Guidelines for Reviewing World Bank Implementation Completion and Results Reports

A Manual for Evaluators

Last updated: August 1, 2014
Note: This is a living document under ongoing revision.

IEG
Independent Evaluation Group

Contents

Annexes

Abbreviations and Acronyms

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 environment objective
IEG	Independent Evaluation Group
ICR	Implementation Completion and Results Report
KPI	Key performance indicator
LA	Lending Agreement
M&E	Monitoring and Evaluation
NCO	Note of Cancelled Operation
NPV	Net Present Value
OPCS	Operations Policy and Country Services Vice Presidency of the World Bank
PAD	Project Appraisal Document
PD	Program Document
PDO	Project Development Objective
PPAR	Project Performance Assessment Report
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 main instruments of self-evaluation of the World Bank.[1] It is prepared by the World Bank at the close of every IDA or IBRD-funded operation or, in the case of a series of programmatic policy operations, at the end of a series of operations.[2]

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 operation.
- Capture and disseminate experience from the design and implementation of an operation in order to: (i) improve the selection of future interventions to achieve the goals of the Country Assistance Strategy (CAS); (ii) improve the design and implementation of future interventions through lessons learned; and (iii) help ensure greater development impact and sustainability of operations.
- Provide accountability and transparency at the level of individual operations with respect to the activities of the Bank, borrower, and involved stakeholders;
- Provide a vehicle for realistic self-evaluation of performance by the Bank and borrowers; and
- Contribute to databases for aggregation, analysis, and reporting, especially by the Independent Evaluation Group (IEG) on the effectiveness of lending operations in contributing to development strategies at the sector, country, and global levels.

The ICR assesses the extent to which the projects achieve their relevant objectives efficiently in the form of an Outcome rating. ICRs also rate the Risk to Development Outcome, the Bank's performance, and the Borrower's performance. They contribute to learning as well as accountability.

The audience for the ICR is both internal (the Board members and Bank managers and staff) and external (governments and their agencies, stakeholders, and beneficiaries in

[1] This material comes from Operations Policy and Country Services (OPCS), "Implementation Completion and Results Report Guidelines", updated May 25, 2010.

[2] ICRs are also prepared for Bank-executed projects that are totally funded by Global Environmental Facility (GEF) grants and for recipient-executed projects, funded by trust funds and managed by the World Bank

partner countries, as well as the general public). The final ICR is publicly disclosed at the time it is submitted to the Board unless otherwise decided in exceptional circumstances.

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 operation's design documents. Based on the evidence provided in the ICR and an interview with the last task team leader, IEG arrives at its own ratings for the project, based on the same evaluation criteria as the Bank[3]. At present, IEG reviews all ICRs of completed operations.

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 task team leader for the operation. 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 of the Bank's self-evaluation.[4]

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

ICR Reviews serve as an independent validation of the results 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 operations that closed from FY2011 onward, are posted on IEG's external website[5]. They are often useful as a starting point for IEG evaluators 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 and what are the main ratings?

The World Bank and IEG share a common, objectives-based project evaluation methodology for World Bank projects that assesses achievements against each operation's stated objectives, while also assessing the relevance of the objectives and design and the

[3] When insufficient information is provided by the Bank for IEG to arrive at a clear rating, IEG will downgrade the relevant ratings as warranted beginning July 1, 2006.

[4] For a sub-set of operations – on the order of 20-25 percent – IEG conducts Project Performance Assessments in the field.

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

efficiency of resource use in achieving the objectives. An advantage of this methodology is that it can take into account country context in terms of setting objectives that are reasonable; the Bank and the governments are accountable for delivering results based on those objectives.

There are four project ratings that IEG validates through the ICR review, and two that are issued by IEG only. The four main ratings are:[6]

- **Outcome**: the extent to which the operation's major relevant objectives were achieved, or are expected to be achieved, efficiently.[7]

- **Risk to Development Outcome:** is the risk, at the time of evaluation, that development outcomes (or expected outcomes) will not be maintained (or realized).

- **Bank Performance:** the extent to which services provided by the Bank ensured quality at entry of the operation 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.

- **Borrower Performance:** the extent to which the borrower (including the government and implementing agency or agencies) ensured quality of preparation and implementation, and complied with covenants and agreements, toward the achievement of development outcomes.

In addition, IEG provides two other ratings based on what is presented in the ICR:

- **Quality of Monitoring and Evaluation:** 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.

- **Quality of the ICR:** 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, conciseness, internal consistency, and the consistency with Bank guidelines.

Structure of this manual

The manual is organized into three parts, with annexes.

---

[6] Definitions of the four main ratings are from the OPCS/IEG Harmonized Evaluation Criteria (2006). These four ratings and IEG's monitoring and evaluation rating are also assigned in Project Performance Assessment Reports (PPARs).

[7] The project outcome rating is a measure of overall project *performance*, including relevance, efficacy, and efficiency; it is not a rating solely of the outcomes of the project (the results).

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

- Chapters 3-12 cover identification of the objectives and the criteria for the six main ratings.

- Chapters 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 – as well as a discussion of issues specific to assessing development policy operations. There is also a chapter on assessing cancelled projects, for which the Bank will issue a Note of Cancelled Operation, in lieu of an ICR. These Notes are also reviewed using the ICR Review form.

Three main Annexes present: detailed guidelines for completing the ICR review form (Annex A); frequently asked questions (Annex B – a work in progress); and the protocol for meeting the Task Team Leader of the project (Annex C).

Chapters 1-11, 13, 14, and 16 apply equally as guidelines for Project Performance Assessment Reports.

2. Procedures for the ICR Review

Responsibilities of the evaluator

The evaluator is responsible for:

- Correctly completing the ICR Review form (Figure 1) following the guidelines and procedures in this Manual and the specific instructions provided in Annex A, and assigning ratings based on the evidence in the ICR and in consulting other key documents (explained below).

- Meeting with the operation's last task team leader (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).

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

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

The review process

The ICR for an operation arrives in IEG after it has been sent to the Board. Either the ICR Review Coordinator or the sector Cluster Coordinator (depending on the unit) then assigns it to an evaluator for review. A blank ICR Review form for the project will automatically be created in the ICR Review database, based on its project ID (see Figure 1, a blank ICR Review form for an investment project[8]). Certain fields in the basic data portion of the form will be automatically populated (the project name, project ID, and sector codes, for example) and the name of the evaluator and the ICR Review coordinator will appear in the reviewer fields.

PREPARING THE INITIAL DRAFT

In preparing the first draft of the ICR Review, the evaluator is provided a package with several key documents but the evaluator 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 operation's original objectives and components. In the event that the legal agreement was amended, the amended agreement(s) will also be provided.
- The Project Appraisal Document (PAD, for investment operations) or Program Document (for development policy lending) – primarily for use in identifying the operation's original objectives, components, planned amounts, co-financiers, results framework, planned M&E and the presence of baseline information, safeguard category (for investment operations), and other aspects of design. If the project has been restructured, there will also be a Project Paper.
- The Country Assistance Strategy (at project closing) – primarily for use in assessing the operation's current relevance.

Implementation Completion and Results Report (ICR) – the main document for review, which is the Bank's self-assessment of the project. It includes information on: *revisions to the design* (restructuring, changes in objectives or components, changes in allocations, changes in co-financiers or expected counterpart contributions, and others); the *implementation of project activities*; the implementing unit's assessment of the project's *outcomes*, the *relevance of the* operation, the *achievement of its objectives,* the project's *efficiency*, and *safeguard and fiduciary compliance*; operational staff's self-*ratings (*on outcome, risk to development outcome, Bank performance, and borrower perfor-

[8] The ICR Review form for development policy operations is discussed in Chapter 14.

mance); and the lessons learned from the experience. These reports often also include in an annex an assessment by the Borrower and, occasionally, an assessment by co- financiers of the results, in addition to financial or economic analysis and the results of beneficiary surveys.[9]

Among the evaluator's most critical tasks is to identify the project's objectives – discussed in the next chapter and not always as straightforward as one would expect. The evaluator is expected to read all of the above documents and use that information as indicated in the explicit instructions in Annex A of this manual, for completing the ICR Review. Many of the specific guidelines are also featured in "help" buttons on the form that fully reflect the material in Annex A. Before finalizing the draft, however, the evaluator must contact the last task team leader (TTL) of the operation to set up an interview.

INTERVIEW WITH THE LAST TASK TEAM LEADER

This interview, which is conducted before the draft ICR Review is finalized, provides an opportunity for the last TTL to offer any additional views or information to the evaluator (beyond what is in the ICR) about the project experience, and for the evaluator 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.[10] The evaluator should not share the draft Review with the TTL, however, or share the proposed ratings. Following the interview, the evaluator writes a summary of the meeting for IEG's project file, copying it to the panel reviewer when he/she is identified. The detailed protocol for the TTL interview is in Annex C.

SUBMITTING THE DRAFT

Following the interview, the evaluator 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). With a final spell check, the draft is saved and submitted to the panel (via a button at the top of the form).[11]

---

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

[10] In the event that the IEG evaluator is a junior staff member, he/she is accompanied to the meeting by the Cluster Coordinator.

[11] The evaluator can go into the ICR Review form multiple times to edit it, but should save and exit ("verify form") each time to ensure that the work is not lost. **He/she must be very careful not to click on the "post" button, as this is not done until the very end of the process.**

Figure 1: ICR Review Form for Investment Projects

IEG **ICR Review**
Independent Evaluation Group

Date Created: 04/11/2011 02:41:45 PI
Last Updated: 04/11/2011 02:42:54 PI
Status: with Evaluatc

? - Help

1. Project Data:		**ICR Review Date Posted:**		
Country:	Cape Verde			
Project ID :	P074055		**Appraisal**	**Actual**
Project Name:	Growth And Competitiveness Project	**Project Costs (US$M):** ?		
L/C Number:	C3755	**Loan/Credit (US$M):**		
Sector Board:	FPD	**Cofinancing (US$M):** ?		
Cofinanciers: ?		**Board Approval Date:**		16
		Closing Date:	16	16
Sector(s):	General public administration sector (35%) Compulsory pension and unemployment insurance (20%) General finance sector (20%) General industry and trade sector (20%) General transportation sector (5%)			
Theme(s):	Other financial and private sector development (25% - P) Regulation and competition policy (25% - P) Legal institutions for a market economy (24% - P) Trade facilitation and market access (13% - S) Standards and financial reporting (13% - S)			
Prepared by:	**Reviewed by:**	**ICR Review Coordinator:**		**Group:**
Chad Leechor		Mark Sundberg		IEGPS2

2. Project Objectives and Components:

a. Objectives: ?

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

c. Components: ?

d. Comments on Project Cost, Financing, Borrower Contribution, and Dates: ?

3. Relevance of Objectives & Design:

a. Relevance of Objectives:

b. Relevance of Design:

4. Achievement of Objectives (Efficacy):

5. Efficiency:

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

	Rate Available?	Point Value	Coverage/Scope*
Appraisal		%	%
ICR estimate		%	%

* 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:

a. Risk to Development Outcome Rating:

8. Assessment of Bank Performance:

a. Quality at entry:

Ensuring Quality-at-Entry Rating:

b. Quality of supervision:

Quality of Supervision Rating:

Overall Bank Performance Rating:

9. Assessment of Borrower Performance:

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: [?]	ICR [?]	IEG Review	Reason for Disagreement/Comments [?]
Outcome:			
Risk to Development Outcome:			
Bank Performance:			
Borrower Performance:			
Quality of ICR:			

NOTES:
- When insufficient information is provided by the Bank for IEG to arrive at a clear rating, IEG will downgrade the relevant ratings as warranted beginning July 1, 2006.
- The "Reason for Disagreement/Comments" column could cross-reference other sections of the ICR Review, as appropriate.

13. Lessons: [?]

14. Assessment Recommended? [?] ○ Yes ○ No

15. Comments on Quality of ICR: [?]

a.Quality of ICR Rating:

Panel review

Once the evaluator presses the "Send to Panel" button, the ICR Review Coordinator will assign a reviewer from the IEG evaluation Panel, comprising cluster coordinators and very senior evaluators. The main task of the Panel reviewer is to review the same documents as the evaluator, read the ICR Review, ensure that the objectives have been properly identified, the guidelines have been properly applied, that the ICR Review is complete, internally consistent, and sufficiently critical of the quality of the data, and to comment on the ratings.

The Panel reviewer provides comments to the evaluator by clicking on a comment field within the ICR Review system. The comments will indicate areas of agreement, but also areas for improvement, areas of disagreement, and queries about the evidence. The system will automatically send an email with the comments to the evaluator. There is also a field in the system for the evaluator to respond. This discussion can go back and forth several times.[12] When the Panel reviewer is satisfied that the ICR Review is ready and there is agreement, he/she will clear it by clicking on a button at the top of the form.

While it is tempting for the discussions between the panel reviewer and the evaluator to take place face-to-face or by regular email, it is important that they be recorded in the ICR Review system, as this is the only way that the discourse, issues, and their resolution are retained for IEG's institutional memory and can be reviewed by the ICR Review Coordinator and IEG Manager.

Regional Review

Following sign-off by the Panel reviewer the ICR Review Coordinator or IEG Manager sends the draft ICR Review to the Country Director for comment.[13] The Country 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. For projects closing in FY11 or beyond, the region will have the option of inviting the Borrower to comment. The standard review period for the Region and the Borrower is 2 weeks (10 business days).[14]

---

[12] When the Panel reviewer and the evaluator cannot reach agreement, the ICR Review Coordinator may ask for a third opinion from the Cluster Coordinator or another Panel member.

[13] The draft Review is also usually copied to regional sector managers and key individuals in the anchor. This is the responsibility of the ICR Review Coordinator.

[14] The Country Director/Region may ask for an extension if necessary (for example if key staff are on annual leave or a mission and unavailable to reply).

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

FINALIZING AND POSTING THE ICR REVIEW

The evaluator should take on board any additional relevant information that is credible and will improve the accuracy of the assessment, but indicate in the text that this additional information was "provided by the region" (to distinguish it from the information in the ICR), and the source, if known.

Since the evaluator will have already fully assessed the project with respect to all of the information in the ICR, a response from the region that simply reiterates that same information would not be expected to result in changes in the ICR Review.

One rating that typically would not be affected by the Region's response is the rating of the Quality of the ICR. An unsatisfactory ICR Quality rating usually leads to substantial comments from the Region, but in this case there is no new evidence to bring to bear on the rating: the ICR itself is the only evidence required to assess its quality.

