Highly Satisfactory**
		Substantial	**Satisfactory**
		Modest	**Moderately Satisfactory**
		Negligible	**Unsatisfactory**
	Substantial	High or Substantial	**Satisfactory**
		Modest	**Moderately Satisfactory**
		Negligible	**Unsatisfactory**
	Modest	High or Substantial or Modest	**Moderately Unsatisfactory**
		Negligible	**Unsatisfactory**
	Negligible	High or Substantial or Modest	**Unsatisfactory**
		Negligible	**Highly Unsatisfactory**
Modest	High	High or Substantial	**Moderately Satisfactory**
		Modest	**Moderately Unsatisfactory**
		Negligible	**Unsatisfactory**
	Substantial	High or Substantial	**Moderately Satisfactory**
		Modest	**Moderately Unsatisfactory**
		Negligible	**Unsatisfactory**
	Modest	High or Substantial	**Moderately Unsatisfactory**
		Modest	**Unsatisfactory**
		Negligible	**Highly Unsatisfactory**
	Negligible	High or Substantial or Modest	**Unsatisfactory**
		Negligible	**Highly Unsatisfactory**
Negligible	High	High or Substantial or Modest	**Unsatisfactory**
		Negligible	**Highly Unsatisfactory**
	Substantial	High or Substantial or Modest	**Unsatisfactory**
		Negligible	**Highly Unsatisfactory**
	Modest	High or Substantial or Modest	**Unsatisfactory**
		Negligible	**Highly Unsatisfactory**
	Negligible	High or Substantial or Modest or Negligible	**Highly Unsatisfactory**

Source: IEG and OPCS

Table 9.2. Deriving the Overall Outcome Rating for a Project, Table View

Rating	Definition	Relevance	Efficacy	Efficiency	Comment
Highly satisfactory	There were no shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance	High on any two criteria – *one of which must be efficacy* – and at least substantial on the third.			Requires efficacy to be one of the high ratings.
Satisfactory	There were minor shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance	*a.* Substantial on all three criteria *or* *b.* Substantial on two criteria and high on the third *or* Substantial efficacy but high relevance and efficiency			*"Minor" shortcomings are implicitly defined as substantially achieving the objectives, and substantial or better on the other two criteria.*
Moderately satisfactory	There were moderate shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance	Substantial (or high) on two criteria—one of which must be efficacy—and modest on the third.			*"Moderate" is implicitly defined as modest on one criterion.*
Moderately unsatisfactory	There were significant shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance.	Modest on any two criteria and substantial (or high) on the third, *or* Modest efficacy with substantial (or high) on the other two criteria.			*"Significant" is implicitly defined as modest on two criteria or modest efficacy. Would also apply if one were high and two were modest*.
Unsatisfactory	There were major shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance.	Modest on all three criteria *or* Negligible on one criterion and modest/substantial/high on the other two			*"Major" is implicitly defined as three modests or at least one negligible.*
Highly unsatisfactory	There were severe shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance.	Negligible on all three criteria Negligible on two criteria and modest/substantial/high on the third one			*"Severe" is implicitly defined as at least two negligible.*

Source IEG and OPCS.

In addition, the Harmonized Evaluation Criteria specifically mention that the variation in achievement of different objectives is to be taken into account:" Shortcomings *in the achievement of objectives may have to do with either the* <u>*number*</u> *of objectives that are not achieved (or are not expected to be achieved) and/or the* <u>*extent*</u> *to which one or more objectives are not achieved (or are not expected to be achieved"* (World Bank, OPCS 2011a, p. 31).

When insufficient information is provided by the Bank for IEG to arrive at a clear rating, this can be a reason for an unfavorable rating.

The rating of Outcome should encompass the extent to which the project's objectives were relevant and were achieved, or are expected to be achieved, efficiently.

- Shortcomings in the achievement of objectives may have to do with either the number of objectives that are not achieved (or are not expected to be achieved) and/or the extent to which one or more objectives are not achieved (or are not expected to be achieved).
- Shortcomings in efficiency may have to do with the extent to which the project fails to achieve (or is not expected to achieve) a return higher than the opportunity cost of capital, and is not the least cost alternative (this criterion may not apply for DPL operations).
- Shortcomings in relevance may have to do with the extent to which a project's objectives are inconsistent with current Bank country strategies (as expressed in PRSPs, and CASs/SCDs/CPFs). It is important to ensure that achievement of objectives reflects continuing priorities at the PDO level, not out-of-date priorities that should have triggered restructuring.

The IEG ICR Reviewer must use judgment in weighing possible shortcomings in the achievement of the project's objectives, in its efficiency, or in its relevance, and arrive at an assessment of how they affect the overall rating.

Rating the outcome of projects with formally revised objectives

If the project's development objectives (and/or key associated outcome indicators targets) have been revised (through a formally approved restructuring or additional financing), the ICR and the ICR Review should take into consideration both the original and formally revised objectives in deriving the project's Overall Outcome rating.

PRINCIPLE

The principle of reporting on results of projects with revised objectives is to ensure that the project demonstrates accountability for reaching the promised level of outcome, without penalizing the project for reasonable adaptations during implementation. Evaluation of project performance in the ICR, therefore, can take into account performance before and after revision of objectives. To do this, the ICR assesses achievement against both the original and revised objectives, assigning separate Outcome ratings (rating performance across the entire project time period against both the original and revised project objectives or outcome targets). Then, to derive the overall Outcome rating, the separate Outcome ratings are weighted in proportion to the share of actual loan/credit disbursements made in the periods before and after approval of the revision. This methodology aims to reward project teams that recognize issues early and restructure accordingly. It also aims to provide an incentive against late "lowering the bar", that is, lowering the ambition of a project to match achievement on the ground rather than to account for changes in context.

RATIONALE

The rationale for the above is based on the following:

- The Bank uses an objectives-based evaluation methodology whereby project Overall Outcome is assessed against the project objectives for which Bank funds have been approved.

- If substantial changes to the original project objectives and/or outcome targets are required, such changes have to be formally approved.

- For accountability purposes, the evaluation of project performance should take into account performance both before and after the revision of project objectives or outcome targets. Weighting pre- and post-revision performance by the share of actual Bank disbursements before and after the revision took place is both practical and transparent.

GUIDANCE AND CRITERIA

Changes in a project's scope or ambition (and therefore its outcome targets) may occur due to a variety of causes: additional financing, change in geographic coverage, change in population coverage/beneficiaries, an external event (such as an earthquake or tsunami), the withdrawal or new availability of an external partner/co-financier, etc. Alternatively, a project's outcome targets may undergo revision because of shortcomings in the original targeting, for example if there was inadequate understanding at appraisal of the appropriate level of ambition. In each of these instances, in determining whether to perform a split rating, the ICR author should assess whether the justification for the revision of outcome targets is valid and consistent with the project's theory of change. If a project receives additional financing, for example, to scale up a project's population coverage, and as a result its outcome targets are commensurately increased, then it is possible that no split rating is necessary and the project can be assessed based on the revised outcome targets. However, if the judgment is that the project's scope/ambition have decreased through a downward revision of targets, then a split rating should be performed.

Note that, during review of the ICR, IEG independently assesses the appropriateness of applying a split rating vs. assessing the entire project based on revised indicators. The ICR Review may apply a split rating even though the ICR does not, and vice versa.

Additional financing

If additional financing is received, but the project's objectives and outcome targets are not revised, a split rating generally should not be performed. The IEG ICR Reviewer should note, however, that additional financing not accompanied by an increase in the scope or ambition of a project may be an indication of lowered efficiency, and as such should be reflected in a downgraded efficiency rating. Particular instances include the following:

- If reason for additional financing was to cover a cost overrun, this can support a decision not to apply a split rating and instead to assess the project against the <u>original</u> indicators and targets.

- If the reason for additional financing was to add to the scope of the project (for example, additional geographical areas, additional beneficiary groups, or some other kind of additional scope) and the targets related to the increased scope were met, then this supports a decision not to apply a split rating and instead to assess the project based on the revised targets. (In such a case, applying a split rating is likely moot and can overly complicate the reporting.)

    - However, in a similar instance of expanded scope and increased targets, but where achievement was mixed -- for example targets for the geographical areas included in the original project were achieved, but targets for the geographical areas added at restructuring were not achieved -- then this supports a decision to apply a split rating.

Changes to project scope

If the scope of the project was changed at restructuring, the ICR should explain the reason(s) for the change. In general,

(1) if the scope of the project expanded and project commitments remained the same, a split rating is not warranted, and the project should be assessed based on the revised objectives/targets.

(2) If the scope of the project shrank and project commitments remained the same, a split rating should be applied.

Some particular instances are:

- If the scope of the project shrank and project commitments remained constant, this supports a decision to apply a split rating.

- If the scope of the project shrank, project commitments decreased (for example through cancellation of funds), and a good case is made that the lowering of project scope was commensurate with the lower commitment size, then this would support a decision not to apply a split rating and to assess the project based on the revised scope (and objectives). {MK says this should be a split.}

- If the scope of the project expanded (say, because of additional financing), this supports a decision not to apply a split rating and instead to assess the entire project based on the revised outcomes and outcome targets.

- If the reason(s) given for the change in scope are problematic or not well-founded, then this can support a decision to apply a split rating.

- If the geographic coverage expanded, such that the targets were more ambitious, then this supports a decision not to apply a split rating and instead assess the project based on the expanded geographical coverage.

- If geographic coverage declined and the funding amount remained the same – for example, the project before revision served 5 regions and after revision serves 3 regions – then this signals a smaller scope and therefore supports a decision to apply a split rating.

