Project Self-evaluation in EBRD

EvD Discussion Paper for Management Working Group and Board of Directors

EvD ID: TRP-085

22 January 2020

EBRD EVALUATION DEPARTMENT

European Bank
for Reconstruction and Development

The Evaluation department (EvD) at the EBRD reports directly to the Board of Directors, and is independent from the Bank's Management.

This report has been prepared by EvD in response to a Recommendation by an Independent External Evaluation. Its principal author is Victoria Millis, Senior Evaluation Manager, with input from Saeed Ibrahim, Principal Evaluation Manager, and Aziza Khanbekova, Principal Evaluation Manager, EBRD Evaluation Department.. It is circulated under the authority of the Chief Evaluator, who is solely responsible for its content. The views expressed herein do not necessarily reflect those of EBRD Management or its Board of Directors.

Contents

Abbreviations & defined terms

ASB	EBRD Advice for Small Businesses Programme
COMPAS	MDBs' Common Performance Assessment System
CSRM	EBRD Country Strategy Coordination & Results Management team
DAC	Development Assistance Committee of the OECD
EBRD	European Bank for Reconstruction and Development
EC	European Commission
ECG	Evaluation Cooperation Group
ESD	EBRD Environment and Sustainability Department
EvD	EBRD Evaluation Department
G20	"Group of 20" developed and developing countries
HIPSO	Harmonized Indicators for Private Sector Operations
IBRD	Internation Bank for Reconstruction and Development (World Bank Group)
IDA	International Development Association (World Bank Group)
IFC	International Finance Corporation (World Bank Group)
IFI	International financial institution
Kirk Report	Independent External Evaluation of EBRD's Evaluation System
MDB	Multilateral development bank
MIGA	Multilateral Investment Guarantee Agency (World Bank Group)
OECD	Organisation for Economic Cooperation and Development
OE&E	EBRD Operational Effectiveness and Efficiency programme
OPA	Operation performance assessment (self-evaluation report)
PCR	Project completion report
TC	Technical cooperation (assignment)
TIMS	EBRD transition impact monitoring system
UN	United Nations
UNEG	United Nations Evaluation Group

1. Introduction and background

Among the key features common to high-performing organisations are internal systems that provide accurate and timely performance assessment, feed this effectively back into operational work and decision-making, and use it actively to improve performance and build competitive advantage. Systematic self-evaluation has long been mandatory in public multilateral organisations, and is seen by shareholders as at the very heart of effective institutional accountability.

The Chairman's Report on the Agreement Establishing the EBRD contained an expectation of the ex post evaluation of all projects, and project-level self-evaluation has been done ever since. Management, validation and quality assurance of Management's self-evaluation have featured prominently in the work of the Evaluation Department (EvD), particularly since independence from management in 2005, and EvD has had a key role in some important aspects of system design.

However few of the current system's intended broader objectives have materialised. While there are individual examples of high-value self-evaluations, they are relatively few. As input for EvD's independent evaluations their overall utility is limited. For the most part self-evaluation has never been meaningfully integrated into Management-side performance monitoring or operational feedback systems, including with respect to resourcing, incentives and agreed priorities.

The Evaluation Department has consistently raised concerns about the value and effectiveness of Bank evaluation systems. Its Board-approved Work Programme for 2019 included an evaluation of self-evaluation at the Bank in order to identify key issues and suggest improvements.

Separately a first-ever "Independent External Evaluation of EBRD's Evaluation System" (Kirk Report), commissioned by the Board of Directors to assess evaluation systems Bank-wide, was delivered in July 2019.[1] It gave particular prominence to issues with self-evaluation and returned many of the same observations that EvD has been making for many years. A central finding was that EBRD Management was not meeting its obligations regarding self-evaluation. The report also made clear that the performance and effectiveness of the self-evaluation system is integral to the value and credibility of the Bank's wider results, monitoring and evaluation architecture.

The Report recommended, and Board and Management endorsed, that Management conduct a comprehensive review of its self-evaluation approach and systems. Management was directed to draw upon an EvD analysis prepared for this purpose, setting out key issues and EvD's perspectives. It has created several Working Groups for this purpose and invited EvD to contribute in an advisory capacity.