The Region's comments 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 saved, and the evaluator drafts a response to the region, also cleared by the Panel reviewer, to be forwarded by the ICR Review Coordinator back to the region. 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, the reasons why other information was not used (if this is the case), with the revised ICR Review attached.[15]

The ICR Review Coordinator or IEG Manager will then forward the response and the final ICR Review to the Region and will instruct a designated staff member (not the evaluator) to post the ICR Review. ICR Reviews of projects that closed prior to FY11 are not publicly disclosed. Following on IEG's new disclosure policy, ICR Reviews of all projects closing in FY11 and beyond will be publicly disclosed 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 Region will request a meeting with IEG. The evaluator, the panel reviewer, and either the ICR Review Coordinator or the Unit Manager generally attend these meetings. The purpose of the meeting is primarily to listen to the concerns of the oper-

[15] It is neither necessary or nor advisable to write lengthy, item-by-item rebuttals to the Region's comments.

ational 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, IEG should be in possession of the Region's written comments on the ICR Review.**

The IEG team is not expected to explain at these meetings 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 Region will be provided in summary form, with the source noted. The Region should not be providing reams of documents for IEG to sift through.

3. Identifying the objectives (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 of the elements of the project outcome rating are linked to the objectives – the relevance of the objectives, the relevance of the design to the objectives, whether the objectives were achieved (efficacy), and whether they were achieved efficiently. Accurately identifying the objectives therefore is essential to the entire evaluation exercise and critical for assuring accountability.

This chapter explains where the objectives can be found and guidelines for interpreting them for the purposes of the review. Identifying the operation's objectives is not as straightforward as one would expect. In practice, we often find:

- Differences between the articulation of the objectives in the PAD and the legal agreement (the lending, credit, or grant agreement)
- Differences the articulation of objectives within the PAD/Program Document
- Vague or unclear objectives
- Compound objectives, listing many outcomes in a single phrase
- The statement of objectives mentions only the operation's outputs, not the intended outcomes OR it contains outputs or intermediate outcomes mixed with outcomes.

Guidelines and examples are discussed below.

What constitutes an operation's objectives?

An operation'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 de-

liverable (output). In its guidelines for the content of the PAD, 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."[16] This applies equally to investment projects and development policy operations (DPOs).[17]

Where to find the objectives?

The evaluator should always assess the objectives reported in the original and revised legal documents or the Program Document – not from the ICR, which is being assessed.

For **investment operations**, the objectives of the project can be found in two documents -- the Project Appraisal Document (PAD, the document approved by the Board) and the lending/development credit/grant agreement (the legally binding document negotiated between the Bank and the government).

- In the PAD, the project development objective(s) (PDO) may be found in more than one place: in the front matter/summary; in the section on "Project Development Objectives;" and in the Technical Annex, "Detailed Project Description."

- In the lending, credit, or grant agreement, the objectives can be found in Schedule 2 at the end of the agreement, entitled "Project Description." If during the life of the project the objectives have been formally revised through a Board-approved restructuring, there should be an amended legal agreement in the pro-

[16] OPCS. Undated. "Guidelines for the structure and the content of the PAD" and OPSPQ 2013, p. 3. The OPCS/IEG Harmonized Criteria 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)." 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)).

[17] OPCS 2009b.

ject files, which would document any changes to the objectives of the operation.[18]

For individual **development policy operations (DPOs),** the objectives of the operation are supposed to be contained in the Program Document, under the heading "The Proposed Operation", and sub-heading "Project Description".[19] While there are occasional exceptions (for example, for sectoral DPOs), the legal agreements for DPOs typically do not include the objectives of the operation.[20]

DPOs that are part of a **programmatic series** – defined as a series of single-tranche loans, credits, or grants in support of a medium-term program – have objectives both for the individual operation and for the overall series. The entire series is conceptualized in the Program Document, with links between the operations in the series. *As of May 25, 2010, IEG is evaluating these operations against the objectives of the overall series, while noting the contribution of each operation to the outcome.* Therefore, the evaluator should note both the series and individual operations' objectives.

Which version of the objectives will be used?

INVESTMENT PROJECTS

For investment projects, the objectives in the legal agreement should be the basis for the evaluation. The statement of the project's development objectives in the lending agreement and the PAD are supposed to be identical, yet often they are not. The project team develops the PAD during appraisal, including the objectives, but during negotiations between the Bank and the government the wording can change and, in some cases, the objectives are substantively altered – by dropping or adding an objective, for example. If the PAD is not revised to reflect these changes before it is submitted to the Bank's Board, then the objectives approved by the Board and those negotiated with the borrower may differ.

Recently, OPCS has been working with the Regions to make sure that the objectives in the PAD and the lending agreements are "materially consistent," and a number of operations have been restructured to bring the objectives into agreement when they were not. Instructions have been disseminated, delineating the responsibility of the legal de-

[18] While a project must be restructured to formally change its objectives, not all restructurings involve changes in the objectives.

[19] OPCS 2009a.

[20] The legal agreements for DPOs typically refer to support for a program detailed in the government's Letter of Development Policy. This letter is sometimes an annex of the legal agreement, but is usually an annex to the Program Document.

partment and the TTL to ensure this consistency[21] and the new guidelines for project preparation emphasize the need for consistency between the legal agreement and the PAD (OPSPQ 2013b, p. 28). Nevertheless, there remains a backlog of projects approved many years ago that are closing for which there are different versions of the objectives, even though there has now been ample time to restructure the projects to clarify the objectives and bring the differing versions into alignment.[22]

For the purpose of reporting the objectives in the ICR Review, the evaluator should quote both the objectives in the PAD and the objectives in the lending agreement. If they are different, both should be quoted, with the comment that the objectives in the legal agreement will be the basis for the evaluation. If they are identical, then only the objectives in the legal agreement need be reported, with the explanation that the PAD objectives are identical.

DEVELOPMENT POLICY OPERATIONS

For development policy operations, the preferred statement of objectives (if there are no objectives stated in the lending agreement) is in the section entitled "The Proposed Operation," under the sub-heading "Project Description." This is what is advocated in OPCS guidelines for writing a Program Document.

Unfortunately, the Program Documents for many operations do not follow OPCS's guidelines, and there are often many different formulations of the objective within the Program Document. There is the added confusion between the objectives of the operation or series of operations (often referred to as "the program") and the objectives of the government program (also often called "the program"), where often the operation supports only part of the government's program. In the instance that the objectives of the operation are not presented in the "Project Description" section of the Program Document, the protocol for identifying the objectives is as follows:

1. <u>Cross-check with:</u> (a) the Summary sheet at the front of the Program Document; (b) the government's Letter of Development Policy (usually an annex to the Program Document); and (c) the Policy Matrix. (Note: the Policy Matrix may have many "sub-objectives" under each pillar or broad grouping - do not use those multiple sub-objectives as the operation's objectives. Just check to see whether there's an articulation of the overall objective.) You may also look elsewhere in the text of the doc-

[21] OPCS, "Legal Issues Related to Processing of IL Operations Submitted to the Board" and "Guidance Note concerning the Responsibilities of Lawyers and Task Team Leaders (TTLs) In respect of Loan Packages Submitted to the Board", both undated.

[22] There is no guidance for Bank staff in the OPCS ICR guidelines, which instruct the author to take them from the PAD "and as reflected in legal documents". However, the ICR guidelines do not explain how to reconcile the two sets of objectives if they are different. Recent communications from OPCS indicate that the Lending Agreement is given greater weight when there are material differences, as these are the objectives that are legally binding.

ument in other unlikely places (there may be other statements strewn throughout the report that are enlightening).

2. <u>When there are different articulations of basically the same objective, the articulation that is most outcome-oriented is the preferred one.</u>

3. <u>In the event that the objectives throughout the Program Document merely say that the objectives are to support the government Program, then use the objectives of the government program.</u> *However, it is often the case that the operation supports a sub-set of the government program's objectives, so be careful only to include those broad objectives of the program actually supported by the operation*.

For the purpose of reporting the objectives in the ICR Review, the evaluator should quote the objectives in the lending agreement (if any) and those in the Program Document. If project objectives do appear in the lending agreement, then they will be used as the basis for the evaluation (but this will be rare). Otherwise, it will be based on the objectives in the Program Document. <u>When there are multiple versions of the objective in the Program Document, list them in the ICR Review with the page numbers, explain which version has been used, and why.</u> <u>If there are remaining ambiguities or inconsistencies about the objectives, consult with the ICR Review Coordinator.</u>

<u>**In the event that the ICR is for a programmatic series of DPLs, the objectives of the series as well as the objectives of the individual operations in the series should be cited, with the source.** The ICR Review will conduct the assessment based on the objectives of the series, while noting the contribution of each operation to the results.</u>

"Overarching" vs. "specific" objectives

When there are both overarching and specific objectives, both should be reported.[23] This is very likely to be the case for operations that are part of Adaptable Program Loans, in which a sequence of 2-3 investment operations are offered over roughly a 10-year period, with objectives for each Phase and for the overall APL.

In the case of the Lesotho Health Sector Reform Project, Phase 2, for example:

> The objectives of the APL are "to achieve a sustainable increase in access to quality preventive, curative, and rehabilitative health care services in Lesotho."
>
> The objectives of this project (Phase II), as stated in the Development Credit Agreement (DCA) and the Project Appraisal Document (PAD), were "to assist the Borrower in achieving a sustainable increase in access to quality preventive, curative, and rehabilitative health services by increasing access to, and quality delivery of, essential health services."

[23] Overarching objectives are not to be confused with "higher level objectives to which the project contributes", which are highlighted in some Project Appraisal Documents. These should not be reported.

Formally revised objectives

If the project's development objectives have been formally revised (through a restructuring approved by the Board and resulting in an amended legal agreement), the ICR Review should record both the original and formally revised objectives.[24] The formally revised objectives can be taken from the amended legal agreement and both sets of objectives will be assessed in the ICR Review. Occasionally an ICR will report that objectives were changed during the course of the project, but not formally approved by the Board. These informal revisions are not taken into consideration in the ICR Review; IEG only assesses objectives reflected in the legal agreement and Board documents.

For example, the objectives of the *St. Kitts and Nevis HIV/AIDS Prevention and Control Project* were formally revised:

> The ***original* objectives**. According to the Loan Agreement (LA), the overall objective was to control the spread of HIV/AIDS with three specific objectives: (i) scaling up programs for prevention, care and control of the epidemic targeted in particular to high-risk groups; (ii) heightening awareness with respect to infection and prevention amongst the population; and (iii) strengthening the institutional capacity of the Ministry of Health (MOH), other related government agencies and civil society organizations to ensure the effectiveness and sustainability of the project.
>
> The project was formally restructured, with the following ***revised* objectives**: to assist the Borrower to control the spread of HIV/AIDS and to mitigate its impact through the following specific objectives: (i) scaling up prevention services for high risk and vulnerable groups and the general population; (ii) expanding and strengthening treatment, care and support for PLWHAs and mitigating its impact on infected and affected persons; and (iii) strengthening the institutional capacity of the MOH, other government agencies, and civil society organizations (CSOs) to ensure an effective multi-sectoral response to the epidemic.

In some cases, the formal revision involves dropping an original objective or adding a new objective, while the other original objectives remain the same.[25]

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 Project Development Objectives. Both the PDOs and the GEOs from the Project Appraisal Document should be listed on the Review form, in addition to

---

[24] There are two levels of restructuring. Level One applies to modifications in a project's Development Objectives or changes in the safeguard category from a lesser category in the appraisal document to a Category A, or a trigger of a new safeguard policy. Level One restructurings are submitted to the Board for approval under absence of objection procedures. Level Two applies to all other modifications to the project, including changes in outcome indicators or targets, modifications in project scope or design, addition or cancellation of components, reallocation of proceeds, new loan closing date, and others.

[25] The instructions for assessing the outcome of projects with formally revised objectives are in Chapter 7 of this manual.

the objectives noted in the Grant/Legal Agreement. However, the project will be assessed based on the articulation in the Grant/Legal Agreement.

4. Relevance of Objectives and Design (Section 3)

Relevance, along with efficacy and efficiency, is one of the three criteria underpinning the Outcome rating. The assessment of relevance comprises two parts, each part assessed separately: (a) the relevance of the project's objectives and (b) the relevance of the project's design to the objectives. Each should be discussed and rated under a separate heading. There is no overall relevance rating.

Relevance of objectives

DEFINITION

RELEVANCE OF OBJECTIVES is the extent to which an operation's objectives are consistent with the *country's current development priorities*[26] and with *current Bank country and sectoral assistance strategies and corporate goals* (expressed in Poverty Reduction Strategy Papers, Country Assistance Strategies (CASs), Sector Strategy Papers, and Operational Policies.) "Current" in this case refers to the time of project closure.

CRITERIA

Relevance of objectives is assessed with respect to: (a) country conditions; (b) *current* (at the time of project closing) World Bank and government strategies; and (c) the framing of the objectives and their ambitiousness.

GUIDELINES

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

For projects with formally revised objectives, the relevance of both the original and revised objectives *at the time of project closing* should be assessed. The "original objec-

[26] And current global priorities for projects funded by the Global Environmental Facility (GEF).

tives” are all of the objectives as originally approved/articulated in the lending agreement; the “revised objectives” are all of those that were changed or added minus those that were dropped through a formal restructuring of the project, plus those that were unchanged.

Occasionally, the CAS at closing does not refer to the particular project objectives because the project achieved its objectives fully and the Bank's strategy shifted elsewhere. In these cases, the ICR Review needs to make an assessment of the extent to which the *outcomes* (that were embedded in the project's objectives and were achieved) are relevant to the CAS at project closing. For example, if a project aimed at eliminating marine pollution succeeded in fully achieving its objectives, and the CAS at project closing, therefore, 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.

RATINGS

Relevance of objectives is rated on a four-point scale: *High, Substantial, Modest, or Negligible.*

Relevance of design

DEFINITION

The relevance of project design is defined as the extent to which the project’s design (its planned activities or policy areas) is consistent with the stated objectives, including an assessment of the Results Framework.

The *Results Framework* represents the underlying project logic linking the project’s inputs and outputs to the outcome(s) that the project seeks to achieve. This is distinct from the design of M&E, which includes measurable indicators that enable tracking of all key links in the causal chain, and arrangements and responsibilities for data collection, analysis, and utilization of the data; the relevance of design does not include an assessment of the indicators or of M&E design.