- If activities were re-aligned or resources were re-allocated to move the geographical coverage from one region to another, and the ICR made a convincing case that the level of ambition or difficulty was similar, then this can support a decision not to apply a split rating and instead assess the project based on revised objectives or under the revised geographical coverage.

- If the target coverage expanded or the definition of beneficiaries changed such that the targets were more ambitious, then this can support a decision not to apply a split rating and instead assess the project based on the targets related to the expanded coverage or the more encompassing definition of beneficiaries.

- If the target coverage declined or the definition of beneficiaries was narrowed (lowering the level of ambition with the same amount of funding), then this supports a decision to apply a split rating.

- If activities were re-aligned or resources were re-allocated to change the target or beneficiary coverage, and the ICR made a convincing case that the level of ambition or difficulty was similar, then this can support a decision not to applying a split rating and instead assess the project based on objectives under the revised coverage or the revised definition of beneficiaries.

- If some donor(s) or stakeholder(s) dropped out (therefore reducing the amount of funds or other resources available), and these stakeholder exit(s) resulted in scope shrinking or lowering the level of ambition within the same amount of funding, then this supports a decision to apply a split rating.

- If some donor(s) or stakeholder(s) joined the project (therefore increasing the amount of funds or other resources available), and stakeholder entry resulted in increasing the scope of the project or raising the level of ambition, then this can support a decision not to apply a split rating and instead assess the entire project based on the revised outcomes and outcome targets.

Changes in outcome indicators

Changes such as dropping old indicators, adding new ones – do not, in themselves (without changes in targets), trigger a split rating, as such changes may reflect different (and presumably better) measure of a project's achievement rather than raising or lowering of a project's ambition. Project teams should be encouraged to adopt better indicators whenever necessary and appropriate. Similarly, changes in a project's components, outputs, and output targets should not trigger a split rating, as these kinds of changes reflect a different (and presumably better) path, in the absence of a revision of PDOs or outcome targets, to achieve the same expected outcomes.

The IEG ICR Reviewer should exercise judgment, based on the project's scope and theory of change, to determine which changes of PDOs or outcome indicators should not trigger a split rating. This judgment should be described and defended in the ICR text.

Changes to targets resulting from adding or dropping indicators

It may be the case that a restructuring replaces an old indicator (with or without target) with a completely new indicator and target. The ICR should explain why the new indicator/target was chosen. In these instances, the ICR should assess whether the new indicator and target represent warrant a split rating, using the following considerations:

- If the change to the indicator/target represented lowering the project's scope or ambition (with the same commitment level), this would support a decision to apply a split rating.

- If the original indicator was relatively weak and there is a strong case to be made that the new indicator/target provides a better measure for achievement of the PDO, then this would support a decision not to apply a split rating and instead assess the entire project based on the revised indicator/target.
    - When making this decision, it is important for the ICR to explain how the revised indicator (with its target) is a robust, valid measure of the objective being assessed. The ICR author should not assume that an indicator listed under "Key PDO Indicators" in the results framework is an appropriate indicator for the particular objective or outcome. Part of the role of the ICR author is to identify areas of weakness in the results framework and, when possible, to supplement the results reported (based on project implementation information) with

information from other sources. Also, part of the role of IEG in reviewing the ICR is to critique the indicators used in the results framework.

- If the original indicator(s) (dropped at restructuring) did not have targets, while the new indicator(s) have targets, it may be possible to make a case that this represented an improvement in the validity of the indicators, and as such this would be a reason not to apply a split rating and instead to assess the project based on the revised indicators and targets (assuming the level of ambition and funding both remain constant). However, if the original indicator(s) were connected to outcome-level results, and new indicator(s) are connected to more intermediate-outcome-level results or output-level results, then this likely represents lowering the level of ambition of the project and therefore would support a decision to apply a split rating (and likely would support a relatively low efficacy rating).

- If project documentation or experience suggests that the reason for selecting a new indicator was mainly convenience (for example, using an indicator that was already being tracked, but that did not represent an improvement in validity), this would support a decision to apply a split rating.

In instances where a project's objectives and/or outcome targets are revised more than once, the above judgments and procedures should be repeated as necessary, determining the overall Outcome rating according to the percent of loan/credit/grant disbursements under each restructuring.

Procedure for applying split rating of Outcome

- Determine the actual total Bank disbursements before and after the date when the revised project objectives were formally approved.

- Rate project outcome against the original project objectives and against the revised project objectives, respectively. Assign a numeric value for each of the outcome ratings: Highly Satisfactory = 6, Satisfactory = 5, Moderately Satisfactory = 4, Moderately Unsatisfactory = 3, Unsatisfactory = 2, Highly Unsatisfactory = 1. It is important to note that, when objectives are revised, the project is rated against both sets of objectives separately, for the <u>entire duration</u> of the project (not just the period for which each of the objectives was in effect). Achievement of each individual objective (Efficacy), both original and revised, is assessed across the project's entire lifetime. Relevance and Efficiency ratings are given for the entire project at closing.

- Arrive at an overall rating by weighting the two ratings by the proportion of actual total disbursement under each set of objectives, and rounding to the nearest whole number (1 to 6).

Example

- Formal approval was obtained for a change in the project objectives of a project with an original loan of $115 million when $20 million was disbursed. $15 million was eventually cancelled.

- The project originally contained four objectives (A, B, C, and D); of those four original objectives, three (A, B, and C) were negligibly achieved across the project's entire lifetime, and one (D) was substantially achieved across the project's entire lifetime. After the restructuring, the project contained two objectives (D and E), one of which was carried over from the original project (D), and the other newly introduced at the restructuring (E). Objectives D and E were both substantially achieved, with those ratings taking into account evidence of achievement across the project's entire lifetime (in other words, objective D, which applied across the entire project period, is not rated separately for the time period before and after the restructuring.) Relevance and Efficiency were both rated as substantial for the entire project.

- Outcome under the original objectives is rated Unsatisfactory (a value of 2 on the 6-point scale), based on Substantial relevance of objectives, negligible Efficacy (negligible achievement of three objectives and substantial achievement of one objective), and substantial efficiency. Outcome under the revised objectives is rated Satisfactory (a value of 5 on the 6-point scale), based on substantial relevance of objectives, substantial achievement of both objectives, and substantial efficiency. The outcome rating is then determined by calculating the share of disbursements before and after restructuring. A total of $100 million was actually disbursed, of which 20 percent was disbursed before the restructuring. The weighted value of the outcome rating under the original objectives is the outcome rating (2) times its weight (0.2) = 0.4. The weighted value of the outcome rating under the formally revised objectives is the outcome rating (5) times its weight (0.8) = 4.0. The weighted average score is the sum of the two: 0.4 + 4.0 = 4.4. Rounding this to the nearest whole number, it amounts to an overall Outcome value of 4, or Moderately Satisfactory on the 6-point scale.

Example of the calculation of split ratings for a skills development project

Table 9.3. Overall Outcome Ratings

Table 1: Overall Outcome Ratings

Rating Dimension	**Original objectives**	**Objectives after first revision**	**Objectives after second revision**
Relevance of Objectives	Substantial		
Efficacy			
Objective 1: Improve access to vocational training	Modest	Substantial	High
Objective 2: Improve quality of vocational training	Modest	Substantial	Substantial

Objective 3: Improve the demand-responsiveness of vocational training	Modest	Modest	High
Overall Efficacy	Modest	Substantial	High
Efficiency	Modest		
Outcome Rating	Moderately Unatisfactory	Moderately Satisfactory	Moderately Satisfactory
Outcome Rating Value	3	4	4
Amount Disbursed (US$ million)	2.9	20.5	5.8
Disbursement (%)	9.9%	70.2%	19.9%
Weight Value	0.30	2.81	0.79
Total weights	3.9 (rounds up to 4)		
Overall Outcome Rating	**Moderately Satisfactory (4.0)**		

Source: adapted from ICRR of P118101.

10. Bank Performance (ICRR Section 8)

Definition

The Bank's performance is defined as the extent to which services provided by the World Bank ensured quality at entry of the project and supported effective implementation through appropriate supervision (including ensuring adequate transition arrangements for regular operation of supported activities after loan/credit closing), toward the achievement of development outcomes.
Bank performance is rated by assessing two dimensions: (a) Bank performance in ensuring quality at entry and (b) quality of Bank supervision. Based on the criteria discussed below, [14] the IEG ICR Reviewer rates the Bank's quality at entry and quality of supervision separately, and uses the OPCS/IEG Harmonized Evaluation Criteria Guidelines to arrive at an overall rating for Bank performance. {Will these still exist with ICR Reform? If so, we need to create and/or revise them}

Quality at entry

DEFINITION

Quality at entry, which is shorthand for "Bank performance in ensuring quality at entry" refers to the extent to which the Bank identified, facilitated preparation of, and appraised the project

[14] The lists of assessment criteria below are taken from the Quality Assurance Group's (QAG's) criteria for its seventh Quality of Entry Assessment (QEA7) (World Bank 2007) and its sixth Quality of Supervision Assessment (QSA6) (Hari Prasad 2005).

so that it was most likely to achieve planned development outcomes and was consistent with the Bank's fiduciary role.

CRITERIA

Bank performance in ensuring quality at entry is rated against the following criteria, as applicable to a particular project. The IEG ICR Reviewer should take account of the operational, sector, and country context in weighing the relative importance of each criterion of quality at entry as it affected outcomes.

- Strategic relevance and approach
- Technical, financial, and economic aspects (for investment lending projects)
- Poverty, gender, and social development aspects
- Environmental aspects[15]
- Fiduciary aspects
- Policy and institutional aspects
- Implementation arrangements
- M&E arrangements
- Risk assessment
- Bank inputs and processes.

Information on which to base the assessment of Quality at Entry may be found throughout the ICR, but sections of particular importance will be the description of Components, and the section on Key Factors Affecting Implementation and Outcome.