This paper provides EvD's initial contribution to the Management review as directed by the Kirk Report. It takes the Kirk Report findings as given and authoritative. It does not draw from any original work, nor does it seek either to provide deeply detailed analysis or to prescribe specific solutions. Instead it mainly consolidates analysis, findings and recommendations that EvD has largely provided in other contexts and that are directly relevant to the Kirk findings and Management's work ahead. The objective is to contribute in an advisory capacity to a wider

[1] CS/AU/19-37

Management-led process in which EvD will also be systematically engaged. It also seeks to inform the Board and assist it to provide effective guidance for further system development. Overall, however, the findings and insights that are already available – from the Kirk Report, from the accumulated work of EvD, and from relevant experience across the IFI system -- provide a high-value basis on which Management can proceed.

The paper is organised as follows:

- Section 2 provides institutional context, including brief background on international standards and EBRD's relevant policies, and summarises EBRD's current arrangements.
- Section 3 sets out the Kirk Report's recommendations and principal findings underpinning them.
- Section 4 identifies important shortcomings of the system as indicated by EvD's work, and summarises its main observations about key changes Management should consider.
- Section 5 discusses a number of specific issues which EvD sees as key drivers of the existing system's deficiencies.
- Section 6 discusses features of a modified system that EvD believes are critical to accomplish significant improvement in both structure and performance.
- Section 7 sets out some specific actions for the working groups established by Management.

2. Context overview

2.1 International standards

This paper refers occasionally to monitoriong and evaluation materials which have international standing.

The OECD Development Assistance Committee (DAC) Network on Development Evaluation Principles for Evaluation of Development Assistance,: Membership of the DAC comprises OECD member states, so the standards were primarily developed for use by bilateral donor countries rather than MDBs. However, they are widely accepted by both MDBs and bilaterals as the international standard for development evaluation. The DAC publishes related documents such as the Evaluation Criteria,[2] Quality Standards for Development Evaluation and the DAC Glossary of Key Terms in Evaluation and Results Based Management.

United Nations Evaluation Group (UNEG) Norms and Standards for Evaluation 2016: Standards applied to evaluation within the various UN bodies.

Evaluation Cooperation Group (ECG) Good Practice Standards: The ECG is a forum for meetings of heads and staff of evaluation departments of MDBs. The ECG "Big Book" contains standards

[2] The DAC updated the Evaluation Criteria in 2019, in part to bring them more closely into line with the Sustainable Development Goals.

agreed by ECG covering evaluation of public and private sector projects and country strategies, independence of evaluation departments, and peer review of evaluation departments. These were developed by and for the MDBs and are the most directly applicable to EBRD, which is a member. In 2018 the ECG also published a Practice Note on Self-evaluation in ECG member institutions following a stock-take of practice among members.

World Bank Group Evaluation Principles 2019: Common principles for evaluation developed by the three institutions (IBRD/IDA, IFC, MIGA) together with the Independent Evaluation Group, in response to the 2015 External Review of the Independent Evaluation Group. They establish core principles of utility, credibility and independence, and underlying principles for selecting, conducting and using evaluations.

Harmonized Indicators for Private Sector Operations (HIPSO) result from joint efforts of senior staff from multilateral and bilateral development institutions to enhance development impact through common development indicators. It is an initiative of Management monitoring and results units rather than independent evaluation departments, in which the EBRD has been closely involved.

2.2 Policy foundations at the EBRD

The Bank's basic documents and the Evaluation Policy provide the foundations for project-level evaluation and self-evaluation at the EBRD.

The Agreement Establishing the Bank (Chairman's Report) refers to operating policies covering:

> *"among other things: ... the detailed requirements for the identification, appraisal, monitoring, implementation and* ***ex-post evaluation of all projects****, including their economic, technical, managerial, financial and environmental aspects."*[3]

The EBRD Evaluation Policy of 2013 establishes that Management:

> *"ensures an effective system of self-evaluation and reports periodically to the Board on its scope and operations."*
>
> *"ensures that proposed operations clearly specify expected results and related performance indicators so as to allow effective evaluation"*
>
> *"ensures that programmes, policies and strategies identify their expected results with sufficient specificity so as to allow effective evaluation."*

Under the Evaluation Policy EvD also:

> *"ensures the integrity of the EBRD's evaluation system by: developing methods and processes for evaluation, in consultation with Board and Management wherever necessary; ... validating and reviewing self-evaluations prepared by Management, and assessing the adequacy of the process."*[4]

[3] Chairman's Report on the Agreement Establishing the Bank, explanatory notes to article 13, in Basic Documents of the EBRD.
[4] BDS12-324 EBRD Evaluation Policy