CRITERIA

The relevance of design is assessed with respect to two elements: (a) the relevance of project design (activities, components, policy areas) to the objectives; and (b) the quality of the results framework. Three important questions for the results framework are:

- Was there a clear statement of objectives, linked to intermediate and final outcomes?
- Was the causal chain between funding and outcomes clear and convincing?
- Were exogenous factors and unintended (positive and negative) effects identified?

Poor relevance of design may reflect a project design, in terms of project components or policy areas, that is not consistent with the project's stated objectives – for example, the absence of key activities or policy areas necessary to achieve the objectives or inclusion of irrelevant or extraneous activities or policy areas.

In addition to the relevance of the activities or policy areas to achievement of the objectives, the choice of lending instrument (for example, investment or development policy operation) can also enter into the relevance of design.

RATINGS

Relevance of design is rated on a four-point scale: *High, Substantial, Modest, or Negligible*.

RELATION BETWEEN RELEVANCE OF DESIGN AND QUALITY AT ENTRY

Relevance of design is often confused with the Bank's Quality at Entry, which is assessed as part of the Bank's performance (see Chapter 9). There is overlap – relevance of design is one dimension of Quality at Entry, but Quality at Entry includes also the Bank's identification, preparation, and appraisal of the operation. These items are not appropriate for the section on relevance of design.

5. Achievement of the Objectives (Efficacy, 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 operation.

For the purposes of this section, the objectives refer to each of the key <u>outcomes</u> indicated in the statement of project development objectives from the Legal Agreement (Credit/Lending/Grant Agreement) in the case of investment projects or the Program Document in the case of DPOs, as discussed in Chapter 3 of this manual and presented in Section 2a of the ICR Review form.

Criteria

The achievement of each objective is assessed based on the level of achievement and concept of "**plausible causality**". This is established for each objective by:

(a) assembling the evidence from the ICR for each part of the results or **causal chain** supported by the project – the inputs and outputs – and the observed intermediate outcomes or impacts for each objective; and

(b) discussing and (to the extent feasible) presenting evidence from the ICR of the contribution of other, non-project factors in leading to these outcomes (**the counterfactual**), with the intent of pointing to the plausible attribution of the outcomes to the government program or project supported by the Bank.[27]

Organizing the Assessment of Efficacy

In Section 4 of the ICR Review, the evaluator will create a heading for each of the outcomes to be achieved in the statement of the objectives.[28] Under each heading, the task of the evaluator is to: (a) assemble the evidence from the ICR (or field work in the case of a Project Performance Assessment Report) documenting 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 supported. These two elements are discussed below, in turn. A rating will be assigned to each objective.

Distinguishing between outcomes and outputs

As noted at the beginning of this chapter, according to OPCS guidelines for PADs, the statement of objectives is supposed to "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."[29] Yet, this advice is often not adhered to and the articulation of the objectives may include multiple outcomes and in many cases all of the components, which are ba-

[27] The importance of attribution is equally supported by OPCS guidelines: "While the Results Framework provides the grounds for judging achievement, it may be helpful to expand in the text on causal relationships between the Bank's intervention and outcomes as distinct from other causal factors (e.g., other interventions, policy changes unrelated to the operation, natural events, and market factors). A common weakness of past ICRs has been failure to establish causal linkage between the operation and claimed benefits – to the exclusion of other events that might have generated benefits for the same target groups."

[28] This instruction holds equally for the chapter of Project Performance Assessment Reports on "Achievement of Objectives".

[29] OPCS (undated), "Guidelines for the structure and content of the PAD."

sically the means to an end. Thus, in addition to sorting out discrepancies between the objectives statement in the legal agreement and the PAD, the evaluator is also faced with distinguishing between outcomes and outputs in the organization of the assessment of the achievement of objectives in Section 4 of the ICR Review, which should be organized according to the outcomes that the objectives seek to achieve.[30]

Compound objectives with multiple outcomes can be "unpacked" for the purposes of the assessment. An example of a compound objective is when several outcomes are linked together in a single sentence. For example, here are the objectives of the *India Third Technical Education Project*, cited in Section 2 of the ICR Review:

> 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 evaluator 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: (a) expand the capacity of technical education; (b) improve the quality of technical education; (c) improve the efficiency of technical education; (d) meet the economic needs of targeted states/UT; and (e) increase access to technician training by disadvantaged groups.

Components, activities, and project outputs should not be included as objectives, even if they appear in the same sentence as the other objectives. The statement of objectives of the Kyrgyz Agricultural Support Services project incorporates both the problem of 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 <u>by means of</u>: 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 three outcomes that the project seeks to affect: agricultural productivity, profitability, and sustainability. These three 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

[30] The distinction between outcomes and outputs is also important for the evaluator in organizing the discussion of the achievement of objectives in Section 4 of the ICR Review.

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.

Establishing a results chain or causal framework for each objective

The results chain or causal framework 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. The different parts of the causal chain for achieving an objective can be described as follows:[31]

The objective of interventions is to improve the welfare **outcomes** of the target group, including longer-term effects or outcomes of an intervention, based upon the achievement of…

- **Intermediate outcomes**, such as increased consumption of goods and services, which are short- and medium-term changes or effects as a result of…
- **Outputs**, such as the goods produced, activities, services delivered, or actions taken as a result of …
- **Inputs**, such as the financing, organizational capacities, and human resources provided by the intervention or program.

Table 2 illustrates the links in the causal chain for a program to improve child health among the poor. Impact evaluation and related research are needed to verify the links in the causal chain underlying the logic of project design.

Table 2. The Causal Chain for a Child Health Program for the Poor

[31] Source: OED 2004..

Table 1.1	The Causal Chain for a Child Health Program for the Poor
Causal chain	**Performance indicators (to be disaggregated by poverty status and gender)**
Outcomes[a]	Infant and child mortality
	Morbidity
	Stunting (height/age) and wasting (weight/height)
	Total fertility rate
Intermediate outcomes	Access to potable water
	Percentage of women who know how to use oral rehydration salts correctly
	Contraceptive prevalence rate
	Number of children using supplements and oral rehydration salts
	Family planning services utilized
Outputs	Number of trained health staff
	Number and coverage of children in well-baby clinics
	Number and coverage of immunization campaigns/services
	Family planning services/contraceptives supplied or sold
	Vaccination campaigns/quantity vaccines delivered
	Nutritional supplements and oral rehydration salts distributed
	New institutions and strategies established
Inputs	Spending on maternal and child health programs
	Private and donor spending on services
	Public policy/regulatory environment
	Strategic planning

Note: Exogenous factors such as household income, food prices, and educational level of the mother will also be important determinants of program outcomes and should be accounted for, in the discussion of the counterfactual.
Source: OED. 2004. *Annual Review of Operations Evaluation*.

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 are also affecting these same 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.

It is rarely the case that the programs supported by the Bank are 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 what parts of a program "worked".

However, most of the projects or programs supported by the Bank involve large-scale and multi-faceted interventions or country or sector-wide policies for which establishing an airtight counterfactual as the basis for attributing outcomes to the project would be difficult, if not impossible.

For the purposes of understanding efficacy, for each objective the evaluator should nevertheless identify and discuss the key factors outside of the project that plausibly might have contributed to or detracted from the outcomes, and any evidence for the actual influence of these factors from the ICR.[32]

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 achieved outcomes can plausibly be attributed to the supported project or program:

- A timeline of key events, showing the relation 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 that the two areas may differ in baseline characteristics and other factors that may be affecting the two types of 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.

How to treat overarching objectives and objectives across operations

When an investment operation has both an overarching objective (such as those for multiple phases of an Adaptable Program Loan or a series of projects)[33] and specific objectives, achievement of each of the specific objectives should be discussed with a comment on (and reasons for) the likelihood of achievement/non-achievement of the overarching development objective.[34] In the case of development policy operations in a

---

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

[33] "A series of projects (SOPs) may be designed to support a single borrower..as part of a program consisting of a series of two or more projects; or multiple borrowers who are facing a common set of devleopment issues or share common development goals." OPSPQ 2013a, p. 4.

[34] OPCS 2011a, p. 24.

series, they are assessed against the series objective, noting the contribution of individual operations (see Chapter 14).

Ratings

The efficacy of each objective (outcome) is rated on a 4-point scale: *High, Substantial, Modest, Negligible*. No overall efficacy rating is required, nor should the evaluator provide one. These ratings are defined as follows:

- **High:** The project exceeded, or is likely to exceed, its objective (intended outcome).
- **Substantial**: The project achieved or nearly achieved its objective (intended outcome), or is likely to do so.
- **Modest**: The project partly achieved or is expected to partly achieve its objective (intended outcome).
- **Negligible**: The project did not achieve nor is it expected to achieve its objective (minimal achievement, if any).

For all ratings, the evaluator needs to assess whether the outcomes achieved are attributable to the operation. When the desired outcome is achieved but there is evidence of a weak results chain and/or that the results are primarily due to other factors, the rating should be adjusted downward, accordingly.

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

- If the anticipated outcome was met or exceeded, but there is evidence that the change was due mainly (or solely) to external factors, a rating of Modest (or Negligible) may be warranted.
- If the outcome deteriorated, falling short of target, but there is evidence that the decline would have been much worse in the absence of the project, a rating of Substantial (or High) could potentially 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 evaluator/ICR to show that improved outcomes were the result of the project and that declining outcomes were not the result of the project.

6. Efficiency (Section 5)

Definition

Efficiency is a measure of how economically resources and inputs are converted to results.[35] It asks whether the costs involved in achieving project objectives were reasonable in comparison with both the benefits and with recognized norms ("value for money").[36] Was the project implemented at least cost? *Efficiency is assessed only for investment-type operations, including technical assistance loans, but not development policy operations.*

Guidelines

This section should report on all available measures of efficiency both ex-ante and ex-post. The analysis should discuss both the traditional measures of efficiency (as applicable and practical)—e.g., 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.[37] 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).

In the event that an ERR/IRR has been calculated, the assumptions should be fully explained and transparent in the ICR. Any data gaps and methodological strengths or weaknesses in the Bank's assessment of efficiency should be noted.[38] The reviewer

---

[35] OECD/DAC 2002.

[36] OPCS 2011a, p. 24.

[37] 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).

[38] The Bank's ICR Guidelines, Appendix G, advocate that if an ERR was calculated at project appraisal, it should also be used for completion reporting. The ICR "should indicate what it was in the Project Appraisal Document (PAD), what it is when re-estimated at completion, and on what percent of total project costs the original and revised estimates were based. A re-estimate should be made even if the ICR is prepared so early in the project operation stage that only a short period of actual benefits can be observed and future investments may still be required. For any new analysis, the revised NPV or ERR at least provides actual latest cost figures and an updated projection of benefits, reflecting changes made during implmenetation. Underlying assumptions about costs and benefits, and other information supporting the

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 operation 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.[39]

The project's efficiency should not be confused with the 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, the repetition and dropout rates in the education system may decline as the result of an education investment, an indication of improved internal efficiency of the 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.

Ratings

Efficiency should be assigned an overall rating, based on a 4-point scale: Negligible, Modest, Substantial, or High.

7. Project Outcome Rating (Section 6)

Definition

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

Thus, the Outcome rating – a measure of the project's overall performance – is derived from the prior assessment of the relevance of objectives and design, efficacy in achieving each objective, and efficiency.

Guidance

The IEG/OPCS Harmonized Evaluation Criteria 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, while restructuring provides opportunity to

---

analysis (e.g., output volumes, major cost items, or prices) should be presented." (ICR Guidelines, Appendix G, "Economic and Financial Analysis", p. 53.)

[39] OPCS 2011a, p. 31 (also in OPCS/IEG 2006).

internalize positive unintended results, or they can be taken into account as additional achievements if convincingly documented.

For evaluation purposes, an operation'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 evaluator from key associated outcome targets (and/or the operation's design features as relevant) (See Chapter 3). The rating of Outcome should encompass the extent to which the operation's institutional objectives 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 operation 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 an operation's objectives, design, or implementation are inconsistent with the country's current development priorities and with current Bank country and sectoral assistance strategies and corporate goals (expressed in PRSPs, CASs, SSPs, OPs).[40]

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

Rating Scale

Highly Satisfactory	There were <u>no</u> shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance.
Satisfactory	There were <u>minor</u> shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance.
Moderately Satisfactory	There were <u>moderate</u> shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance.

[40] The ICR Guidelines and IEG practice have been identical on this point for some years. 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.

Moderately Unsatisfactory	There were significant shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance.
Unsatisfactory	There were major shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance.
Highly Unsatisfactory	There were severe shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance.

Deriving the outcome rating from IEG's sub-ratings

For the purpose of assuring consistency across evaluators, IEG has developed guidelines for deriving the project outcome rating from the sub-ratings on relevance, efficacy, and efficiency in the previous sections. It is difficult to define every contingency, but IEG's guidelines derive from two general cases:

Case #1: Meets the following criteria:

(i) relevance of design=relevance of objectives

(ii) efficacy of objective 1 = efficacy of objective 2 …=efficacy of objective n.

In other words, the efficacy ratings are uniform, and the relevance ratings are uniform. In effect, there are overall ratings for relevance, efficacy, and efficiency, though the three ratings may differ. Guidelines for the derivation of the outcome rating in this case are in Table 2.**Case #2:** The ratings for relevance of objectives and design may differ or the ratings on different objectives may vary, so there's no longer an implicit overall rating and the variation in ratings within relevance and/or efficacy must be taken into account. In this instance, the evaluator will have to derive the outcome rating by using Table 2 and interpolating across the guidelines for different ratings.

CASE #1: WHEN THE SUB-RATINGS FOR RELEVANCE MATCH AND THE SUB-RATINGS FOR EFFICACY OF EACH OBJECTIVE MATCH

The Guidelines in Table 2 apply when the two sub-ratings on relevance are the same and the sub-ratings on all of the objectives are the same. In this case, implicitly there is an overall relevance and efficacy rating that can be considered along with efficiency.

Consider the guidelines for a project for which the outcome is rated fully **satisfactory** (Table 1). This would apply when all of the sub-ratings are substantial (relevance of objectives, relevance of design, efficacy of all of the objectives, and efficiency), which are interpreted as the equivalent of "minor shortcomings" in the Harmonized Criteria. Satisfactory would also apply when two criteria are substantial but the third is rated high, or when efficacy is rated substantial but the other two ratings are high.

Table 1 Definition of a Satisfactory Outcome when Sub-Ratings Match

Rating	Definition	Relevance of objectives & design	Efficacy	Efficiency
Satisfactory	Minor shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance	Substantial on all three criteria		
		Substantial on two criteria, high on the third		
		Substantial efficacy but high relevance and efficiency		

Source: Table 2.