With respect to Components, the ICR should discuss whether the components were reasonable in relation to the following, and this discussion should be considered when assessing Quality at Entry:

- **Achieving the project/program objectives**;
- The **capacity** of the implementing agency (including administrative and financial management capacity);
- The likely **availability of local funding**; and
- **Lessons learned** in relevant prior projects.

[15] This would include provisions for safeguard policy compliance.

RATING OF BANK PERFORMANCE IN ENSURING QUALITY AT ENTRY

With respect to the relevant criteria that would enhance development outcomes and the Bank's fiduciary role, rate Bank performance in ensuring quality at entry using the following scale:

Highly satisfactory	There were no shortcomings in identification, preparation, or appraisal.
Satisfactory	There were minor shortcomings in identification, preparation, or appraisal.
Moderately satisfactory	There were moderate shortcomings in identification, preparation, or appraisal.
Moderately unsatisfactory	There were significant shortcomings in identification, preparation, or appraisal.
Unsatisfactory	There were major shortcomings in identification, preparation, or appraisal.
Highly unsatisfactory	There were severe shortcomings in identification, preparation, or appraisal.

Quality of supervision

DEFINITION

Quality of supervision refers to the extent to which the Bank proactively identified and resolved threats to the achievement of relevant development outcomes and the Bank's fiduciary role.

CRITERIA

Bank performance in quality of supervision is rated against the following criteria, as applicable to a particular project. The IEG ICR Reviewer should take account of the operational, sector, and country context in weighing the relative importance of each criterion of quality of supervision as it affected outcomes.

- Focus on development impact
- Supervision of fiduciary and safeguard aspects (when applicable)
- Adequacy of supervision inputs and processes
- Candor and quality of performance reporting
- Role in ensuring adequate transition arrangements (for regular operation of supported activities after loan/credit closing).

In assessing Bank performance, it can be helpful to keep in mind three groups of implementation factors that remain outside the control of the Bank:

Factors **outside the control of the Bank or government or implementing agencies**, such as:

- Changes in world markets and prices
- Unexpected and unforeseeable technical difficulties
- Natural disasters (including extraordinary weather and sudden disease epidemics)
- War and civil disturbances, including the effects overflowing from neighboring territories (such as refugees)

Factors generally **subject to government control**, such as:

- Macroeconomic and sector policies
- Government commitment
- Governance and corruption (see Handbook on Governance and Corruption)
- Appointment of key staff
- Provision of counterpart funds, and
- Efficient administrative procedures

Factors general **subject to implementing agency control**, such as:

- Management effectiveness
- Staffing adequacy and quality

RATING OF QUALITY OF SUPERVISION

In consideration of relevant criteria that would enhance development outcomes and the Bank's fiduciary role, rate quality of supervision using the following scale:

Highly satisfactory	There were no shortcomings in the proactive identification of opportunities and resolution of threats.
Satisfactory	There were minor shortcomings in the proactive identification of opportunities and resolution of threats.
Moderately satisfactory	There were moderate shortcomings in the proactive identification of opportunities and resolution of threats.
Moderately unsatisfactory	There were significant shortcomings in the proactive identification of opportunities and resolution of threats.
Unsatisfactory	There were major shortcomings in the proactive identification of opportunities and resolution of threats.
Highly unsatisfactory	There were severe shortcomings in the proactive identification of opportunities and resolution of threats.

Rating of overall Bank performance

The rating of overall *Bank performance* is based on the ratings for each of the two dimensions (i) Bank performance in ensuring quality at entry and (ii) the quality of supervision. Quality at entry and quality of supervision should each be rated using their respective 6-point rating scales, and for transparency, the individual ratings for quality at entry and quality of supervision should be presented separately. The quality-at-entry and quality-of-supervision ratings should be combined into a rating of overall Bank performance.

In general, the lower of the two ratings (of Bank performance in ensuring quality at entry and Quality of supervision) determines the rating of overall Bank Performance. The reason for this is that Bank Performance is considered to be entirely within the control of the Bank. A mistake or shortcoming in either one would have been avoidable. Also, mistakes or shortcomings in Bank Performance can be self-replicating. For example, shortcomings in design and preparation that present issues in Quality at Entry are likely to lead to issues during supervision, which means that time and Bank resources must be spent to fix problems that the Bank itself created.

Ratings for the most common combinations of ratings of quality at entry and quality of supervision are provided below, followed by additional guidance on other combinations.

Highly Satisfactory	Bank performance was rated highly satisfactory on both dimensions.
Satisfactory	Bank performance was rated satisfactory on both dimensions, ***or*** was rated satisfactory on one dimension and highly satisfactory on the other dimension.
Moderately Satisfactory	Bank performance was rated moderately satisfactory on both dimensions, ***or*** was rated moderately satisfactory on one dimension and satisfactory or highly satisfactory on the other dimension. (Also see guidance below.)
Moderately Unsatisfactory	Bank performance was rated moderately unsatisfactory on both dimensions. (Also see guidance below.)
Unsatisfactory	Bank performance was rated unsatisfactory on both dimensions, ***or*** was rated unsatisfactory on one dimension and moderately unsatisfactory on the other dimension.
Highly Unsatisfactory	Bank performance was rated highly unsatisfactory on both dimensions, ***or*** was rated moderately unsatisfactory or unsatisfactory on one dimension and highly unsatisfactory on the other dimension.

When the rating for one dimension is in the satisfactory range (moderately satisfactory or better), while the rating for the other dimension is in the unsatisfactory range, the rating for overall Bank performance normally depends on the outcome rating. Thus, overall Bank performance is rated moderately satisfactory ***if*** outcome is rated in the satisfactory range, or moderately unsatisfactory ***if*** outcome is rated in the unsatisfactory range, except when Bank performance did not significantly affect the particular outcome.

Note that ICRR Section 9, Borrower Performance, is not applicable for ICRs initiated starting July 1, 2017. The section remains part of the ICRR Form, however, during the transition to accommodate ICR Reviews completed under the previous guidelines.

11. Quality of Monitoring and Evaluation (ICRR Section 10)

Definition

The M&E quality rating is based on an assessment of three main elements: (i) M&E design, (ii) M&E implementation, and (iii) M&E utilization. Monitoring and evaluation are distinct, and the rating is informed by both the quality of monitoring and the quality of evaluation.

Guidance

In assessing the M&E quality rating, the IEG ICR Reviewer should note that there may be good M&E mechanisms located outside the project as well as inside - for example, national surveys related to child educational achievements. Such alternative arrangements, provided they exist and serve the purpose, are fully acceptable as the basis for assessing the quality of M&E rating, and they are often more sustainable than project-specific M&E systems. Moreover, while monitoring is an essential part of any project management system, impact studies relevant to a sector as well as a project, such as impacts on child health, may be more efficiently done through broader national assessments.

Rating M&E quality is not intended to call for a focus only on quantitative evidence. In addition, good M&E will always rely on sound qualitative evidence, on the triangulation of that evidence with quantitative findings, and on the linkage of the array of evidence with the postulated causality chain. Such triangulation is essential to reach the ultimate goal of understanding what happened as a result of the intervention.

In rating M&E quality, the IEG ICR Reviewer is asked to look at three sequential elements: (i) *M&E design*, as reflected in the project design and proposed methodologies mapped out in the documents up to the point of Board approval; (ii) *M&E implementation*, as reflected in the actual project M&E inputs and the methodologies applied over the period of project effectiveness; and, (iii) *M&E utilization,* as reflected in the changes made in the ongoing project or changes in subsequent interventions attributable to this work. These three elements are common to both investment and policy lending.

The IEG ICR Reviewer is asked to discuss each of the three elements of M&E quality separately and to arrive at an overall *quality of M&E* rating on a 4-point scale.[16]

Criteria

M&E DESIGN

The IEG ICR Reviewer should assess to what extent the M&E design was sound and was designed to collect, analyze, and provide decision makers with methodologically sound assessments, given the stated objectives. The IEG ICR Reviewer also needs to assess the extent to which the methodology proposed in the PAD would enable the assessment of attribution. The specific questions in assessing M&E design are:

- To what extent was the theory of change (documenting how the key activities and outputs led to the outcomes) sound and reflected in the results framework?
- To what extent were the objectives clearly specified?

[16] As of this writing, IEG does not provide ratings on the three separate elements of M&E quality.

- To what extent were there indicators encompassing all outcomes of the PDO statement?
- To what extent were the intermediate results indicators adequate to capture the contribution of the operation's components (activities) and outputs toward achieving PDO-level outcomes?
- To what extent were the indicators specific, measurable, achievable, relevant, and time-bound? To what extent were there baselines and targets available for all indicators?
- To what extent were the proposed sampling methods, data collection methods, and analysis appropriate for all indicators? How were comparators selected and handled?
- To what extent did the design ensure that a baseline, if relevant, would be done in time?
- To what extent were the M&E design and arrangements well-embedded institutionally?

M&E IMPLEMENTATION

In *M&E implementation*, the IEG ICR Reviewer should assess to what extent the input, output, outcome, and impact evidence anticipated in the design was actually collected and analyzed in a methodologically sound manner. Specifically:

- To what extent was planned baseline data collection actually carried out?
- To what extent were the indicators included in the Results Framework measured and reported?
- To what extent were any weaknesses in M&E design, including specification of indicators, corrected during implementation?
- To what extent did the agency responsible for M&E (as well as any other relevant stakeholders) ensure attention to effective M&E implementation?
- To what extent are the data found to be reliable and of good quality? Important elements of this include sound methodology, independence of analysts, and quality control.
- If relevant, to what extent were beneficiaries involved in defining target indicators and assessing their achievement?
- To what extent are M&E functions and processes likely to be sustained after project closing?