2.3 Evolution of project-level evaluation at the EBRD

2.3.1 Investment operations

The self-evaluation instrument. Project-level evaluation, including self-evaluation has been a feature of the Bank since its inception. The first Evaluation Policy envisaged a project completion report by operational staff for every project, with EvD preparing "performance audits" of selected reports.[5] In practice, the project completion report was operationalised as the "expanded monitoring report" (XMR), a one-off, expanded version of the annual or biannual project monitoring report. From 2008 onwards, a shorter version of the XMR was used for projects not scheduled for independent evaluation or validation. In 2013 at EvD's urging the self-evaluation report was detached fully from the monitoring system and turned into a stand-alone document called an operation performance assessment (OPA). The intention was to elevate the profile and content of self-evaluations and establish them as a stronger basis for performance assessment.

Coverage of self-evaluation. Self-evaluation reports have always aimed at 100% coverage of stand-alone, Board-approved operations. Investments approved under frameworks have been evaluated at either framework or operational level, the approach varying over time and depending on the characteristics of the specific framework.

Coverage of independent evaluation or validation. Initially, EvD evaluated every evaluation-ready project in depth, conducting 5-10 full project evaluations per year. The number of full project evaluations increased progressively to 23 per year by the mid-2000s but could not keep up with increasing numbers of annual project approvals. EvD progressively retreated from a model of in-depth evaluation of all projects. In 1996 EvD introduced XMR Assessments to provide a shorter (two-day) desk-based validation of (initially) a majority of XMRs. The XMR Review, introduced in 1998 and applied to around 40-50% of evaluation-ready operations, saw EvD advising bankers on the quality of draft XMRs but not validating the findings. The selection of projects to be validated through XMR Assessments (later OPA Validations) was fully randomised from 2009 to 2016, with the aim of producing a statistically representative and significant sample of independently evaluated or validated projects. In 2017 EvD dropped random selection and reduced the number of validation reports to 10-15 per year.[6]

Frameworks were incorporated imperfectly into this system, taking the same case-by-case approach as for self-evaluation.

Report content. The XMR/OPA template has always been designed and owned by EvD, which consults Management on significant revisions. It has always comprised several sections addressing project performance against selected criteria, plus a final section on lessons emerging from the project. Until 2015, it assessed performance against criteria closely tied to EBRD project approval documents (additionality, operational objectives, transition impact, financial performance, environmental performance, investment performance and Bank handling) and used

[5] Until 2005, EvD was known as the Project Evaluation Department (PED). This report uses "EvD" throughout to avoid confusion.
[6] The reasons for this were discussed in detail in the EvD Work Programme 2017-19 and Budget 2017 (BDS16-241), section 3.2.2.

related terminology. In 2016 EvD introduced a new version of the OPA intended to align better both with international standards of development evaluation and with the emerging EBRD results architecture and anticipated introduction of project-level results frameworks for which EvD had been pushing strongly. It assessed projects according to evaluation criteria established by the Development Assistance Committee of the OECD: relevance, effectiveness, efficiency, impact and sustainability. Instead of classifying project objectives by type as operational, environmental or transitional, it classified them by level as outputs, outcomes and impacts and organised them into results frameworks – again consistent with international standards. At the same time, the short-form OPA template dropped project performance ratings, in favour of a description of performance and lessons arising.[7]

Handling and administration. The scheduling of projects for evaluation follows a process substantially unchanged since the 1990s: EvD prepares an annual list of possible candidates “ready for evaluation” in the following year and discusses it with individual Banking teams. The resource intensive process takes up most of the fourth quarter and part of the third quarter of each year, with EvD notifying Banking of the final selection and OPA type (long-form or short-form) at the start of the new year. Banking delivers the agreed OPAs to EvD, which reviews them and suggests quality improvements. Banking is responsible for the content and for consulting other departments; EvD does not validate the report unless the project is one of the small number scheduled for an independent evaluation or validation. In both cases OPA content remains fully under Banking control; validations of long-form OPAs are a separate EvD-owned product. EvD oversees the administration of report delivery and filing.

Use. The completed reports are essentially not disseminated. Filed in ProjectLink, the Bank's access-controlled electronic filing system, they can be viewed only by staff on the operation team and their line management. As far as EvD is aware, Management does not make any use of OPAs either individually or in aggregate. EvD extracts lessons and enters them in the Lessons Investigation Application, a database accessible to operational staff. EvD circulates its own OPA validations internally, including to the Board of Directors. EvD further reports on validated OPAs in aggregate in the Annual Evaluation Review, an annual report to the Board which is also published on the Bank's external website. Up to 2016, when EvD validated a representative sample of OPAs, the analysis in the AER compared performance across evaluation years, sectors and regions, and Management used EvD's ratings in inter-MDB comparison exercises such as COMPAS.[8] This is no longer possible given the small number of validations now prepared. OPAs are not used in staff training and there is no longer a requirement for new project proposals to refer to previous experience or self-evaluation findings as a matter of course.