A **moderately satisfactory** outcome is assigned when two of the criteria are substantial (or high), one of which must be efficacy, while the third is modest, the equivalent of "moderate shortcomings." (Table 3) Thus, moderately satisfactory is interpreted as modest on at least one of the sub-ratings.

Table 2 Guidelines for Investment Projects when relevance sub-ratings are the same and the efficacy of each objective is the same

Rating	Definition	Relevance	Efficacy	Efficiency	Comment
Highly satisfactory	There were <u>no</u> 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 <u>minor</u> 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 <u>moderate</u> 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 <u>significant</u> 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 <u>major</u> 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 <u>severe</u> 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.*

Table 3 Definition of a Moderately Satisfactory Outcome when Sub-Ratings Match

Rating	Definition	Relevance of objectives & design	Efficacy	Efficiency
Moderately Satisfactory	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.		

Source: Table 2

A **moderately unsatisfactory** outcome results from modest on any two criteria and substantial or higher on the third criterion, or modest efficacy, either of which is interpreted as "significant" shortcomings (Table 4), while a fully **unsatisfactory** outcome ("major" shortcomings) is defined as modest on all three criteria, or negligible on one and higher on the other two (Table 5). These four cases plus the highest and lowest ratings (highly satisfactory – no shortcomings -- and highly unsatisfactory – severe shortcomings) are all presented in Table 2.

Table 4 Definition of a Moderately Unsatisfactory Outcome when Sub-Ratings are not Split.

Rating	Definition	Relevance of objectives & design	Efficacy	Efficiency
Moderately Unsatisfactory	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; modest efficacy and substantial (or high) on the other two criteria.		

Source: Table 1

Table 5 Definition of an Unsatisfactory Outcome when Sub-Ratings are not Split

Rating	Definition	Relevance of objectives & design	Efficacy	Efficiency
Unsatisfactory	Major shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance	Modest on all three criteria Negligible on one criterion and higher on the other two		

Case #2: When the sub-ratings within relevance and efficacy do not match

It is often the case that the sub-rating for relevance of objectives and relevance of design are not identical, and even more common that there are several objectives that weren't equally achieved. Consider the four examples in Figure 7.5. How would one determine the outcome rating in these cases where the sub-ratings are split within relevance or efficacy?

Figure 7.5: Examples of Split Sub-Ratings within Relevance and Efficacy

Example	Relevance		Efficacy		Efficiency
	Objectives	Design	First objective	Second objective	
1	Substantial	Modest	Substantial	Modest	Substantial
2	High	Negligible	Substantial	Modest	Substantial
3	Substantial	Substantial	High	Modest	Modest
4	High	High	High	Substantial	Substantial

One possible way of dealing with these examples of split ratings would be to "average" the ratings within relevance and efficacy – essentially assigning an overall relevance and overall efficacy rating -- and then using Table 1 to determine the outcome. However, in instances like the first example in Figure 7.5, the outcome rating could be as high as Satisfactory (if sub-ratings for both criteria are rounded "up") or Moderately Unsatisfactory (if sub-ratings for both criteria are rounded "down") or Moderately Satisfactory (if one is rounded "up" and the other rounded "down")(see Figure 7.6). While there may be good reasons for weighting the results for one objective more than for the other, this can seem arbitrary and would lead to additional inconsistencies across evaluators.

Figure 7.6: "Averaging" Split Sub-Ratings within Relevance and Efficacy Doesn't Necessarily Lead to Consistency

Averaging rule	Relevance		Efficacy		Efficiency	Outcome rating?
	Objectives	Design	First objective	Second objective		
Original Sub-ratings	**Substantial**	**Modest**	**Substantial**	**Modest**	**Substantial**	**?**
If both relevance and efficacy are rounded "up"	Substantial		Substantial		Substantial	**Satisfactory**
If one is rounded "up" and the other "down"	Modest		Substantial		Substantial	**Moderately Satisfactory**
If both are rounded "down"	Modest		Modest		Substantial	**Moderately Unsatisfactory**

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."*

The Harmonized Criteria also imply that a shortcoming in relevance can be due either to objectives or design: *"Shortcomings in relevance may have to do with the extent to which an operation's objectives, design, or implementation are inconsistent with the country's current development priorities and with current Bank country and sectoral assistance strategies and corporate goals…"*

What is the evaluator to do?

- **IEG assumes that all objectives are equally important, unless the lending agreement or PAD explicitly indicates their relative importance**.[41]
- When the sub-ratings within relevance and within efficacy are variable, this should be taken into account (along with the efficiency rating) in determining whether there are minor, moderate, significant, major, or severe shortcomings.
- The evaluator should interpolate between ratings, using Table 2 as a starting point. For example, in the case considered in Figure 7.6, there are moderate shortcomings in part of relevance and part of efficacy, suggesting that the outcome isn't fully Satisfactory. Yet, because relevance of the objectives and the efficacy of one of the objectives were rated substantial, it is clearly better than a Moderately Unsatisfactory outcome. If the two objectives were equally important, a Moderately Satisfactory outcome rating would seem reasonable.
- To take another common example, the relevance of objectives and design and efficiency may be substantial, but the efficacy ratings may be split, with one objective modestly achieved and the other one substantially achieved. In this case, again, a Moderately Satisfactory outcome rating would apply. If both objectives were modestly achieved, then the project would be Moderately Unsatisfactory.
- When insufficient information is provided by the Bank for IEG to arrive at a clear rating, IEG will downgrade the relevant ratings as warranted beginning July 1, 2006.

Rating the Outcome of Projects with Formally Revised Objectives

These guidelines concern a subset of restructured projects, namely those whose objectives have been formally revised. The outcome of restructured projects where neither the Project Development Objectives (PDO) nor key associated outcome targets have

[41] The relative importance of objectives is never measured by the resources allocated to them, as some activities are inherently more expensive than others (for example, civil works vs. institutional reforms).

been formally revised will continue to be assessed against the (unchanged) original project objectives and targets.

Note that in many cases the restructuring involves dropping an original objective or adding an objective not in the original project design. The instructions below apply equally to these situations as to situations in which objectives are altered.

PRINCIPLE:

For projects whose project objectives (as encompassed by the stated PDOs **and** key associated outcome targets) have been formally revised – through approval by the Bank authority that approved the original loans/credits/grants – project outcome will be assessed against both the original and revised project objectives. To assist in arriving at an overall outcome rating following this principle, **separate outcome ratings (against original and revised project objectives) will be weighted in proportion to the share of actual loan/credit disbursements made in the periods before and after approval of the revision**.

RATIONALE:

The rationale for the above is based on the following:

- The Bank uses an objective-based evaluation methodology whereby project outcome is evaluated against the project objectives for which Bank funds have been approved.

- If substantial changes to the original project objectives and/or key associated outcome targets are required, such changes have to be approved by the same authority that approved the original loans/credits/grants.

- For accountability purposes, the evaluation of project performance should take into account performance both before and after the revision of project objectives. Weighting pre- and post-revision performance by the share of actual loan disbursements before and after the revision took place is both practical and transparent, and at the same time rewards early revision.

EVALUATION APPROACH:

- Rate project outcome against the original project objectives and against the revised project objectives, respectively.[42]
- Assign a numeric value for each of the outcome ratings (under the original project objectives and under the revised project objectives): Highly Satisfactory = 6, Satis-

[42] IEG elaboration: based on these guidelines, the project is rated against both sets of objectives separately, for the entire duration of the project – not just the period for which each of the objectives was in effect.

factory = 5, Moderately Satisfactory = 4, Moderately Unsatisfactory = 3, Unsatisfactory = 2, and Highly Unsatisfactory = 1.

- Derive the actual total Bank/IDA disbursements before and after the date when the revised project objectives were formally approved.
- Arrive at an overall rating by weighting the two ratings by the proportion of actual total disbursement in each period, 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 a loan of $115 million when $20 million was disbursed. It then performed quite well and closed with an additional $80 million disbursed ($15 million was cancelled). The project was rated ***Unsatisfactory*** against the original project objectives (a value of 2 on the 6-point scale), while it was rated ***Satisfactory*** against the revised project objectives. The outcome rating is calculated following these steps:

- **Calculate 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. Thus, *the proportion disbursed before restructuring was 0.2 and after restructuring was 0.8.*

- **Assign a value of each of the two outcome ratings:** The *Unsatisfactory* outcome rating under the original project objectives amounts to a value of 2 on the 6-point scale, while the *Satisfactory* rating under the revised project objectives amounts to a value of 5 on the six-point scale.

- **Multiply the value of each outcome rating by its share in disbursements and add them.** 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) = 4.8. 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 outcome value of 4, or* ***Moderately Satisfactory*** *on the 6-point scale*.[43]

Additional examples can be found in the OPCS *Implementation Completion and Results Report Guidelines* (updated October 5, 2011).

[43] In comparison, had the project been rated Highly Satisfactory with the revised objectives, the weighted value is 6 x 0.8 = 4.8 and the weighted overall outcome value is 0.4 + 4.8 = 5.2, or Satisfactory with rounding.

8. Risk to development outcome (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 rating has two dimensions:

(i) the likelihood that some changes may occur that are detrimental to the ultimate achievement of the operation's development outcome; and

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

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

Other risks arise from factors outside the operation, be they country level, such as price changes, or global, 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 operation's design to withstand that change.

The rating helps to identify operations that require close attention in managing risks that may affect the long-term flow of net benefits. Rating the Risk to Development Outcome requires an assessment of the uncertainties faced by an operation over its expected remaining useful life and whether adequate arrangements are in place to help avoid or mitigate the impact of those uncertainties. The impact will increase if the design or implementation of the operation is not well aligned with the operating environment, or mitigation measures are inappropriate to deal with foreseeable risks.

Whereas the Outcome rating reflects the evaluator's best estimate of the expected overall development outcome, the Risk to Development Outcome rating reflects the evaluator's judgment of the uncertainties faced by the operation's development outcomes over its expected remaining useful life, taking account of any risk mitigation measures already in place at the time of evaluation.[44] Risk to Development Outcome

[44] In statistical terms, Outcome would be analogous to the "expected value/mean" whereas Risk to Development Outcome would be analogous to the "dispersion/standard deviation". Of course, in considering risk we are interested only in the downside risk – that is the possibility that the positive development outcomes, even for unsatisfactory projects, will be reduced. This is akin to a one-tailed test, which is the

says nothing about the absolute level of the expected net benefits. So, for example, an operation 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.

Criteria

The overall Risk to Development Outcome is rated by assessing both the probability and likely impact of various threats to outcomes, taking into account how these have been mitigated in the operation's design or by actions taken during its initial implementation. The evaluator should take account of the operational, sector, and country context in weighing (in each case) the relative importance of these individual criteria of risk as it may affect planned outcomes.

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

Rating Scale

Taking into account the above, the overall *Risk to Development Outcome* — *the risk at the time of evaluation that development outcomes (or expected outcomes) will not be maintained (or realized)* — should be rated using the following four-point scale:

Negligible to Low
Moderate
Significant
High

statistical basis for much hypothesis testing in the research literature. In that sense, the analogy is to a test that the benefits will be less than expected.

Whenever the lack of sufficient information, or other circumstances, makes it impossible to assign one of the above ratings, "*Non-evaluable*" should be recorded.

9. Bank performance (Section 8)

Definition

The Bank's performance is defined as the extent to which services provided by the Bank ensured quality at entry of the operation 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,[45] the evaluator separately rates the Bank's quality at entry and quality of supervision, and uses the OPCS/IEG harmonized evaluation criteria guidelines to arrive at an overall rating for Bank performance.

Quality at entry

DEFINITION

Quality at Entry refers to the extent to which the Bank identified, facilitated preparation of, and appraised the operation such that it was most likely to achieve planned development outcomes and was consistent with the Bank's fiduciary role.

CRITERIA

Bank performance is rated against the following criteria, as applicable to a particular operation. The evaluator 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 IL operations)
- Structural, Financial and Macro-economic Aspects (for DPL operations)
- Poverty, Gender and Social Development Aspects

[45] The lists of assessment criteria below are taken from the Quality Assurance Group's (QAG's) criteria for QEA7 and QSA6.

- Environmental Aspects[46]
- Fiduciary Aspects
- Policy and Institutional Aspects
- Implementation Arrangements
- Monitoring and Evaluation Arrangements
- Risk Assessment
- Bank Inputs and Processes

RATING SCALE

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

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

CRITERIA

Bank performance is rated against the following criteria, as applicable to a particular operation. The evaluator 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)

RATING SCALE

With respect to 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 six-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. Ratings for the more 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.

GUIDANCE

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.

10. Borrower performance (Section 9)

Definition

Borrower performance is defined as the extent to which the borrower (including the government and implementing agency or agencies) ensured quality of preparation and implementation, and complied with covenants and agreements, towards the achievement of development outcomes.

Borrower Performance is rated by assessing two dimensions: (i) government performance (central and/or local government as relevant); and (ii) implementing agency or agencies performance. *Where the government and implementing agency are indistinguishable, particularly for DPL operations, only an overall rating is necessary taking into account relevant criteria from both lists below*.

Government performance

Government performance is rated against the following criteria, as applicable to a particular operation. The evaluator should take account of the operational, sector, and country context in weighing the relative importance of each criterion of Government Performance as it affected outcomes.

Criteria

- Government ownership and commitment to achieving development objectives
- Enabling environment including supportive macro, sectoral, and institutional policies (legislation, regulatory and pricing reforms etc.)
- Adequacy of beneficiary/stakeholder consultations and involvement
- Readiness for implementation, implementation arrangements and capacity, and appointment of key staff
- Timely resolution of implementation issues
- Fiduciary (financial management, governance, provision of counterpart funding, procurement, reimbursements, compliance with covenants)
- Adequacy of monitoring and evaluation arrangements, including the utilization of M&E data in decision-making and resource allocation
- Relationships and coordination with donors/ partners/stakeholders
- Adequacy of transition arrangements for regular operation of supported activities after Loan/Credit closing

Rating Scale

With respect to the relevant criteria either in identifying opportunities for, or resolving threats to, development outcomes or sustainability, rate government performance using the following scale:

Highly Satisfactory	There were no shortcomings in government performance
Satisfactory	There were minor shortcomings in government performance
Moderately Satisfactory	There were moderate shortcomings in government performance
Moderately Unsatisfactory	There were significant shortcomings in government performance
Unsatisfactory	There were major shortcomings in government performance
Highly Unsatisfactory	There were severe shortcomings in government performance

Implementing agency performance

Implementing agency or agencies' performance is rated against the following criteria, as applicable to a particular operation. The evaluator should take account of the operational, sector, and country context in weighing the relative importance of each criterion of agency performance as it affected outcomes.