USE OF M&E DATA

The IEG ICR Reviewer should assess, first, to what extent the M&E findings were communicated to the various stakeholders and, second, to what extent this informed strategic redirection and resource reallocation, or is expected to lead to these in follow-on interventions. The specific questions to be answered in assessing M&E utilization are:

- To what extent were M&E findings communicated to the various stakeholders?
- To what extent can positive (or negative) shifts in the implementation direction of the project or program be attributed to the M&E activities? (Examples of formal shifts in implementation direction include restructuring or additional financing, but other shifts that affected implementation should also be noted.)
- To what extent was the M&E data used to provide evidence of achievement of outcomes, as opposed to only providing evidence of application of inputs or achievement of outputs?
- To what extent did the M&E data or findings inform subsequent interventions (if known by the time of assessment), or to what extent is it expected to influence subsequent interventions in the near term?

Determining the overall M&E quality rating

The quality of M&E is rated on a 4-point scale—negligible, modest, substantial, or high.

- **High:** There were minor shortcomings, or only minor shortcomings, in the M&E system's design, implementation, or utilization. The M&E system as designed and implemented was more than sufficient to assess the achievement of the objectives and to test the links in the results chain. M&E findings were disseminated and used to inform the direction of the project, strategy development, and/or future projects.
- **Substantial:** There were moderate shortcomings in the M&E system's design, implementation, or utilization. The M&E system as designed and implemented was generally sufficient to assess the achievement of the objectives and test the links in the results chain, but there were moderate weaknesses in a few areas.
- **Modest**: There were significant shortcomings in the M&E system's design, implementation, or utilization. There were significant weaknesses in the design and/or implementation of the M&E system, making it somewhat difficult to assess the achievement of the stated objectives and test the links in the results chain, and/or there were significant weaknesses in the use and impact of the M&E system.
- **Negligible**: There were severe shortcomings in the M&E system's design, implementation, or utilization. The M&E system as designed and implemented was insufficient to assess the achievement of the stated objectives and test the links in the results chain, and the use and impact of the M&E system were limited.

Relation of M&E quality to other ratings

Strengths and weaknesses in M&E design should also be reflected in the Bank quality at entry rating, and those in M&E implementation should be reflected in the Bank supervision rating as well as the implementing agency performance rating. The relevance of project design includes an assessment of the results framework and its relevance to the objectives. Rating overlap is acceptable, and may be expected.

Beyond these direct links to other ratings, weak M&E may often impact effective project management, and thus it can indirectly affect the project efficacy and efficiency rating. If M&E is insufficient, it can affect the IEG ICR Reviewer's ability to assess the achievement of the project's objectives.

Standards for M&E quality for projects with learning objectives

IEG holds projects with an explicit objective of learning, such as pilots or Learning and Innovation Loans, to higher M&E standards. It does so on the grounds of particularly high relevance of monitoring and, in particular, evaluation, to the explicitly declared learning objectives of these projects. Strong M&E would be assumed to be the primary means for achieving the learning objective of these projects, even if it is not stated as such, and is thus a critical part of the results chain in and of itself.

12. Quality of the ICR (ICRR Section 15)

Because the ICR Review is almost entirely based on the information found in the ICR, the reliability of IEG's ratings based on the desk review depends critically on the accuracy and quality of the evidence it provides. For this reason, IEG rates the quality of the ICR.

Criteria

The assessment of the quality of the ICR is based on the following criteria:

- Quality of Evidence – is the evidence from a credible source, appropriately referenced and presented in a parsimonious fashion? Does the ICR, including annexes or appendices, present a complete and robust evidence base to support the achievements reported?
- Quality of Analysis – has there been sufficient interrogation of the evidence, concise summarizing of salient points and clear linking of evidence to findings?
- Extent to which lessons are based on evidence and analysis – are the lessons appropriately responding to the specific experiences and findings for the project? Are they sufficiently linked to the narrative and ratings included in the report?
- Results orientation – does the report emphasize and highlight how activities inform outcomes which in turn is linked to the impact of the project's intervention? Is the report focused on what occurred as a consequence of the project?
- Internal consistency – is there a logical linking and integration of the various parts of the report and are the results mutually reinforcing?

- Consistency with guidelines – has the report followed and responded to the guidelines, both with regards to ratings and the performance narrative?
- Conciseness – is there sufficient clarity in the report's messaging? Is the performance story direct, well informed and tightly presented?

Guidelines

In commenting on the quality of the ICR, it is generally a good strategy to begin by highlighting the strengths of the ICR before touching on the weaknesses. Candor, for instance, is highly valued. See the ratings profiles below for more complete information on what to look for in a high-quality ICR. Problems in ICR quality that should be flagged include: inadequate evidence; incomplete ICRs (such as missing data in tables, no discussion of efficiency); failure to assess the objectives; and relying too much on monitoring indicators instead of using all available data. IEG does not downgrade the ICR quality simply because of a difference in opinion about the ratings. There is also not necessarily a relationship between the project's outcome rating and the quality of the ICR. Some of the best ICRs have been written for projects that were unsatisfactory.

Rating of the Quality of the ICR

Based on these criteria the rating profile is as follows:

- **High** – The ICR is a tightly written and complete critique of the project. There is a clear link between the narrative, the ratings and the evidence. It provides a candid, accurate and substantiated set of observations that are aligned to the project development objective. The report is concise, follows the guidelines, seeks to triangulate data to reach conclusions, and is focused on results. The quality of evidence and analysis is substantial and informs all aspects of the ICR and there are few lapses in the quality of data and information. There is a well-articulated theory of change informing the reader as to how the ratings have been reached and the lessons are specific, useful and based on evidence of what actually occurred in the project.
- **Substantial** – The ICR provides a detailed overview of the project. The narrative supports the ratings and available evidence. It is candid, accurate and generally aligned to the project development objective. The report is concise, follows the majority of the guidelines, makes an attempt to triangulate data to reach conclusions, and is focused on results. The quality of evidence and analysis is aligned to the messages outlined in the ICR, though there may be some minor shortcomings in the completeness of data and information. There is a reference to the project's theory of change that helps the reader to understand how the ratings have been reached. The ICR's lessons are clear, useful and based on evidence outlined in the ICR.
- **Modest** – The ICR provides a comprehensive overview of the project. The narrative loosely supports the ratings and there are some gaps in evidence. It is relatively candid, predominantly accurate and generally aligned to the project development objective.

The report covers a wide range of issues, follows most of the guidelines and is focused on results. There is an attempt to link the quality of evidence and analysis to the messages outlined in the ICR. There is an effort to articulate how the ratings have been reached, and there may be some gaps in data and in various sections of the ICR. The ICR's lessons are generally useful and based on evidence outlined in the ICR, but may be overly general and not targeted to specific events or actions.

- **Negligible** – The ICR provides a basic overview of the project. The narrative loosely supports the ratings and there are obvious and consistent gaps in evidence. It is not particularly candid, may have evident inaccuracies and is not always aligned to the project development objective. The report covers a range of issues, follows some of the guidelines and is irregularly focused on results. The link between the quality of evidence and analysis in the ICR is not always reflected in the reports messages. There is some attempt to articulate how the ratings have been reached but there are evident gaps in data and in various sections of the ICR. The ICR's lessons are likely to be overly general and not targeted to specific events or actions.
- **Not Rated/Not Applicable** – there is insufficient material to articulate a rating.

13. Risk to Development Outcome (ICRR Section 7)

Definition

The risk to development outcome is the risk, at the time of evaluation, that development outcomes (or expected outcomes) will not be maintained (or realized). This refers to outcomes that have actually been achieved (or are expected to be achieved).

Guidance

The risk to development outcome has two dimensions:

(i) The likelihood that changes may take place that are detrimental to the ultimate achievement of the project's development outcome

(ii) The impact of some or all of these changes on the project's development outcomes.

Some risks are internal or specific to a project. They are primarily related to the suitability of the project's design to its operating environment.

Other risks arise from factors outside the project. These may appear at the country level, such as price changes, or on a global scale, such as technological advances. The impact on outcomes of a change in the operating environment depends on the severity and nature of the change, as well as the adaptability (or lack thereof) of the project's design to withstand that change.

Assessment of risk to development outcome requires an assessment of the uncertainties faced by a project over its expected useful life and of whether adequate arrangements are in place to help avoid or mitigate the impact of those uncertainties. This risk is understood to be higher if the design or implementation of the project is not well aligned with the operating environment, or mitigation measures are inappropriate to deal with foreseeable risks.

While the outcome rating reflects the IEG ICR Reviewer's best estimate of the expected overall development outcome, the assessment of risk to development outcome reflects the IEG ICR Reviewer's judgment of the uncertainties faced by the project's development outcomes over its expected useful life, taking account of any risk-mitigation measures already in place at the time of evaluation. Risk to development outcome says nothing about the absolute level of the expected net benefits. So, for example, a project can have a high expected rate of return and a satisfactory outcome, but the risk to development outcome may be high in its particular operating environment.

{Include a discussion about how the risk is treated when little was achieved: If very minimal outcomes were achieved, but the risk to that minimal achievement is small – do we rate risk low or high? Is there a threshold below which we decide that the risk refers to what wasn't achieved rather than what was?}

Criteria

The overall risk to development outcome is based on both the probability and the likely impact of various threats to outcomes, taking into account how these have been mitigated in the project's design or by actions taken during its initial implementation. The IEG ICR Reviewer should consider the operational, sector, and country context in weighing (in each case) the relative importance of these individual criteria of risk as they may affect planned outcomes.