Criteria

- Agency commitment to achieving development objectives
- Adequacy of beneficiary/stakeholder consultations and involvement
- Readiness for implementation, implementation arrangements and appointment of key staff
- Timely resolution of implementation issues
- Fiduciary (financial management, governance, procurement, reimbursements, compliance with covenants)
- Adequacy of monitoring and evaluation arrangements, including the utilization of M&E data in decision-making and resource allocation
- Relationships and coordination with partners/stakeholders
- Adequacy of transition arrangements for regular operation of project supported activities after Loan/Credit closing

Rating Scale

With respect to the relevant criteria either in identifying opportunities for, or resolving threats to, development outcomes or sustainability, rate implementing agency performance using the following scale:

Highly Satisfactory	There were no shortcomings in implementing agency or agencies' performance
Satisfactory	There were minor shortcomings in implementing agency or agencies' performance
Moderately Satisfactory	There were moderate shortcomings in implementing agency or agencies' performance
Moderately Unsatisfactory	There were significant shortcomings in implementing agency or agencies' performance
Unsatisfactory	There were major shortcomings in implementing agency or agencies' performance
Highly Unsatisfactory	There were severe shortcomings in implementing agency or agencies' performance

Rating of Overall Borrower Performance

The rating of Borrower Performance is based on the ratings for each of the two dimensions (i) government performance; and (ii) implementing agency or agencies' performance. Government performance and implementing agency or agencies' performance should each be rated using their respective six-point rating scales, and for transparency, the individual ratings for them should be presented separately. The ratings for government performance and implementing agency or agencies' performance should be combined into a rating of Borrower Performance. Ratings for the more common combinations of the ratings for government and implementing agency performance are provided below, followed by additional guidance on other combinations.

Highly Satisfactory	Performance was rated Highly Satisfactory on both dimensions.
Satisfactory	Performance was rated Satisfactory on both dimensions, OR was rated Satisfactory on one dimension and Highly Satisfactory on the other dimension.
Moderately Satisfactory	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	Performance was rated Moderately Unsatisfactory on both dimensions (Also see guidance below.)
Unsatisfactory	Performance was rated Unsatisfactory on both dimensions, OR was rated Unsatisfactory on one dimension and Moderately Unsatisfactory on the other dimension.
Highly Unsatisfactory	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.

GUIDANCE

When the rating for one dimension is in the satisfactory range while the rating for the other dimension is in the unsatisfactory range, the rating of overall Borrower Performance normally depends on the Outcome rating. Thus, overall Borrower 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 Borrower performance did not significantly affect the particular outcome.

11. Quality of monitoring and evaluation (Section 10)

Why IEG rates monitoring and evaluation quality

The quality of monitoring and evaluation (M&E) is a rating introduced by IEG on July 1, 2006, with the objectives of systematically tracking the progress of projects in improving M&E and creating an incentive for better performance that would ultimately improve the quality of evaluations and the operations themselves. [47]

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.

[47] This section is based largely on a note prepared by Ridley Nelson in 2006.

Guidance

In assessing the M&E quality rating, the evaluator 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 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. Good M&E will always rely also 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 evaluator 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 on-going project, or as reflected in attributable changes made in subsequent interventions. These three elements are common to both investment and policy lending.

The evaluator is asked to discuss separately each of the three elements of M&E quality and to arrive at an overall *quality of M&E* rating on a 4-point scale.[48]

Criteria

M&E DESIGN

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

- To what extent were the objectives clearly specified and to what extent did the indicators reflect those objectives?

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

- To what extent were the indicators measurable in terms of numbers, timing, and location?
- To what extent were the proposed data collection methods and analysis appropriate for the purpose? In particular, how was sampling to be done? How were data to be collected? How were comparators handled to understand attribution along the logical results chain?
- To what extent did the design ensure that a baseline, if relevant, would be done in time?
- To what extent was the M&E design well-embedded institutionally and have sufficient stakeholder ownership?

M&E IMPLEMENTATION

In *M&E implementation*, the evaluator 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 PAD-enumerated indicators actually measured?
- To what extent were any weaknesses in design, including indicator specification, fixed during implementation?
- To what extent did the agency responsible for M&E function effectively and efficiently?
- To what extent was M&E owned by the various stakeholders?
- To what extent are the data reliable and of good quality – what evidence is there of 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 can the system designed and implemented be sustained?

USE OF M&E DATA

The evaluator should assess, first, to what extent the M&E findings were communicated to the various stakeholders and, second, to what extent this actually 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 did M&E focus on assessing whether the theory of change within the project causality logic was sound and did this lead to any reframing of strategy?
- To what extent did the findings of M&E measure outcomes as opposed to simply input application or outputs?
- To what extent can positive (or negative) shifts in the project's or program's direction and outcome be attributed to the M&E activities?

- To what extent did the M&E impact 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 four-point scale – negligible, modest, substantial, or high.

- **High:** There were at most 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 strengths and weaknesses 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 on effective project management and therefore indirectly affect the project efficacy and efficiency rating. If M&E is insufficient, it can affect the evaluator's ability to assess the achievement of the project's objectives.

Standards for M&E quality for projects with learning objectives

Projects with an explicit objective of learning, such as pilots or Learning and Innovation Loans (LILs), would be held to higher M&E standards by IEG 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 (Section 15)

As 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 provided in the ICR. For this reason, IEG rates the quality of the Implementation Completion and Results Report.

Criteria

The following criteria should be used to assess ICR Quality:

- Quality of evidence
- Quality of analysis
- Extent to which lessons are based on evidence and analysis
- Results-orientation (ICR should be outcome-driven, not an implementation narrative)
- Internal consistency
- Consistency with (OPCS) guidelines
- Conciseness

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. Typical problems include: inadequate evidence; incomplete ICRs (missing data in tables, no discussion of efficiency); failure to assess the objectives; relying too much on indicators instead of using all available data. These shortcomings would only justify a downgrade if they were severe. IEG does not downgrade the ICR quality simply because of a difference in opinion about the ratings. There is also necessary no relation 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 Scale

The quality of the ICR is rated on a three-point scale: Exemplary, Satisfactory, or Unsatisfactory. Generally speaking, the exemplary rating is assigned quite rarely. Examples of exemplary ICR reviews can be found on the IEG awards section of the IEG website.

13. Safeguards compliance, fiduciary, and unanticipated impacts (Section 11)

OPCS requires that the ICR "summarize key *safeguard and fiduciary issues* in the operation, compliance with the Bank policy and procedural requirements, and any problems that arose and their resolution, as applicable." It also asks that the ICR record "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.[49] This chapter 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 potentially adverse impacts of Bank-supported programs on the environment and people are avoided or minimized and 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:[50]

OP 4.01, *Environmental Assessment*
OP 4.04, *Natural Habitats*
OP 4.09, *Pest Management*
OP 4.10, *Indigenous Peoples*

[49] World Bank OPCS 2011b, p. 21.

[50] The Bank is undertaking a review of its safeguards policies and is expected to revise its safeguard framework in 2012.

OP 4.11, *Physical Cultural Resources*
OP 4.12, *Involuntary Resettlement*
OP 4.36, *Forests*
OP 4.37, *Safety of Dams*
OP 7.50, *Projects on International Waterways*
OP 7.60, *Projects in Disputed Areas*

During project identification, the Bank screens a project from the point of view of these safeguards policies and classifies the project into one of four categories – A, B, C, and FI – based on the significance of environmental and social risk (Box 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 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 (e.g., 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 Involuntary Resettlement. These impacts may affect an area broader than 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 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 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 (ISDS) and disclose it publicly, and to ensure that the appropriate safeguards of adequate quality are implemented by the Borrower as it supervises the operation.

The primary responsibility for implementing the safeguards actions lies with the Borrower. Specifically, the Borrower is responsible for preparing the appropriate assessment instrument and disclosing it, implementing and monitoring the mitigation plan consistent with the requirements, and making arrangements for independent verification of implementation.[51]

What Information is 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 is a Category A or B project, it is supposed to have an Annex 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 EI), and (for Category A and B projects) the assessment instrument and mitigation plan.[52] 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, whether the sub-projects 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 Category B or FI.
- If the physical components of the project that generated environmental or social effects were modified, such as through additional financing or project restructuring,

[51] 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.

[52] Any instruments or plans mentioned for other Categories should also be mentioned, but they are required for Category A.

the reviewer should note whether the environmental or social assessment was updated or a new assessment was prepared.

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

ARE ANY OF THE RATINGS AFFECTED BY PERFORMANCE ON SAFEGUARD COMPLIANCE?

At present, there is 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, then 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 terms of 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 terms of 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/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, whether or not the external auditors' opinions were qualified, the nature of the qualifications (that is, whether they were serious or merely administrative) and what measures were taken to address them. 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 to which financial covenants were complied with (should be reported in the ICR).
- Whether all Bank, IDA, and (where relevant) Trust Fund resources were fully accounted for by project closure.
- 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 project closure.

PROCUREMENT

Procurement issues include: the extent to which Bank procurement guidelines were followed; significant implementation delays due to procurement-related issues and their causes; 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 that arise 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 Bank or IDA funds are found to be disbursed for ineligible expenditures, then they are supposed to be refunded by the Borrower, preferably before project closure (see financial management above).
- For DPOs, compliance with the provision not to disburse for items included in a negative list, as part of the Loan/Credit Agreement.

HOW ARE FIDUCIARY ISSUES REFLECTED IN THE RATINGS?

While there is no rating for fiduciary issues, they can be related to Bank or Borrower performance and in some cases 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 or not 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, then this should be reflected in Bank performance.

Unanticipated positive and negative effects

Even when a project's objectives are not achieved, there are often many benefits from implementation. 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.[53] 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 "internalized" through project restructuring by modifying either project objectives or key associated outcome targets.

14. Evaluating development policy operations and programmatic lending

What is a development policy operation?

A **development policy operation** (DPO) provides rapidly-disbursing financing to help a borrower address actual or anticipated development financing requirements of domestic or external origins. It supports a program of policy and institutional actions: for example, strengthening public financial management; improving the investment climate; addressing bottlenecks to improve service delivery; diversifying the economy. Funds from a DPO cannot be earmarked to a specific sector.

The funding of the program is contingent on:

- Maintenance of an adequate macroeconomic policy framework (with inputs from IMF assessments);
- Satisfactory implementation of overall reform program; and
- Completion of a set of critical policy and institutional actions agreed between the Bank and the client.

This is in contrast to **investment operations**, which finance goods, works, and services in support of specific economic and social development objectives in a broad range of sectors. There are many types of lending instruments for investment operations, including:

[53] **From the OPCS ICR guidelines:** 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.

Adaptable program loans (APL); Emergency recovery loans (ERL); Financial intermediary loans (FIL); Learning and innovation loans (LIL), Specific investment loans (SIL); Sector investment and maintenance loans (SIM); and Technical assistance loans (TAL).
Both DPOs and investment operations can be funded by loans, credits, or grants.

Key terms involved in the design of DPOs are defined in Box 14.1.

Box 14.1. Vocabulary for Development Policy Operations

Actions are the policy and institutional actions of the government that are expected to bring about desired results. They are defined as implementation of substantive changes (reforms) in policies or institutions.

A ***program*** is the borrower's set of policy and institutional actions which sustainably promote growth and poverty reduction.

Prior actions are the actions that are deemed critical to achieving the results of the program supported by the DPO. Prior actions are taken by a country before presentation of a development policy operation to the Bank's Board and are included in the operation's Financing Agreement.

Triggers, in the context of a programmatic series of DPOs, are the planned actions in the second or later years of a program that are deemed critical to achieving the results of the program and that will be the basis for establishing the prior actions for later operations. Given their indicative nature, they are not included in the operation's Financing Agreement. The Bank evaluates the achievements with respect to triggers to decide on the scope, timing, and prior actions for each later operation.

Tranche release conditions are actions that a country agrees to take before a tranche is released in a multi-tranche operation. They are included in the operation's Financing Agreement.

The ***Policy Matrix*** summarizes all of the policy and institutional actions directly supported by the DPO (prior actions, triggers (for a programmatic series), tranche release conditions) and the results expected from these actions. …[It] is not meant to outline the entire borrower reform program but instead only the subset of the program which the development policy operation directly supports.

Source: Adapted from World Bank/OPCS 2011b, Box 1.1.

How is evaluation of a DPO different from evaluation of an investment project?

Both DPOs and investment projects use the same objectives-based methodology and have the same major ratings discussed in earlier chapters. They are both evaluated by the sponsoring region at the end of the operation in an ICR. But there are some differences:

- DPOs have a Program Document instead of a Project Appraisal Document
- The lending agreements for DPOs rarely include the objectives of the operation, but the lending agreements often have a Letter of Development Policy from the government as an annex, which describes the program.

- DPOs don't have "components", but they do have "policy areas." The latter are groupings of reforms or actions by topic that constitute the "outputs" of a DPO, in the same way that components are activities that are the outputs of investment projects. For both investment lending and DPOs these outputs (be they activities or actions) are intended to lead to achievement of the objectives (outcomes). In some Program Documents the policy areas are labeled as "pillars."
- The project outcome rating for DPOs is based on relevance (of objectives and of design) and efficacy (achievement of the objectives), but not efficiency.
- They often occur over a very short time frame – single-tranche DPOs can be approved and closed in 6 months' time – while investment projects generally last 5 years. Thus, it is more difficult to observe results for some DPOs, underscoring the need for good indicators of intermediate outcomes.

The other major difference is that DPOs can also occur in a programmatic series of successive development policy operations. Each operation in the series must be approved by the Board and the first operation in the series describes the indicative sequencing of future policy actions and expected end-of-program results.

DPOs that are part of a series are evaluated based on the overall objectives of the series, and a single ICR is produced discussing all of the operations in the series, with a single, series-level rating.[54] Programmatic DPOs in IDA-eligible countries that support the implementation of a country's Poverty Reduction Strategy are called Poverty Reduction Support Operations, which can be credits or grants.

What are prior actions and triggers, and how are they reflected in the ICR Review?

Prior actions are a set of mutually agreed policy and institutional actions that are deemed critical to achieving the objectives of the program supported by a DPO and that a borrower agrees to take before the Bank Board approves a loan (credit or grant).