- Technical (where innovative technology and systems are involved)
- Financial (including the robustness of financial flows and financial viability)
- Economic (at both the country and global levels)
- Social (in terms of the strength of stakeholder support and/or mitigation of any negative social impacts)
- Political (for example, volatility of the political situation)
- Environmental (including both positive and negative impacts)
- Government ownership/commitment (for example, continuation of supportive policies and any budgetary provisions)
- Other stakeholder ownership (for example, from private sector/civil society)
- Institutional support (from project entities; and/or related to legal/legislative framework)
- Governance
- Exposure to natural disasters.

14. Safeguards Compliance, Fiduciary, and Unanticipated Impacts (ICRR Section 11)

{need to check the section below with Lauren Kelly?}

OPCS requires that the ICR "summarize key *safeguard and fiduciary issues* in the project, compliance with the Bank policy and procedural requirements, and any problems that arose and their resolution, as applicable." {to be updated based on revised ICR guidelines} {also to be updated based on input of safeguards focal point} It also asks that the ICR "record any significant *deviations or waivers* from the Bank safeguards/fiduciary policies and procedures" and provides a list of the operational policies/Bank policies that apply (World Bank, OPCS 2011b, p. 21). This section provides some background on what to look for and how to interpret this information in reviewing the ICR. More detailed review of primary documents outside of the ICR would be appropriate for IEG's field-based PPARs.

Safeguards

WHAT IS THE PURPOSE OF ENVIRONMENTAL AND SOCIAL SAFEGUARDS?

The World Bank's environmental and social safeguard policies are designed to ensure that the potentially adverse impacts of Bank-supported programs on the environment and on people are avoided or minimized and that unavoidable adverse impacts are mitigated.

WHAT ARE THE DIFFERENT TYPES OF SAFEGUARDS AND HOW ARE THEY TRIGGERED?

There are currently 10 safeguard policies to address adverse environmental and social impacts of Bank-financed projects:[17]

OP 4.01, *Environmental Assessment*

OP 4.04, *Natural Habitats*

OP 4.09, *Pest Management*

OP 4.10, *Indigenous Peoples*

OP 4.11, *Physical Cultural Resources*

OP 4.12, *Involuntary Resettlement*

OP 4.36, *Forests*

OP 4.37, *Safety of Dams*

[17] The Bank is undertaking a review of its Safeguards policies and is expected to revise its Safeguard framework in 2012. {to be updated based on present scenario}

OP 7.50, *Projects on International Waterways*

OP 7.60, *Projects in Disputed Areas*

During project identification, the Bank screens a project with these safeguards policies in mind and classifies the project into one of four categories - A, B, C, and FI - based on the significance of environmental and social risk (box 14.1). The assigned category signals the appropriate level of environmental and social review required for the proposed project prior to project appraisal to identify the mitigation actions required.

Box 14.1. Safeguard Categories

Category A: *Projects likely to have significant adverse environmental impacts that are sensitive, diverse, or unprecedented. Potential impact is considered "sensitive" if it may be irreversible (for example, lead to loss of a major natural habitat) or it raises issues covered by Operational Policies on Natural Habitats, Indigenous Peoples*, Physical Cultural Resources, or In*voluntary Resettlement. These impacts may affect an area beyond the sites or facilities subject to physical works.*

Category B: *Projects that have potentially adverse environmental impacts on human populations or environmentally important areas—including wetlands, forests, grasslands, and other natural habitats—but these are less adverse than those of category A.*

Category C: *Likely to have minimal or no adverse environmental (or social) impacts.*

Category FI*: Applies when the Bank provides funds to participating national banks, credit institutions, and other financial intermediaries for financing subprojects that may result in environmental (or social) impacts. The financial intermediary must screen each subproject proposed for financing and classify it into any one of three categories: A, B, or C.*

HOW ARE THESE SAFEGUARD REQUIREMENTS ADDRESSED IN PROJECTS?

Depending on the safeguard policies that are triggered and the category, an appropriate choice is made from a range of instruments to satisfy the requirements. These include Environmental Assessments, Resettlement Action Plans or Resettlement Action Frameworks, and Indigenous Peoples Plans or Planning Frameworks. These instruments set forth what action will be required in terms of mitigation when the issues can be identified in advance, or they set forth a framework for how they will be handled if they cannot be identified in advance.

WHO IS RESPONSIBLE FOR COMPLYING WITH SAFEGUARD POLICIES?

The Bank's role is to determine the safeguard category during project appraisal, to prepare an Integrated Safeguards Data Sheet and disclose it publicly, and to ensure that appropriate safeguards of adequate quality are implemented by the borrower as it supervises the project. The primary responsibility for implementing the safeguards actions resides with the borrower. The borrower is responsible for preparing the appropriate assessment instrument and disclosing it, implementing and monitoring a mitigation plan consistent with the requirements, and making arrangements for independent verification of implementation.[18]

WHAT INFORMATION ARE THE PAD AND ICR EXPECTED TO PROVIDE ON SAFEGUARD COMPLIANCE?

The PAD should include a section on safeguard policies, including a table on the specific policies triggered, and will also note the safeguard category. If the project falls into category A or B, the PAD is supposed to have an appendix that explains the assessment instrument and mitigation plan.

The ICR should have a section (2.4) on "Safeguards and Fiduciary Compliance" that describes what was done and includes quantitative indicators of the extent to which the safeguard objectives were achieved. If the project is category A, the ICR should also summarize the findings of the Independent Panel of Experts and/or the impact assessment. The ICR for category B and FI projects may also summarize the results of any third-party impact assessment.

WHAT SHOULD THE REVIEWER RECORD IN THE SAFEGUARD SECTION?

The reviewer should note the following information:

- The applicable safeguard policies, if any, the environmental category of the project at appraisal (A, B, C, or FI), and (for category A and B projects) the assessment instrument and mitigation plan.[19] These can be found in the PAD.
- For category A and B projects, evidence that the project completed the planned mitigation activities, from the ICR.
- For category B or FI projects that relied on environmental and/or social policy frameworks, it should be noted whether the subprojects generated environmental or social impacts during implementation and, if so, how the project addressed them, from the ICR.
- The findings of any independent review of safeguards implementation (for high-risk projects) or monitoring reports (for others). This would include, for category A projects, the findings of the Independent Panel of Experts, or other types of impact assessment for categories B and FI.
- If the physical components of the project that generated environmental or social effects were modified—through additional financing or project restructuring, for example—the reviewer should note whether the environmental or social assessment was updated or a new assessment prepared.

The absence of any of the above required information in the PAD or ICR should be noted.

[18] In the case of category A projects, this takes the form of reports by an Independent Advisory Panel for Safeguards. Category B and FI projects would typically not have an Independent Panel of Experts, but may rely on some form of third-party impact assessment.

[19] Any instruments or plans mentioned for other categories should also be mentioned, but they are required for category A.

Are any of the ratings affected by performance on safeguard compliance?

There is currently no formal rating for safeguard compliance. However, the results in the safeguards section will affect other ratings.

- If the ICR fails to document any of these issues, that should also be mentioned in the ICR Quality section and contribute to the assessment of that rating. Likewise, an exemplary explanation of safeguard issues should also feed into the ICR quality rating.
- Good or poor performance in preparation (identifying the applicable policies, preparing the assessment and mitigation plans) should be a factor in the Bank performance/quality at entry rating.
- The Bank's performance in supervising safeguard compliance should be reflected in the Bank performance/quality of supervision assessment.
- Good or poor performance in adequately mitigating the impacts of safeguard issues should enter into the Bank supervision rating and the borrower performance rating.

Fiduciary issues

What constitutes a fiduciary issue?

Fiduciary issues refer to compliance with operational policies on Financial Management (OP/Bank Procedure [BP} 10.02), Procurement (OP/BP 11.00), and Disbursement (OP/BP 12.00). This material is to be culled from throughout the ICR.

Financial management

Financial management issues involve the adequacy of the project's institutional financial management arrangements, reporting and accounting provisions, internal control procedures, planning and budgeting, counterpart funding, flow-of-funds arrangements, external audit reporting, and project financial management and accounting staff issues. Particular attention should be paid to the timeliness of project external audits and whether the external auditors' opinions were qualified. If so, the nature of the qualifications (that is, whether they were serious or merely administrative) and the measures taken to address them should be included. If the ICR does not offer comments on the latter, the review should note the absence of information.

Other important aspects of financial management include:

- The extent of compliance with financial covenants (this should be reported in the ICR)
- Whether all Bank, IDA, and (where relevant) trust fund resources were fully accounted for by the time of project evaluation
- Issues of corruption or misuse of funds associated with the project, and how they have been addressed
- Whether all audit recommendations had been addressed by the time of project evaluation.

PROCUREMENT

Procurement issues include: the extent to which Bank procurement guidelines were followed; significant implementation delays due to procurement-related issues and their causes; and evidence of timely Bank intervention in resolving procurement difficulties, providing procurement advice, or in giving non-objections. Common causes of procurement-related delays or issues include misprocurement, low procurement capacity in the implementing agency, and lack of consistency between Bank and national procurement laws and regulations. Any issues of this nature should be discussed in the section and also mentioned in the Bank Performance/Quality of Supervision section.

DISBURSEMENT

Disbursement as a fiduciary issue rarely arises as a separate issue in the ICR, but would include:

- For investment projects, eligibility of expenditures (there may be special conditions attached to some disbursement categories - for example, for food and severance pay). If it is found that Bank or IDA funds were disbursed for ineligible expenditures, they are supposed to be refunded by the borrower, preferably before project closure (see Financial Management, above). {to check with SC and MG}

HOW ARE FIDUCIARY ISSUES REFLECTED IN THE RATINGS?

While there is no rating for fiduciary issues, they can be related to Bank performance, and in some cases in the efficiency component of the outcome rating. The information provided in the ICR on fiduciary issues is often incomplete, particularly with regard to financial management dimensions. For example, some ICRs do not discuss external project audits or whether they were qualified. In such cases, clarification should be sought from the TTL, and mention should be made of the shortcoming in the section of the review on ICR Quality. If lack of clarity on important fiduciary questions persists even after consulting the project team, this should be reflected in the Bank performance rating.

Unanticipated positive and negative effects

Even when a project's objectives are not achieved, implementation often yields many benefits. However, those benefits are not taken into account in the assessment of the objectives. An unanticipated benefit is a positive or negative benefit or externality that occurred outside the framework of the stated objectives of the project.[20] To be included in this section, they must be truly unanticipated (in the PAD or program document), attributable to the project, quantifiable, of significant magnitude, and at least as well evidenced as the project's other outcomes. Where there are unintended benefits, an assessment should be made of why these were not

[20] From the OPCS ICR guidelines (World Bank, OPCS 2011a, p. 29): "To the extent not previously covered, if any, assess the operation's positive or negative effects on the population or its subgroup(s) (including unintended or unexpected) with regard to the Bank's social objectives. Discuss whether the effects were foreseeable, identify causes of success or shortcomings, and assess how they will affect the future operation of the project."

"internalized" through project restructuring by modifying either project objectives or key associated outcome targets.

15. Ratings Summary (ICRR Section 12)

This table displays main ratings from the ICR and compares them with IEG's ratings in the ICR Review. The ICR ratings are automatically pulled from the Operations Portal, and the IEG ratings are pulled from the earlier sections in the ICR Review Form.

In cases where the ratings diverge, the IEG ICR Reviewer should explain the reasons for divergence. The explanations can be short and link back to summary statements in earlier sections of the ICR Review. The explanation for divergences in outcome rating can repeat elements in section 6, which summarizes the outcome rating.

16. Deriving Lessons (ICRR Section 13)

The purpose of the ICR is twofold: accountability and learning. Getting the ratings right is important, but learning what works, what doesn't work, and why is the key to greater effectiveness in the future. An ICR without good lessons is a missed opportunity to learn and do better. ICRs for projects that do not achieve their objectives often produce some of the most valuable lessons. {Add more explanation about how to formulate good lessons.}

The ICR Review typically presents three-to-five key lessons that emerge from the information in the report. They may come from the ICR, or they may be reflections on this project from the IEG ICR Reviewer based on the ICR, compared with other projects the IEG ICR Reviewer has reviewed (or, for example, confirming that the findings in this ICR underscore evaluative findings or lessons from other IEG evaluations). Whatever the case (whether the lesson is from the ICR or from IEG), it is incumbent on the IEG ICR Reviewer to identify the source.

Even if lessons in the ICR are not well formulated, the ICR Review should formulate them well. The two biggest issues in formulating lessons—in the ICR and by IEG ICR Reviewers—are: (a) they are formulated as facts, findings, or recommendations, rather than lessons; and (b) they are not underpinned by the evidence in the ICR. Table 16.1 below distinguishes between facts, findings, lessons, and recommendations.

Table 16.1. The Difference Between Facts, Findings, Lessons, and Recommendations

	What is it?	**Example**
Fact	What happened—an event and data (results). Not in dispute.	"The project manager was dismissed in year 5."
Finding	What the analyst interpreted or concluded from the facts specific to the project. Can be disputed.	"Mainly because replacement of the project manager was delayed, the project did not meet its targets."
Lesson	The broader significance of a finding. it draws a conclusion from experience that may be applicable beyond the project under review.	"Poor performance by project managers can critically affect project outcomes."
Recommendation	Suggests how to proceed in the future in the light of this experience. Proposes actions.	"The borrower should ensure that key project management positions are filled with competent staff. The Bank should help ensure this through appropriate covenants and prompt supervision."

Source:

Facts and findings will be found throughout the ICR; they are the material from which lessons are built. If you find that you are repeating something verbatim that is already in the ICR, then it probably is not a lesson, but a fact or finding. If you find that your draft lesson has the words "should" or "needs to" or "ideally," then it is very likely a recommendation and not a lesson.

The second major problem encountered is that the lessons as formulated are often not based on the evidence in the ICR.

- *When something has been found not to work, do not suggest what should have been done instead.* There are two reasons for this. First, when some aspect of the project has been shown not to work, in most cases the ICR will not have any evidence on what would have worked, only what did not. One cannot assume that approaches being used elsewhere would have worked in the context of the country under study—that would be speculation. Second, suggesting what should have been done or should be done in the future would be a recommendation, not a lesson.
- *If there is good evidence, a valuable lesson can point to why an intervention or project worked or did not work in the context of the country under study*. Often there is variation in project performance across areas—in some places the intervention had better results than in others. Pointing to the factors that led to success in some cases but not in others—including contextual differences—can be extremely helpful, not just for the country under study, but for other countries as well.
- *Lessons can also usefully point to the contextual factors under which an intervention succeeded or not.* Explaining the context of the results yields valuable insights into the constraints faced—for example, low-income households, fragile states, middle-income countries, and weak or strong institutions. Good results in spite of difficult contexts, or weak results despite more favorable contexts are good fodder for lessons. The context also

includes aspects of Bank and borrower performance, such quality at entry, M&E, country ownership, capacity of the implementing agency, and so forth.

Finally, the lessons should be clearly and concisely stated.

In the ICR Review form, the “lessons” stated (whether from IEG or the ICR) should be properly formulated and evidenced, and the source cited (whether from the ICR or IEG). Comments on the quality of the lessons (including the extent to which they are evidence-based) belong in the section on the quality of the ICR.

Remember that even projects rated unsatisfactory with poor M&E generate important lessons.

17. Other Considerations for the ICR Review

DOES THE ICR REVIEW CONFINE ITSELF ONLY TO EVIDENCE ON THE PROJECT'S KEY PERFORMANCE INDICATORS THAT WERE IDENTIFIED IN THE PAD OR ISRS?

No, all evidence, regardless of the source, is to be brought to bear in preparing the ICR Review, so long as the evidence is of quality. The ICR should provide enough information for IEG to be able to assess the quality of the evidence (e.g. the methodology of the Beneficiary Assessment, how the control group or comparison group was selected, and so on).

HOW SHOULD INTERLINKAGES BETWEEN RATINGS BE ADDRESSED?

For the most part, each of the key ratings – Outcome, Bank Performance – measure distinct dimensions of development effectiveness and are independent of each other. So, for example, Outcome may be rated highly unsatisfactory (say in a fragile state where a political coup erodes government commitment and the project objectives remain unachieved), while Bank Performance may be rated highly satisfactory (if the political coup was wholly unpredictable and the Bank had done the best it could under the circumstances), or vice versa.

In practice, however, there can be a number of interlinkages among the ratings which must be borne in mind in order to ensure internal consistency among the ICR Review ratings. Some of these interlinkages are deliberate and obvious while others are not so obvious:

- The IEG-OPCS Harmonized Evaluation Criteria introduce a deliberate interlinkage between the Outcome and the Bank Performance ratings – when the two elements of Bank Performance are in opposite directions (one above the line and the other below), the Outcome rating becomes the tie-breaker.

- There is another deliberate linkage between the Bank Performance rating and the following sub-ratings/dimensions: M&E Quality rating, safeguard compliance, fiduciary compliance, unintended positive and negative effects. These sub-ratings/dimensions were introduced as separate sections in the ICR Review form to specifically zoom-in on, and give prominence to, particular aspects of Bank Performance. So, if, say M&E Quality is rated unfavorably, Bank Performance cannot be rated too favorably. Similarly, if there is weak fiduciary compliance, Bank Performance may be affected.

- Not-so-obvious interlinkages can manifest themselves between the Bank Performance ratings and the Outcome rating depending on the extent and nature of weaknesses in the above mentioned sub-ratings/dimensions. So, for example, if M&E Quality is extremely weak or M&E is non-existent, that would raise questions about how effectively project implementation could have occurred and, therefore, how favorable Efficacy (and hence Outcome) could be. Similarly, if there were significant unintended negative effects attributable to the project (for example, a project in which a road was built in line with the stated project objectives and was also efficiently built, but the surrounding areas were deforested in the process), Outcome could not be rated favorably. Also, if fiduciary compliance was weak and there was evidence of substantiated corruption, then that would signal an inefficient use of project resources and Outcome could not be rated favorably.

- There is a final set of not-so-obvious interlinkages that the IEG ICR Reviewer needs to also be aware of: While the Outcome rating and the Bank Performance ratings can certainly go in different directions, the IEG ICR Reviewer should be able to explain the reasons for any divergence. Generally, if Outcome is rated unfavorably, Bank Performance would typically also be rated unfavorably. There are some exceptions, however: First, an unfavorable Outcome rating could be associated not with poor implementation performance on the part of the Bank but rather on the part of other donors/cofinanciers, if there were such other donors/cofinanciers in the project. Second, an unfavorable Outcome rating could be associated with an exogenous shock (e.g., an earthquake that wipes out the project roads). Finally, it could be that the Bank supports a high-risk project, the risk materializes, the Bank makes an informed decision not to cancel or restructure the project because the rewards could be extraordinarily high if the project succeeded – in such a case, if the project fails to achieve its objectives, Outcome could be rated unfavorably while Bank Performance could both be rated favorably. In any event, when the ratings diverge, the IEG ICR Reviewer should be able to explain the reasons for the divergence. Otherwise, the divergence may be unjustified.