Triggers, as used in the context of a Programmatic DPO, are the planned actions in the second or later year of a program that are deemed critical to achieving the results of the program and that will be the basis for establishing the prior actions for later operations.

There is no place on the ICR Review form that requires a listing of the prior actions, and in many cases the prior actions have little to do with the achievement of the objectives. However, to the extent that some of these prior actions *are* relevant to achievement of the objectives, they may be included in the ICR Review as evidence of implementation of part of the results chain.

[54] Programmatic series are now assigned a single rating for the whole series; prior to May 10, 2010, an end-of-series ICR was prepared, each operation was rated separately, with no overall series rating.

ICR Reviews of Programmatic DPO series

ICRs for programmatic development operations are prepared after the closing of the last operation in the series. As of May 24, 2010 ratings are issued for the *overall program, not by individual operation*.[55] OPCS's instructions for preparing ICRs for programmatic DPO series were revised (see Appendix C of OPCS's ICR Guidelines) and the ICR template for development policy lending was updated with footnotes reflecting the new instructions for programmatic DPO series.

What are ICRs supposed to report for the programmatic series and for individual operations, according to OPCS?

The ICR at the end of the series is expected to provide "a comprehensive evaluation of the series of operations as a coherent program. For each operation, the ICR should discuss the extent to which the operation's activities and inputs have contributed (or may be expected to contribute) to outputs, outcomes (or expected outcomes) and impacts (or expected impacts) of the series, as relevant. The ICR draws on program documents for individual operations in a series and builds on the cumulative findings and internal ratings in the ISRs."[56]

In terms of the format of the ICR,

- The Data Sheet fields in the front matter (A-Basic Information, B-Key Dates, C3-Quality at Entry, D-Sector and Theme Codes, E-Bank Staff, F-Results Framework Analysis, G-Ratings of Program Performance in ISRs, and H-Restructuring) are supposed to be completed for each operation individually. The first two parts of Section C-Ratings Summary are supposed to be completed for the programmatic series only; the main text block of Section F-Results Framework Analysis for program development objective is at the program level, but there will be separate tables (a) and (b) with the baselines, original targets, revised targets, and actual results for each of the operations in the series. There is no table for "overall program results" that are not captured at the operation level.
- Section 2, part 2.1, is supposed to include a table showing the status of prior actions for each operation separately.
- Annex 1, on Bank lending and implementation support, is supposed to present the task team members and staff time and cost by operation.

[55] Prior to this date, a single ICR was written at the end of the series and the overall achievements of the series were discussed, but each operation was rated separately.

[56] World Bank OPCS 2011a, Appendix C.

- The remaining sections of the ICR and Annexes are supposed to assess the overall series, but it is expected that the contributions or shortcomings from each operation will be part of that discussion (per the guidelines above).

How are the results for a programmatic series captured on the ICR review form?

The ICR Review form for DPOs can accommodate either individual DPOs or series (see the form below).

Section 1, includes blocks to record the details of each individual operation in the series that was approved. If there were two operations approved, two project data fields will be generated, one for each project. Some of these fields will be automatically populated, while others will need to be completed by the evaluator (amounts, dates). If fewer projects were approved than were planned, data fields will only appear for the projects that were approved. The fact that fewer were implemented than planned is also captured in Section 1.

If the assessment is for a single DPO, the first question in section 1 ("Is this part of a programmatic series?") will be ticked with "no" and only a single project data field will appear.

Section 2a. on the objectives, should present both the overall programmatic series objectives from the Program Document and the objectives of each of the operations. The assessment will take into account the contributions of each operation, but the overall assessment will be against the objectives of the series.

Section 2c, on the policy areas, should discuss the policy areas for each of the specific operations that were part of the series.

Section 2d should discuss the financing, etc., for all of the operations in the series.

Section 3 will discuss the <u>relevance of the overall series objectives</u> and the objectives of each operation, with a summary rating for the relevance of the series objectives. It will also discuss and rate the <u>relevance of the series design</u> (including the design of the individual operations), and the policy areas in relation to the objectives, with a summary rating for the series on relevance of design.

Section 4 should assess the achievement of the series objectives (outputs, outcomes, impacts), taking into account the contribution of individual operations to each objective.

Sections 6-11 should all assess based on the overall series performance, taking into account the contribution of each of the individual operations.

What are the guidelines if the programmatic series was not completed?

Occasionally a series of DPOs will be planned to support a program, but fewer operations are actually approved than were planned. For example, a Poverty Reduction Support Credit series may have been designed as a package of three successive operations, but for whatever reason, perhaps only two were actually approved. This is called an 'incomplete series'.

When fewer operations are approved than were planned, the projects that were approved are rated based on the series-level objectives. This would be the case, for example, if only 2 of the 3 planned operations were approved, or even if only one of the three operations was approved. In all of these cases, whatever was actually approved is assessed against the overall series objectives. It is likely when the series is incomplete, that there will be shortfalls in the inputs and outputs that were part of the results chain leading to the desired programmatic series outcomes.

ICR Review Form for DPOs

IEG **ICR Review**
Independent Evaluation Group

Date Created: 04/12/2011 08:27:04 AM
Last Updated: 04/13/2011 03:09:51 PM
Status: New

? - Help

1. Project Data:		**ICR Review Date Posted:**		
Country:	Moldova			
	Is this Review for a Programmatic Series?			Yes No
	How many operations were planned for the series?			3
	How many were approved?			2
Series ID:	S099166			
First Project ID :	P099166		**Appraisal**	**Actual**
Project Name:	Poverty Reduction Support Credit (prsc)	**Project Costs (US$M):**		
L/C Number:	C4236	**Loan/Credit (US$M):**		
Sector Board:	EP	**Cofinancing (US$M):**		
Cofinancers:		**Board Approval Date:**		16
		Closing Date:	16	16
Sector(s):	Central government administration (34%) General industry and trade sector (33%) General agriculture fishing and forestry sector (11%) Compulsory pension and unemployment insurance (11%) Agricultural marketing and trade (11%)			
Theme(s):	Regulation and competition policy (29% - P) Administrative and civil service reform (29% - P) Rural policies and institutions (14% - S) Trade facilitation and market access (14% - S) Public expenditure financial management and procurement (14% - S)			
Second Project ID :	P103941		**Appraisal**	**Actual**
Project Name:	Poverty Reduction Support Credit 2 (prsc 2)	**Project Costs (US$M):**		
L/C Number:	C4413	**Loan/Credit (US$M):**		
Sector Board:	EP	**Cofinancing (US$M):**		
		Board Approval Date:		16
Cofinancers:		**Closing Date:**	16	16
Sector(s):	General public administration sector (30%), Roads and highways (20%), General education sector (20%), General industry and trade sector (15%), Agricultural marketing and trade (15%)			
Theme(s):	Rural policies and institutions (20% - P), Education for all (20% - P), Regulation and competition policy (20% - P), Infrastructure services for private sector development (20% - P), Administrative and civil service reform (20% - P)			

Evaluator:	Panel Reviewer:	ICR Review Coordinator:	Group:
		Ismail Arslan	IEGPS2

2. Project Objectives and Components:

a. Objectives:

b. If this is a single DPL operation (not part of a series), were the project objectives/ key associated outcome targets revised during implementation?

c. Policy Areas:

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

3. Relevance of Objectives & Design:

a. Relevance of Objectives:

b. Relevance of Design:

4. Achievement of Objectives (Efficacy):

5. Efficiency (not applicable to DPLs):

6. Outcome:

a. Outcome Rating:

7. Rationale for Risk to Development Outcome Rating:

a. Risk to Development Outcome Rating:

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:

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: [?]	ICR [?]	IEG Review	Reason for Disagreement/Comments [?]
Outcome:			
Risk to Development Outcome:			
Bank Performance:			
Borrower Performance:			
Quality of ICR:			

NOTES:
- When insufficient information is provided by the Bank for IEG to arrive at a clear rating, IEG will downgrade the relevant ratings as warranted beginning July 1, 2006.
- The "Reason for Disagreement/Comments" column could cross-reference other sections of the ICR Review, as appropriate.

13. Lessons: [?]

14. Assessment Recommended? [?] ◯ Yes ◯ No

15. Comments on Quality of ICR: [?]
a.Quality of ICR Rating:

15. Note on Cancelled Operations

What is a Note on Cancelled Operation (NCO)?

A Note on Cancelled Operation (NCO) is prepared for an operation that fails to become effective or is cancelled before significant implementation is initiated.[57] The cut-off point for "significant implementation" is defined as final actual disbursement of less than five percent of the initial commitment or US$ 1 million (whichever is smaller), excluding any Project Preparation Facility and front-end fees. The NCO is sent to the Board describing the operation and explaining why it was not implemented. The ICR Guidelines also cover NCO requirements.

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

The OPCS guidelines for writing ICRs (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 operation'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 operation."

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 sub-ratings 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

For the purposes of assessing its quality, the NCO is expected to discuss the main events leading to cancellation, steps taken to resolve problems, exogenous factors, identifica-

[57] This information is from Appendix E of the OPCS *Implementation Completion and Results Report Guidelines*, updated October 5, 2011.

tion of causes, and responsible parties if the operation failed, and implications of failure. Above all, the purpose of the NCO is to clearly explain why the operation was cancelled; if the NCO does not convincingly explain or document the reasons for cancellation, the quality would be unsatisfactory.

16. Deriving lessons

The purpose of the ICR is both 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 don't achieve their objectives often produce some of the most valuable lessons.

The ICR Review typically presents 3-5 key lessons emerging from the information in the report. They may come from the ICR, or they may be reflections on this project from the IEG evaluator based on the ICR and compared with other projects reviewed by the evaluator (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 evaluator to identify the source.

Often, however, the lessons in the ICR are not well formulated. The two biggest issues in formulating lessons – in the ICR and by IEG evaluators -- 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. The table below distinguishes between facts, findings, lessons, and recommendations.

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 operation 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"

Facts and findings will be found throughout the ICR; they are the material from which lessons are built. If you find that you're repeating something verbatim already in the ICR, then it probably isn't 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, don't 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 didn't. One can't 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 didn't 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 it to succeed 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.
- *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 – low income households, fragile states, middle-income countries, weak or strong institutions, for example. Good results in spite of difficult contexts or weak results in spite of less-difficult contexts are good fodder for lessons. The context also includes aspects of Bank and borrower performance – factors such quality at entry, monitoring and evaluation, 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 monitoring and evaluation generate important lessons.

References

OED. 2004. *Annual Review of Operations Evaluation 2004*. Washington, DC: World Bank.

OECD/Development Assistance Committee, Working Party on Aid Evaluation. 2002. *Glossary of Key Terms in Evaluation and Results Based Management*. Paris: Organization for Economic Cooperation and Development.

OPCS/IEG. 2006. "Harmonized Evaluation Criteria for Implementation Completion and Results Report (ICR) and Independent Evaluation Group (IEG) Evaluations."

OPCS. Undated. "Guidelines for the structure and the content of the PAD".

OPCS. 2009a. "DPL appraisal template." http://intresources.worldbank.org/INTOPCS/Resources/380831-1175264321504/DPLAppraisaltemplate0609.pdf

OPSPQ. 2013a. "Investment Project Financing – Preparing the Project Appraisal Document (PAD)." April 9.

--------. 2013b. "Investment Project Financing – Project Preparation Guidance Note." April 9.

World Bank OPCS. 2011a. "Implementation Completion and Results Report: Guidelines". August 2006, last updated on 10/5/2011.

World Bank, OPCS. 2011b. "Designing Development Policy Operations," *Good Practice Note for Development Policy Lending*. January.

World Bank, OPCS. 2011c. "Designing Development Policy Operations in Situations of Fragility," *Good Practice Note for Development Policy Lending*. December.

Annex A: Guidelines by section of the ICRR

Section 1: Project data

Appraisal amounts: The source of the project costs, loan/credit amounts, and cofinancing amounts at appraisal is the Project Appraisal Document (for investment lending) or President's Report/Program Document (for development policy lending). The project costs include the contribution of the Bank, the government (counterpart funding) and any official cofinaciers.[58] 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 $US million.

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 operation 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:* an operation 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 providing *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.

Section 2: Project objectives and components

Objectives:

The project's objectives include the statement of objectives, as articulated in the Loan/Credit/Grant Agreement and the PAD, and key associated outcome targets, if any. The statement of objectives should be lifted directly out of the PAD and lending agree-

[58] 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..

ment (for investment lending) or Program Document/President's Report (for development policy lending) – not from the ICR.[59] When there are both overarching objectives and specific objectives for the operation, both should be presented.

In case of inconsistencies between the statement of objectives in the PAD and in the Credit/Loan Agreement, both should be cited but the statement in the legal agreement will be the benchmark for evaluation in the ICR Review. See Chapter 3 of the Evaluator Manual for additional guidance on identifying the objectives. If the objectives were formally revised through a restructuring with the Board, both the original and revised objectives should be reported.

Key associated outcome targets are targets for the project's main intended outcome(s), if any. They should not include project outputs and other key performance indicators unless they pertain to *targets* for the main outcomes of the project's objectives.

For projects with revised objectives/key associated outcome targets, both the original and revised objectives/key associated outcome targets should be reported and clearly marked "original" and "revised".

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

Check "yes" only if the project's objectives (as encompassed by the stated Project Development Objective and key associated outcome targets) have been formally revised through approval by the Bank authority that approved the original loans/credits/grants. Such projects will be assessed against *both* the original and revised project objectives. (See Chapter 7 of the Evaluator Manual for guidance on assessing restructured projects.)

c. Components (for investment lending) or Policy Areas (for DPLs), as appropriate:
For investment lending:
A description of components (matching the PAD/lending agreement, not the ICR) should be provided with sufficient detail to make clear the activities on which project funds were spent. 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.

Each component should be listed 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, respec-

[59] OPCS ICR guidelines instruct the author to take them from the PAD "and as reflected in legal agreements". They do not explain how to reconcile the two if they are different. However, recent communications from OPCS indicate that the Lending Agreement is given greater weight when there are material differences, as these are the objectives that are legally binding.

tively.[60] 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.

For development policy lending:
A brief description of the main policy areas covered in the operation in support of the government's program should be provided, as described in the Program Document/President's Report. It is not appropriate to list all of the prior actions or triggers in this section.