“RED FLAGS” OR EXAMPLES OF RATINGS PATTERNS TO CHECK:

Scenario	Red flag or likely OK?	Explanation
Outcome MS+ M&E Quality Modest	Red flag	While you can have a good M&E system showing a lack of outcomes, it would be strange to have evidence of good outcomes with a poor M&E system, although it could be that an independent impact evaluation was done that showed the achievement of outcomes despite poor M&E.
Outcome MU- M&E Quality Substantial	likely OK	
Outcome MS+ Bank Performance MU-	Red flag	Unless there was a negative unexpected shock that caused the outcome to disappear, it would not make sense to have an outcome rating rated MS or higher together with low rating for Bank and performance. (If there was a positive shock that created the outcome then it can’t be attributed to the Bank.) However, if there was a negative shock (say earthquake) that eliminated the outcomes, we could still have Bank performance being rated MS or higher (they did a good job but could not have foreseen the earthquake).
Outcome MU- Bank Performance MS+	likely OK	
Efficacy Substantial + Efficiency Modest -	likely OK	It is very possible for a project to achieve outcomes but achieve them inefficiently; however, if the project did not achieve outcomes then efficiency would be quite unlikely to be high, because efficiency, in essence, compares project achievements to project costs.
Efficacy Modest - Efficiency Substantial +	Red flag	

18. Note on Canceled Operations

What is a Note on Canceled Operation (NCO)?

A Note on Canceled Operation (NCO) is prepared for a project that fails to become effective or is canceled before significant implementation is initiated.[21] The cut-off point for “significant implementation” is defined as final actual disbursement of less than 5 percent of the initial commitment or US$1 million (whichever is smaller), excluding any Project Preparation Facility and front-end fees. The NCO, which describes the project and explains why it was not implemented, is sent to the Board. The ICR guidelines also cover NCO requirements.

[21] This information is from appendix E of World Bank, OPCS 2011a.

Which sections of the ICR Review should be completed, and what ratings assigned?

The OPCS guidelines for writing ICRs (World Bank, OPCS 2011a, appendix E) do not indicate which ratings are to be completed for the NCO: “the text should generally follow the relevant sections of the ICR Guidelines and cover briefly the project’s rationale and objectives, main events and factors leading to cancellation, and any lessons learned. Special attention should be paid to the roles of the Bank and borrower with respect to design and implementation problems, their attempted resolution, and to Bank/borrower responses to any changed circumstances threatening the project.” (p. 49).
For the purposes of the ICR Review, sections 1, 2, 3, 8 (Bank quality at entry), 9 (government performance, if applicable), 10 (M&E design), 14 (lessons, if appropriate), and 15 (quality of the ICR) should be completed. The following ratings or subratings should be assessed:

- Relevance of objectives
- Relevance of design
- Bank quality at entry
- Government performance
- ICR quality (in this case, the NCO quality).

Rating the quality of the NCO

To enable an assessment of its quality, the NCO is expected to discuss the main events leading to cancellation, steps taken to resolve problems, exogenous factors, identification of causes, and parties responsible if the project failed, and the implications of failure. Above all, the purpose of the NCO is to clearly explain why the project was canceled; if the NCO does not convincingly explain or document the reasons for cancellation, the quality would be rated unsatisfactory.

Appendix A. ICR Review Checklist

{add add VA's panel reviewer checklist} **Overall**	
	Text is frank but diplomatic, with a neutral tone.
	Review is critical, not just a summary of the ICR.
	Statements are substantiated with evidence.
	Acronyms are minimized and spelled out the first time they're introduced
	No typos and/or grammatical errors (English fixed for non-English natives) {clarify; a task for panel reviewer}
Section 1: Project data	
	All required fields are populated.
	The loan or credit amount is less than or equal to the total project costs, at appraisal and closing.
	If there was any cofinancing, the cofinancers are mentioned; if cofinancers are mentioned, then there is some cofinancing recorded.
	If the actual is different from the planned, there is an explanation in section 2d.
	If there was additional financing (check section 2d), it has been added to the actual (not appraisal).
Section 2: Objectives and components	
	2a. The objectives from the legal agreement stated, with page numbers.
	2a. Key outcome targets are mentioned, if any; the section does not list all of the KPIs.
	2a. If the project's objectives were changed by formal restructuring (see section 2b), then the revised objectives are also presented, with the source and page numbers.
	2a. Both overarching and specific objectives are mentioned; if one phase of an APL, the APL objectives and the phase-specific objectives are listed.
	2b. If the project's objectives were changed due to a formal restructuring, the revised objectives should be in 2a, the relevance of objectives is assessed for the revised objectives (section 3a), achievement of the original and revised objectives is assessed (section 4), efficiency of the original and the revised project is assessed (section 5), and a weighted outcome rating is in section 6.
	2c. Components are listed and summarized with the planned and actual expenditure (for investment projects); any components added after approval are also listed with the same information.
	2d. Differences between the actual amount (total or the credit/loan) and the planned amount are explained (for example, cancellations, additional financing, favorable or unfavorable exchange rates, and so forth)
	2e. Extent to which the counterpart contribution was paid and the timeliness is discussed.

	2f. There is an explanation for any project extension(s).
Section 3: Relevance of objectives	
	Relevance of the objectives/outcomes is rated with respect to the current Bank country strategy document (CPF), at project evaluation or validation (whichever is later).
	The selection of the implementing agency, quality of preparation, M&E design, choice of indicators, project complexity, and other quality-at-entry material is *not* included.
	Relevance of objectives is rated on a 4-point scale.
Section 4: Efficacy	
	There is a separate subsection for each objective/outcome whose achievement is assessed. Each subsection contains a field for the wording of the objective/outcome (the heading), a field for the text explaining the assessment (the rationale), and a rating. If an objective was revised at restructuring, there is a separate subsection for each revision.
	The headings all represent outcomes or intermediate outcomes (the expected changes to result from the project), not outputs or components.
	The wording of the objectives/outcomes is taken from the legal agreement, in section 2a.
	If the statement of objectives was expressed in terms of outputs, there is a discussion and justification of the outcomes that will be assessed.
	If the statement of objectives has multiple outcomes and intermediate outcomes, there is a heading for each one.
	Under each heading, the assessment and evidence for the achievement of the entire results chain (outputs and outcomes) is presented, including intermediate outcomes.
	In each case, there is a discussion of attribution of the results to the project, other factors that might have affected the outcome beyond the project.
	There is an overall efficacy rating.
Section 5: Efficiency	
	If there is an ERR or net present value (NPV), the table showing the coverage of the ERR or NPV is completed (section 5a) and the ICRR addresses the assumptions and their realism or points to lack of transparency regarding the assumptions
	The Review presents evidence of cost-effectiveness and efficient use of project resources, efficient implementation, or efficient design.
	Improved efficiency of the sector is *not* included as evidence of project efficiency.
	There is a single rating for efficiency on a 4-point scale
Section 6: Outcome	
	The proposed outcome rating is consistent with the guidelines for combining relevance of objectives, efficacy, and efficiency.

	The rationale for the rating is explained in terms of results for the elements and *couched in the language of the harmonized criteria*. {clarify; inserting a sentence about moderate or significant etc shortcomings?}
	If the project was restructured to change the objectives, the outcome rating has been correctly calculated (assessing the outcome rating for the entire project with the original and revised objectives, then weighting the results according to the share disbursed before and after restructuring).
Section 7: Risk to development outcome	
	Identifies the major risks that could occur in the future, the likelihood that they may occur, and the consequences if they do. (For example, technical, financial, economic, social, political, environmental, ownership of government or other stakeholders, institutional support, governance, and exposure to natural disasters.) Starting July 1, 2017, there is no rating of risk to development outcome.
Section 8: Bank performance	
	Bank quality at entry incorporates comments from the relevance of objectives and M&E design, in addition to other criteria. Deficiencies in M&E design detract from the rating.
	Bank supervision incorporates comments from the M&E implementation section. Deficiencies in the latter reduce the rating.
	Shortcomings in the identification of safeguard issues or compliance are taken into account in the Bank quality-at-entry and Bank supervision ratings, as appropriate. (See section 11 on safeguards.)
	Shortcomings in fiduciary arrangements or performance are taken into account in the Bank's quality-at-entry and supervision ratings, as appropriate. (See section 11 on fiduciary arrangements.)
	The overall Bank performance rating is correctly calculated from the two subratings and, when split between the satisfactory and unsatisfactory scales, according to the outcome rating.
Section 9: Borrower performance is not applicable to ICRs initiated starting July 1, 2017	
Section 10: Monitoring and evaluation	
	On monitoring, the M&E design considers the choice of indicators, whether they are adequate to measure each element of the results chain, whether the implementation arrangements for M&E were identified, and whether there were adequate baseline data.
	On evaluation, M&E design discusses planned evaluations.
	The M&E implementation section comments on the extent to which the M&E plan was fully implemented and the data were collected in a timely manner, and were of good quality.
	The M&E utilization section provides evidence of the extent to which timely data were used to improve project performance and results on the ground.

	Overall M&E quality is rated on a 4-point scale; the three individual elements are not rated.
Section 11: Other issues	
	For investment loans, the safeguard category is mentioned, as well as the presence of a mitigation plan (if required), and whether the mitigation plan was successfully implemented. (Or the ICRR comments on lack of information in the ICR in this regard.)
	If unanticipated impacts are mentioned, they are truly unanticipated, adequately evidenced, attributable to the project, and of sufficient magnitude to be important.
Section 12: Ratings	
	When the IEG ratings differ from ICR ratings, there is an explanation in the last column.
Section 13: Lessons	
	There are no more than five lessons, and it is clearly noted whether they are from the ICR or from the IEG ICR Reviewer.
	The lessons clearly build on results evidenced elsewhere in the ICR Review.
	The lessons are not findings.
Section 14: Assessment recommended?	
	If an assessment is recommended, the reason given makes sense.
Section 15: ICR quality	
	Shortcomings in the ICR that were mentioned in the other sections are collected in this section.
	If the quality is rated modest or low, the explanation points to one or more “fatal flaws,” not a series of small errors.
	If the quality is rated high, the reasons are well documented and there should not be more than small, incidental shortcomings. (It shouldn’t simply say that it was frank and well-written.)