In the event that the ICR Review is for a programmatic series of DPLs, list the policy areas for the overall program and the specific focus of each planned operation in the series, according to the original program document, including operations that were planned but not approved.

d. Comments on Project Cost, Financing, Borrower Contribution, and Dates:
The following information should be recorded in this section:
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 Chapter 5),

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.

For development policy lending:

- comments on the date of tranche release and reasons for any delays
- the share of a package of adjustment/development policy operations executed at the same time.

Section 3: Relevance of objectives and design

This section should cover both: (i) the relevance of project objectives; and (ii) the relevance of project design. Each should be rated separately on a four-point scale (negligible, modest, substantial, high); no overall relevance rating should be reported.

The relevance of the objectives and design are assessed *at the time of project closing* – not project appraisal. While the ICR Review may note that the project objectives and

[60] 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 evaluator should note that the appraisal estimates for the components exclude contingencies.

design were relevant to strategies at project appraisal, the ICR Review should assess and rate the extent to which project objectives and design remained relevant towards the closing years of the project (even if the ICR is silent on this). See Chapter 4 of the Evaluator Manual for detailed instructions on assessing relevance.

Relevance of objectives

The relevance of project objectives is assessed in relation to country conditions, the Bank's strategy (as reflected in sector and country assistance strategies), and the government's strategies (sectoral or poverty reduction strategies, for example) *at the time of project closing*.

Relevance of design

The relevance of the project's design is assessed relative to its stated project objectives. This section should explicitly discuss the project's results framework – the extent to which the causal chain between the main investments or policy areas and intended outcomes was clear and convincing, and the extent to which exogenous factors and unintended (positive and negative) effects were identified. *This section should not include discussion of other issues in quality-at-entry of the project, which will be assessed in the section on Bank performance. For example, the choice of implementing agency is not discussed here, nor is the complexity of the project in relation to borrower capacity. Nor should the monitoring and evaluation design or choice of indicators be discussed. These latter items all belong in the section on Bank quality at entry.*

For projects with *revised objectives*, relevance of both the original and the revised objectives should be assessed and rated separately under the sub-heading for relevance of objectives, and the relevance of design in relation to the original and revised objectives should be assessed and rated separately under the sub-heading on relevance of design.

Section 4: Achievement of objectives

Organization of the section

This section should separately discuss and assess the achievement of each of the project's objectives. A heading should be introduced, corresponding to each outcome. Under the heading for each objective's anticipated outcome, evidence should be presented from the ICR documenting the extent to which the anticipated inputs, outputs, intermediate outcomes, and impacts were achieved. The purpose is to assess the degree of plausible association between the project's inputs and outputs with outcomes and impacts. If this is in any doubt, the lack of plausible association should be noted. If the project's design itself was weak with respect to plausible association between project activities and expected outcomes, this should already have been noted and discussed in the section "Relevance of Design." See Chapter 6 of the Evaluator Manual for more detailed instructions on assessing the achievement of objectives.

For projects with *formally revised objectives*, both the original and revised objectives should be assessed.

Parsing out the objectives
As statements of objectives often have “by” or “through”, denoting the activities, actions, or components to be undertaken, it is very important that the evaluator correctly identify which part of the objective statement is the desired outcome and which part is the means to that end (the components and outputs), as part of the results chain. In the event that the single objective statement includes parts with multiple intermediate or final outcomes, the evaluator should discuss and assess/rate each of the parts separately for the purposes of assessing efficacy. For more information on identification of the objectives, see Chapter 3 of the Evaluator Manual.

Achievement of the objectives should be assessed against stated objectives and their outcomes. Only in the rare cases where objectives are so totally vague or incoherent should objectives be inferred. Where objectives are inferred, they should be restricted to inferring "what the project meant to do" (based, for example, on project design or the discussion of expected benefits/outcomes in the appraisal documents), and not on "what the project should have done". Nor should objectives be inferred based on what actually happened (otherwise we would have a situation where objectives chase performance rather than the other way around).

Plausible attribution
Beyond tracking the plausibility of the results chain between project outputs and desired outcomes, the reviewer should comment to the extent possible on available information concerning attribution of the results to the project or government program being supported and the counterfactual (what would have happened in the absence of the project). This would include mention of other factors that might equally have contributed to the same outcome, independent of the project’s outputs – for example, weather patterns, economic crises, the presence of other donors – and any evidence supporting the influence of factors outside the project or program on its outcomes.

Ratings: The actual or expected achievement of each objective is rated on a 4-point scale: *High, Substantial, Modest, or Negligible*. When a project has both an overall objective and specific objectives, each of the specific objectives should be rated, with a comment included on the overall objective. In the case of projects with *formally revised objectives*, each of the original and revised objectives should be rated.

In the rare instances where lack of sufficient information or other circumstances make it impossible to assign one of the above ratings, "Not Rated" should be recorded. The comment on the overall objective could take the form of " data are insufficient to assess the extent to which the overall objective will be achieved".

Each objective is rated separately; there is no overall efficacy rating.

Section 5: Efficiency

Applicability. Development policy operations are not rated on efficiency; this section is to be completed only for investment-type operations, including technical assistance loans.

Measures of efficiency. Efficiency is a measure of how economically resources and inputs are converted to results. This section should report on all available measures of efficiency both ex-ante and ex-post, and would highlight any data gaps and methodological weaknesses in the Bank's assessment of efficiency. In the event that an ERR/IRR has been calculated, the reviewer should also include other available indicators of efficiency, including efficient use of project funds, in the assessment. See Chapter 6 of the Evaluator Manual for additional details on assessing efficiency.

"Coverage/Scope" refers to the percent of total project cost for which the Economic Rate of Return/Financial Rate of Return was calculated. A comment should be included, where possible, on the reliability of the ERR/IRR calculation(s) presented in the ICR.

Efficiency rating. Efficiency should be assigned an overall rating, based on a 4-point scale*: Negligible, Modest, Substantial, or High.*

Section 6: Outcome

In this short section, the reviewer must review the ratings on relevance of objectives, relevance of design, achievement of each objective, and overall efficiency in light of the OPCS/IEG Harmonized Evaluation Criteria, to arrive at an overall outcome rating on a six-point scale: *Highly Satisfactory, Satisfactory, Moderately Satisfactory, Moderately Unsatisfactory, Unsatisfactory, Highly Unsatisfactory*. Guidelines for this process, as well as for assessing the outcome of formally restructured projects, are in Chapter 7 of the Evaluator Manual.

In justifying the outcome rating, the reviewer can summarize the main results or achievements, but s/he should also link it to the relevance of objectives and design, the achievements or lack of achievement of certain objectives, and efficiency.

Section 7: Risk to development outcome

This section should present evidence to substantiate IEG's risk to development outcome rating for the project. See Chapter 8 of the Evaluator Manual for the definition and criteria.

Section 8: Bank performance

This section should present evidence to substantiate IEG ratings on the Bank's quality at entry, its quality of supervision, and the overall Bank performance rating, based on the

OPCS/IEG Harmonized Evaluation Criteria. All three ratings are on a six-point scale – *Highly Satisfactory, Satisfactory, Moderately Satisfactory, Moderately Unsatisfactory, Unsatisfactory, and Highly Unsatisfactory*. See Chapter 9 of the Evaluator Manual for a detailed explanation.

Note that when one of the two components of Bank performance is in the Satisfactory range and the other is in the Unsatisfactory range, the overall Bank performance rating is based on whether the Outcome rating is in the Satisfactory or Unsatisfactory range.

In addition, when both components of Bank performance are rated in the Satisfactory range, the overall Bank performance rating is constrained to be the lower of the two. The same rule holds when both are in the Unsatisfactory range.

Section 9: Borrower performance

This section should present evidence to substantiate IEG ratings on Government performance, Implementing Agency performance, and overall Borrower performance, based on the OPCS/IEG Harmonized Evaluation Criteria. See Chapter 10 of the Evaluator Manual for more detailed instructions. All three are rated on a six-point scale – Highly Satisfactory, Satisfactory, Moderately Satisfactory, Moderately Unsatisfactory, Unsatisfactory, and Highly Unsatisfactory.

Note that according to the Harmonized Criteria, when one of the two components of Borrower performance is in the Satisfactory range and the other is in the Unsatisfactory range, the overall Borrower performance rating is based on whether the Outcome rating is in the Satisfactory or Unsatisfactory range.

In addition, when both components of Borrower performance are rated in the Satisfactory range, the overall Bank performance rating is constrained to be the lower of the two. The same rule holds when both are in the Unsatisfactory range.

In instances in which it is impossible to distinguish between the Government's performance and the implementing agencies' performance – such as for development policy lending – the evaluator should propose only an overall Borrower performance rating.

Section 10: M&E design, implementation, and utilization

The section "M&E Design, Implementation and Utilization" should include a separate textual assessment under separate headings of:

- M&E Design
- M&E Implementation

- M&E Utilization

A summary rating of M&E Quality should be included using one of the following ratings: High, Substantial, Modest, Negligible, or Non-evaluable. *For the time being, IEG is not rating the three individual components separately.*

M&E weaknesses identified in this section should always be factored into the Bank Performance rating (M&E design should also factor into quality at entry and M&E implementation should be factored into quality of supervision). It also indicates weaknesses in Borrower Performance and should also be factored into the Borrower Performance rating. See Chapter 11 of the Evaluator Manual for greater detail.

Section 11: Other issues

This section should separately discuss: (a) relevant safeguards and compliance with them; (b) fiduciary compliance; and (c) positive and negative unintended impacts. None of these is rated; they are all discussed. See Chapter 13 of the Evaluator Manual for greater detail on these issues.

Safeguards refer to compliance with the following operational policies: *Environmental Assessment* (OP 4.01), *Natural Habitats* (OP 4.04), *Pest Management* (OP 4.09), *Indigenous Peoples (*OP 4.10), *Physical Cultural Resources (*OP 4.11), *Involuntary Resettlement (*OP 4.12), *Forests (*OP 4.36), *Safety of Dams (*OP 4.37), *Projects on International Waterways* (OP 7.50), and *Projects in Disputed Areas (*OP 7.60).

The evaluator should report on the environmental classification of the project (A, B, C, or FI, which is in the PAD), the safeguard policies that were triggered (noted in the PAD and in the ICR) and the extent to which they were successfully addressed (from the ICR). In the case of category A projects the ICR would normally relate in some detail, the expected impacts on affected people and the environment and mitigation measures taken.

While there is no safeguard rating at the present time, the success or failure in addressing safeguard issues should be reflected in discussions of both the Bank's and Borrower's performance. If the ICR is unclear as to the resolution of the safeguard, this is something that should be pursued in the meeting with the Task Team Leader before the ICR review is finalized; poor reporting in the ICR on safeguards can also be a factor in the quality of the ICR.

Fiduciary issues refer to compliance with operational policies on financial management (OP/BP 10.02), Procurement (OP/BP 11.00), and Disbursement (OP/BP 12.00). This material is to be culled from throughout the ICR. While there is no rating for fiduciary issues, they can be related to Bank or Borrower performance and in some cases the efficiency component of the Outcome rating.

Unintended positive or negative benefits. An unintended benefit is defined as a benefit that occurred outside the objectives framework of the project. Where there are unintended benefits, an assessment should be made of why these were not "internalized" through project restructuring by modifying either project objectives or key associated outcome targets. When not internalized through restructuring, any unintended impacts-positive or negative-should be taken into account only if significant and substantiated. Even when a project's objectives are not achieved, there are often many benefits from implementation. However, those benefits are taken into account in the assessment of the objectives. The material described in this section should include positive or negative benefits or externalities arising outside of the stated objectives.

Section 12: Ratings summary

This table summarizes the project ratings from the ICR (which the reviewer must enter him/herself) and the IEG ratings from this review (which automatically populate the field), and explains 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.

Section 13: Lessons

This section presents generally up to five key lessons emerging from the information provided in the ICR. They may be relevant and well-documented lessons directly from the ICR, or they may be reflections on this project (based on the ICR) and compared with other projects reviewed by evaluator (or, for example, confirming that the findings in this ICR serve as an example of other IEG evaluations). It is important, however, to clearly note which lessons are from the ICR and which are from the IEG evaluator.

See Chapter 16 of the Evaluator Manual for guidance on the difference between a finding, lesson, and recommendation.

Section 14: Assessment recommended?

An Assessment can be recommended by the reviewer for any one (or more) of the following reasons:

i. if requested by Region, Executive Directors, or Co-financiers;
ii. to support a cluster/country sector review;
iii. as input into sector/thematic study;
iv. as input into Country Program Evaluation;
v. as a potential impact evaluation;
vi. because it is an innovative project/new instrument;
vii. due to major disagreement with the ICR rating and/or poor ICR quality;
viii. because the project is from an under-evaluated country/sector/theme; and

ix. because of a major safeguard compliance issue.

The first six reasons correspond to IEG's learning function, and the remaining three to its accountability function. If the project is recommended for a field assessment, the reviewer should offer a rationale. This will flag the project as a possible candidate, but it will not necessarily lead to a field visit.

Section 15: Quality of the ICR

The quality of the ICR is rated on a three-point scale: *Exemplary; Satisfactory; and Unsatisfactory*. The following criteria should be used to assess ICR Quality:

- Quality of evidence
- Quality of analysis
- Extent to which lessons are based on evidence and analysis
- Results-orientation (ICR should be outcome-driven, not an implementation narrative)
- Internal consistency
- Consistency with guidelines
- Conciseness

See Chapter 12 of the Evaluator Manual for additional instructions in rating the quality of the ICR.