Appendix B. Guidelines for the IEG ICR Reviewer’s Meeting with the Last Project Task Team Leader {to be checked}

1. What is the purpose of the meeting?

The purpose of the meeting is twofold: (i) to gain a better understanding of the project experience to improve the accuracy and quality of IEG’s ICR Reviews; and (ii) to ensure due process by providing the project TTL and the IEG ICR Reviewer an opportunity to discuss the project experience. *The meeting is explicitly not intended to discuss any possible ICR Review ratings.*

This meeting is conducted before IEG sends the draft ICR Review to the Global Practice. The meeting with the TTL is different from the meeting that the Global Practice might request to discuss the draft ICR Review after receiving it from IEG (please see point 4 below for further details on the timing of the meeting).

2. Who should initiate and attend the meeting from IEG's side?

The ICR reviewer should initiate and attend the meeting from IEG's side. As a general rule, a new ICR reviewer (regardless of his/her seniority) should be accompanied by a more experienced reviewer to the meeting, in conducting his/her first ICR Review. It will be the responsibility of the hiring IEG staff (typically, the ICR Review coordinator) to determine when the new ICR reviewer is ready to plan and conduct his/her meetings with the last project TTL without assistance from more experienced reviewers.

Depending on who will be attending the meeting from the Global Practice's side, the relevant ICR Review coordinator or IEG manager may also choose to attend the meeting (please see point 3 below).

3. Who should attend from the Global Practice's side?

The meeting should be held with the last TTL of the project or, in the case of a programmatic series, the TTL of the final project. The meeting should not be held with the ICR author alone, unless the last TTL and the ICR author are the same person, or unless the last TTL specifically delegates to the ICR author responsibility for the meeting behalf of the Global Practice. In the rare instances when the last TTL of the project has left the Bank, the IEG ICR reviewer, upon consultation with the ICR Review coordinator, should contact the concerned sector manager for an alternative suggestion. It would be up to the project TTL to invite other Global Practice staff at his/her discretion.

The IEG ICR reviewer should ensure that she/he has advance notice of the complete list of attendees from the Global Practice. If the list includes the sector manager, the ICR reviewer should inform his/her manager, giving him/her the option of attending the meeting.

4. At what point during the ICR Review process should the meeting be conducted?

The meeting should be conducted only after the ICR reviewer has prepared an advanced draft of the ICR Review, and after the feedback on the first draft is received from the Panel reviewer. The ICR reviewer is expected to indicate in the relevant sections of the draft ICR that information will be sought to substantiate the assessment when submitting the draft to the Panel reviewer, along with the list of questions that he/she intends to ask.

In the rare instances when there is a delay in the availability of the last TTL, the ICR reviewer (with the concurrence of the ICR Review coordinator) may meet with the project TTL after obtaining the Panel reviewer's formal clearance and sign-off on the draft ICR Review. In such cases, it will be the responsibility of the ICR reviewer to obtain an e-mail indicating the Panel reviewer's concurrence with any substantive changes made to the draft ICR Review after the meeting with the project TTL and to include that e-mail in the ICR Review package going to the IEG manager.

5. What should the length of the meeting be?

The meeting should be between 30 minutes and 1 hour. This should also be clearly indicated to the project TTL in the e-mail inviting him/her to the meeting.

6. What should be discussed during the meeting?

The ICR reviewer should inform the meeting participant(s) that additional information obtained during the meeting as well as their comments may be used in the ICR Review. The ICR reviewer should focus on missing or ambiguous information in the ICR that is necessary to answer IEG's evaluative questions, including any additional evidence that may be needed to substantiate the ratings. In addition, the ICR reviewer should use the meeting to confirm his/her understanding of the project context, gain a better understanding of the factors that might explain the project's (good or bad) performance, and probe what the project TTL might have done differently had she/he had the option.

Under no circumstances should the ICR reviewer share the draft ICR Review, or even discuss the ICR Review ratings, at the meeting. However, the ICR reviewer should be totally responsive to any other questions, including questions about IEG evaluation methodologies, posed by the project TTL.

The ICR reviewer should record any additional information (not already contained in the ICR) obtained during the meeting with the project TTL in the draft ICR Review, noting "In a meeting between IEG and the project task team, the project team stated that…"

If an agreement is reached at the meeting that the project TTL will provide additional information to IEG, the ICR reviewer will confirm the specific additional information to be provided as well as the format in which it will be presented, both at the meeting itself and in a follow-up e-mail to the project TTL. With regard to the format, the ICR reviewer should make it clear that the additional information is to be provided in a separate note written specifically for this purpose, and not in the form of volumes of documents.

7. Should the meeting be conducted in person, by telephone, or by video?

Every attempt should be made to have a face-to-face meeting. If the project TTL is traveling for an extended period and/or based in the field, which would make a face-to-face meeting impossible (given time/budget constraints), the meeting should be conducted by telephone, personal conference line, video chat (e.g. Skype), or a similar means that does not involve additional cost.

8. What are the instructions if the last TTL is unresponsive to attempts to set up a meeting of if the TTL is traveling?

The ICR reviewer should prepare for this eventuality by contacting the project TTL and checking his/her travel schedule early in the process and getting the draft ICR Review ready soon after it is assigned. If the TTL is traveling, or if the TTL is based in the field office, an audio meeting should be suggested. If the TTL is not responsive after three weeks to setting up a meeting, the ICR reviewer may proceed with finalization of the draft. He/she should inform the Panel reviewer and the ICRR coordinator that repeated attempts to meet with the TTL were unsuccessful.

9. Should IEG circulate an agenda to the project TTL before the meeting?

Yes. Good practice is to develop a list of questions in advance, and to share this list with the TTL before the meeting.

10. Should a written record be kept of the discussion at the meeting?

It is mandatory to prepare a written summary of the general topics covered in the meeting. The names of attendees, date, and time of the meeting should be noted in the summary.

The summary of the meeting with the TTL must be attached to the Activity History in the IEG ICRR Portal record for this ICR Review.

If a meeting did not take place (because, for example, repeated requests for a meeting were not answered), the IEG ICR Reviewer enters a note explaining the reason for not having a meeting with the TTL in the Activity History.

11. How will the IEG manager sending the ICR Review to the Global Practice know that the meeting has taken place?

The Activity History contains the record of the ICR Review progress, including either the summary of the meeting with the TTL, or the note indicating the reason for not having a meeting with the TTL.

References

Hari Prasad, P. 2005. "Quality of Supervision in FY03–04 (QSA6): A QAG Assessment." Working Paper 46146. Washington, DC: World Bank. http://documents.worldbank.org/curated/en/155411468340864677/pdf/461460v20WP0QS1Box0334079B01PUBLIC1.pdf

IEG (Independent Evaluation Group; formerly OED). 2005. *Annual Report on Operations Evaluation 2004*. Washington, DC: World Bank. http://lnweb90.worldbank.org/oed/oeddoclib.nsf/24cc3bb1f94ae11c85256808006a0046/8e49d9d2a5f86e7a85256fc500565a1e/$FILE/2004_aroe.pdf

______. 2008. "Guidelines for IEG's ICR Review Meeting with the Project Task Team Leader." IEG Discussion Draft. Washington, DC: World Bank.

______. 2014. "Guidelines for Reviewing World Bank Implementation Completion and Results Reports: A Manual for Evaluators." IEG Discussion Draft. Washington, DC: World Bank. https://ieg.worldbankgroup.org/Data/ICRR_EvaluatorManualAugust2014.pdf

OECD (Organization for Economic Cooperation and Development/DAC (Development Assistance Committee). 2002. *Glossary of Key Terms in Evaluation and Results Based Management*. Working Party on Aid Evaluation. Paris: OECD. https://www.oecd.org/dac/evaluation/dcdndep/39249691.pdf

World Bank. 2007. *Quality at Entry in FY04–05 (QEA7): A QAG Assessment*. Washington, DC: World Bank. http://documents.worldbank.org/curated/en/114281468779431693/Quality-at-entry-in-FY04-05-QEA7-a-QAG-assessment

World Bank, OPCS (Operations Policy and Country Services). 2009. "DPL Appraisal Template." http://intresources.worldbank.org/INTOPCS/Resources/380831-1175264321504/DPLAppraisaltemplate0609.pdf

______. 2014. "Implementation Completion and Results Report: Guidelines." 1st edition, August 2006, last updated on 7/22/2014. Washington, DC: World Bank. http://intresources.worldbank.org/INTOPCS/Resources/380831-1175264321504/ICR_Guidelines_Updated.pdf

______. n. d. a. "Guidance Note Concerning the Responsibilities of Lawyers and Task Team Leaders (TTLs) in Respect of Loan Packages Submitted to the Board." Washington, DC: World Bank.

______. n. d. b. "Guidelines for the Structure and the Content of the PAD." Washington, DC: World Bank.

______. n. d. c. "Legal Issues Related to Processing of IL Operations Submitted to the Board." Washington, DC: World Bank.

World Bank OPCS/IEG. 2006. "Harmonized Evaluation Criteria for Implementation Completion and Results Report (ICR) and Independent Evaluation Group (IEG) Evaluations." World Bank and IEG, Washington, DC. https://ieg.worldbankgroup.org/Data/reports/ieg-wb-project-performance-ratings-codebook_092015.pdf {this does not seem like the correct link for this source}

World Bank OPSPQ. 2013a. "Investment Project Financing – Preparing the Project Appraisal Document (PAD)." April 9. World Bank, Washington, DC.

______. 2013b. "Investment Project Financing – Project Preparation Guidance Note." World Bank, Washington, DC. http://siteresources.worldbank.org/PROJECTS/Resources/40940-1365611011935/Guidance_Note_Project_Prepration.pdf