Annex B: ICR Review Checklist

Overall	
	Text is frank but diplomatic, 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)
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/Policy Areas	
	2a. The objectives from the PAD and legal agreement are both mentioned, with page numbers
	2a. If the objectives are not identical, the lending agreement's objectives are used.
	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.
	2a. If a programmatic series of DPOs, the series objectives and the objectives of the individual operations are listed, with page numbers from the Program Documents.
	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 original and 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 operations); any components added after approval are also listed with the same information.
	2c. For DPOs, policy areas are listed; the prior actions are not listed.
	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's an explanation for any project extension(s); for DPOs, the date(s) of tranche release(s) are mentioned.
	2g For DPOs, if series was not completed, there's an explanation why not.
Section 3: Relevance of objectives and design	
	Relevance of the objectives/outcomes is rated with respect to country conditions, the current CAS, and the country's current development priorities, at project closing.
	Relevance of design is rated according to the relevance of the project's causal chain (activities/components) to achieving the objectives.
	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 and relevance of design are separately rated on a 4-point scale, no overall rating.
Section 4: Efficacy	
	There's a separate heading and rating for each objective/outcome to be achieved
	The headings all represent outcomes or intermediate outcomes (the expected changes to result from the operation), 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 evidence for the entire results chain (outputs and outcomes) is presented, including intermediate outcomes.
	In each case, there's a discussion of attribution of the results to the operation, other factors that might have affected the outcome beyond the project.
	There's no overall efficacy rating.
Section 5: Efficiency	
	If there's an ERR or 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 on 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's a single rating for efficiency on a four-point scale
Section 6: Outcome	
	The proposed outcome rating is consistent with the guidelines for combining relevance of objectives and design, efficacy of each objective, 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.
	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 did. (For example, technical, financial, economic, social, political, environmental, ownership of government or other stakeholders, institutional support, governance, and exposure to natural disasters.)
Section 8: Bank performance	
	Bank quality at entry incorporates comments from the relevance of objectives and design 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 sub-ratings and, when split between the satisfactory and unsatisfactory scales, according to the outcome rating.
Section 9: Borrower performance	
	Actual vs. planned counterpart contributions and their timeliness are mentioned in the government performance section.
	The implementing agency performance section clearly identifies what entity(ies) is being assessed.
	Shortcomings in implementation of safeguards or fiduciary performance are taken into account in the Borrower's performance ratings.
	Shortcomings in implementing M&E and using the data are part of the implementing agency performance.
	The overall Borrower performance rating is correctly calculated from the two sub-ratings and, if partly above and below the line, takes into account the outcome rating.
Section 10: Monitoring and Evaluation	
	On monitoring, the M&E design considers the choice of indicators, whether they are adequate to measure 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, the data were collected in a timely manner, and 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, 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 operation, and of sufficient magnitude to be important.
Section 12: Ratings	
	When the IEG ratings differ from ICR ratings, there's an explanation in the last column
Section 13: Lessons	
	There are not more than five lessons and it is clearly noted whether they are from the ICR or from the evaluator.
	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 unsatisfactory, the explanation must point to one or more "fatal flaws," not a series of small errors.
	If the quality is rated exemplary, the reasons are well documented and there shouldn't be more than small, incidental shortcomings. (It shouldn't simply say that it was frank and well-written.)

Note: The treatment of the macro framework for DPOs is to be added (requires research).

Annex C: Guidelines for IEG's meeting with the last project Task Team Leader[61]

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 so as to improve the accuracy and quality of IEG's ICR Reviews; and (ii) to ensure due process by providing the project Task Team Leader and the IEG evaluator 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 Region and is different from the meeting that the Region 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 ICR Reviewer to the meeting, in conducting his/her first ICR Review. It will be the responsibility of the hiring IEG staff (typically, the cluster coordinator) to determine when the new ICR Reviewer is ready to plan and conduct his/her meetings with the last project task team leader without assistance from more experienced reviewers.

Depending on who will be attending the meeting from the Region's side, the relevant Cluster Coordinator or IEG Manager may also choose to attend the meeting (please see point 3 below).

3. Who should attend from the Region's side?

The meeting should be held with the last Task Team Leader of the project or, in the case of a programmatic series, the Task Team Leader of the final operation. In neither case should the meeting be held with the ICR author, unless these two individuals are the same or such arrangement is proposed by the last Task Team Leader. In the rare instances where the last Task Team Leader of the project has left the Bank, the IEG ICR

[61] Based on "Guidelines for IEG's ICR Review Meeting with the Project Task Team Leader", DMT Discussion Draft, July 28, 2008, with the addition of point 12, following an internal assessment of the procedures in early 2009.

Reviewer, upon consultation with the Cluster Coordinator, should contact the concerned Sector Manager for an alternative suggestion. It would be up to the project Task Team Leader to invite other Regional staff at his/her discretion.

The IEG ICR Reviewer should ensure that s/he has advance notice of the complete list of attendees from the Region's side. 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 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 because of delay in the availability of the last TTL, the ICR Reviewer (with the concurrence of the ICR Review Coordinator) may meet with the project Task Team Leader 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 email indicating the Panel Reviewer's concurrence with any substantive changes to the draft ICR Review made subsequent to the meeting with the project Task Team Leader and to include that email 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 Task Team Leader in the email 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 Task Team Leader might have done differently had s/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, while being totally responsive to any other questions, including on IEG evaluation methodologies, from the project Task Team Leader.

The ICR Reviewer should record any additional information (not already contained in the ICR) obtained during the meeting with the project Task Team Leader in the draft ICR Review, noting "In a meeting between IEG and the project task team, the project team stated/clarified that….".

If an agreement is reached at the meeting that the project Task Team Leader 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 provided both at the meeting itself and in a follow-up email to the project Task Team Leader. 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 the purpose and not in the form of volumes of documents.

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

Attempt should be made to have a face-to-face meeting. If the project Task Team Leader is travelling for an extended period and/or based in the field making a face-to-face meeting impossible (given time/budget constraints), the meeting should be conducted by telephone. Given the cost implications of a video conference, clearance from the relevant IEG Manager should be sought before selecting this option.

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

The ICR Reviewer should prepare for this eventuality by contacting the project Task Team Leader and checking his/her travel schedule early on and getting the draft ICR Review ready soon after it is assigned. If the TTL is travelling or if the TTL is based in the field office, audio meeting should be suggested. If the TTL is not responsive after three weeks to set up a meeting the ICR Reviewer may proceed with finalization of the draft. He/she should inform the panel member that repeated attempts to meet with the TTL were unsuccessful.

9. Should IEG circulate an agenda to the project Task Team Leader before the meeting?

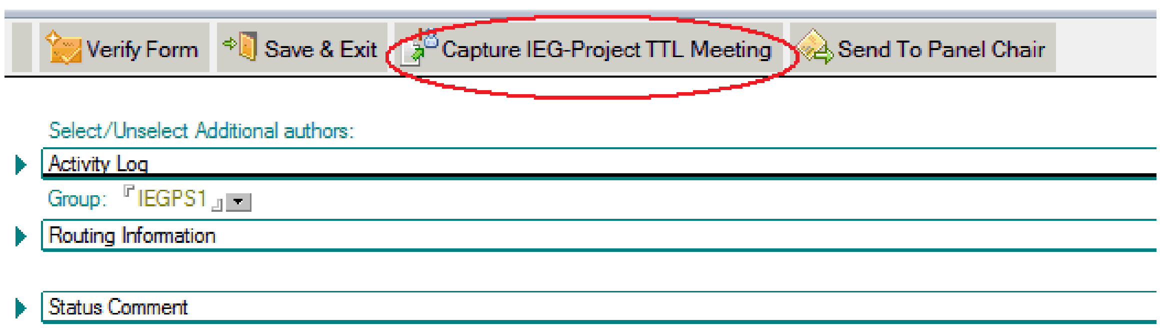
It is not mandatory to circulate an agenda before-hand, but the ICR Reviewer should feel free to do so if s/he finds it useful.

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 (without going into the specific details of the discussion of each topic). The names of attendees, date, and time of the meeting should be noted in the summary.

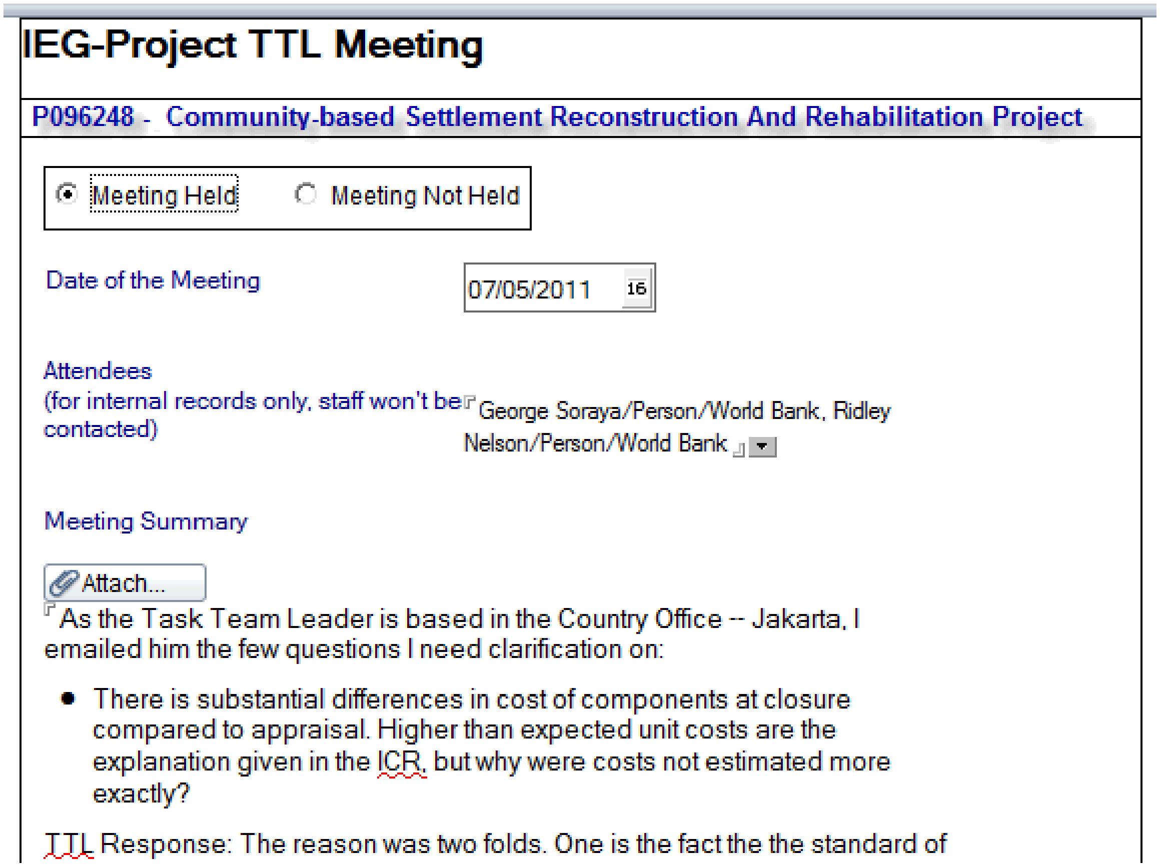
The summary of the meeting with the TTL must be entered into or appended to the ICR Review Database before the Review is sent to the Panel Reviewer. Enter into the ICRR work space by clicking on "Work on Form" for the ICR Review. At the top of the screen will be a button, "Capture IEG-Project TTL Meeting" (Figure C.1).

Figure C.1: Screen shot of the "Capture IEG-Project TTL Meeting" button.



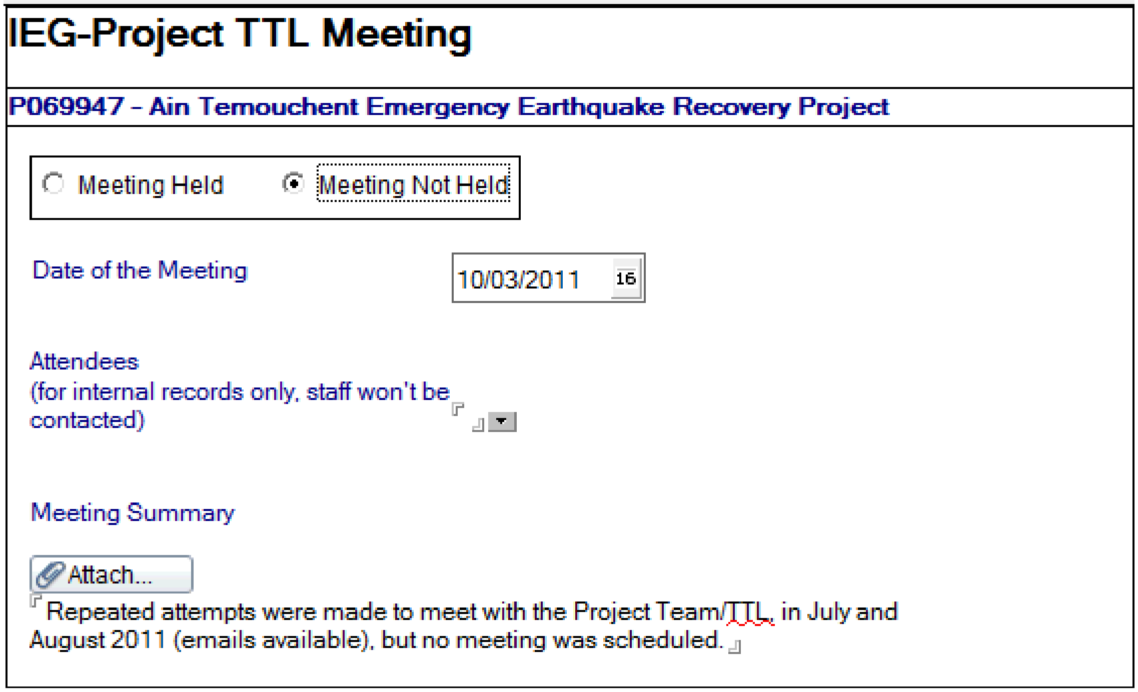
After clicking on the Capture button, the screen that appears will show two options: "Meeting Held" or "Meeting Not Held." If the meeting was held, the evaluator will need to click on that button and enter the meeting date and participants, along with the meeting minutes (either as a narrative or as an attachment (Figure C.2).

Figure C.2: Recording the summary of the TTL meeting



If a meeting did not take place (because, for example, repeated requests for a meeting were not answered), there is also the option of clicking "Meeting Not Held" and explaining the reason (under "meeting summary") why the meeting wasn't held (Figure C.3). However, the evaluator needs to get prior approval from the ICR Review Coordinator before submitting "Meeting Not Held" as an option.

Figure C.3: Screen shot if the meeting was not held.


IEG-Project TTL Meeting

P069947 - Ain Temouchent Emergency Earthquake Recovery Project

Meeting Held   Meeting Not Held

Date of the Meeting 10/03/2011

Attendees
(for internal records only, staff won't be contacted)

Meeting Summary

Attach...

Repeated attempts were made to meet with the Project Team/TTL, in July and August 2011 (emails available), but no meeting was scheduled.

11. How will the IEG Manager sending the ICR Review to the Region know that the meeting has taken place?

IEGCS will develop a way to electronically record the names of attendees, date and time of the meeting in the Activity Log in IEG's ICR Review database enabling the concerned IEG Manager to quickly confirm that the meeting with the project Task Team Leader did, indeed, take place.[62]

[62] As of July 2010, this modification has not yet been operationalized.