[7] The 2016 changes were introduced in the Work Programme 2016-18 and Budget 2016 (BDS15-316) p.13, and described in more detail in the 2014 Annual Evaluation Review (BDS15-103) section 6.5.2.
[8] COMPAS: the MDBs' Common Performance Assessment System, an annual comparison exercise conducted between 2008 and 2013. It was developed by the MDB Working Group on Managing for Development Results as part of its efforts to develop and strengthen corporate results monitoring systems.

2.3.2 *Related activities such as TC and Policy Dialogue*

Evaluation Department guidance specifies that project-level evaluations and self-evaluation reports should cover other activities directly associated with the investment, including transactional TC, non-TC grants and directly relevant policy dialogue activities. In practice, compliance has been inconsistent. XMRs/OPAs have not generally rated the performance of other activities nor consistently drawn lessons from them as they have from the investment operations. Evaluation of non-transactional TCs and other stand-alone activities has been much more ad hoc.

TC assignments. Management operates an entirely separate self-evaluation system for TC assignments, without EvD involvement, and has done so for over 20 years. Operational staff prepare Project Completion Reports (PCRs) on completed TC assignments, using a template designed by Management. Its principal purpose is for reporting to donors. The PCRs are quality-reviewed by the Donor Co-Financing team and forwarded to donors, and the outputs also feed into the Annual Report, Sustainability Report and other reporting. Between 1998 and 2010, EvD reviewed and assessed a selection of PCRs each year, though this was never intended to be a representative sample. EvD has periodically evaluated large TC-funded programmes such as the Advice for Small Businesses Programme (ASB).

Non-TC grants and policy dialogue. There has never been a systematic process for evaluation or self-evaluation of these activities. Their variety is such that some (for example, incentive payments) are inseparable from the associated investment operations, while others (for example, stand-alone policy dialogue) fall far outside any definition of project-level evaluation.

3. Kirk Report recommendations and main findings

This section sets out the Kirk Report's key findings and recommendations.

> **"Building on the findings of this independent external evaluation, [EvD should] identify key issues and develop practical options for improving the EBRD self-evaluation system,** drawing on experience gained under the current system and evaluated experience of other MDBs, and present a report on issues and options to Board and Management by end-2019. Issues to be addressed include rating methodologies, alignment with ECG standards and assessment of contribution to transition impact."
>
> **"Management and EvD [should] work jointly to develop an effective and appropriate self-evaluation system for EBRD.** [Management should] set up a working group to consider the EvD analysis and its broader findings on self-evaluation in the Bank, and develop proposals, for Board review and approval, for an appropriate and more effective self-evaluation system in EBRD."
>
> **"Management [should] prepare a medium-term plan to strengthen the self-evaluation system and integrate it with the Bank's results architecture.** The plan will set out:
>
> – clear objectives and targets with appropriate performance metrics, detailing the steps required and the anticipated resource requirements.

- outline requirements for the transfer from EvD to Management of key responsibilities for managing the self-evaluation system.
- measures to strengthen knowledge management and organizational learning, drawing on experience of other MDBs, including mechanisms to systematically integrate relevant findings and lessons from evaluation into the design of new policies and projects.

Preparatory work for the external evaluation included a survey and interviews with many members of the Bank's Board and Management, and a review of practices at other MDBs and the UK Department for International Development. The final report's key findings relating to self-evaluation were:

- The self-evaluation system at EBRD is extensive and involves significant time and attention from operational staff and senior management.
- Among senior managers, this is widely perceived as an overhead rather than an essential feature of organizational learning and accountability.
- While many in Management appreciate the rationale for evaluation, some senior managers confuse monitoring with evaluation and evince scepticism about the benefits of the self-evaluation process ('too much, too late and too little learning').
- Although EBRD's Evaluation Policy assigns responsibility for ensuring 'an effective self-evaluation system' to Management, in practice it is EvD which runs the system and reports results. At present, Management does not 'own' the system.
- In the past, a key purpose for the self-evaluation system was to provide comprehensive, independently validated reporting on institutional performance. Following the changes in EBRD's results reporting system, this is no longer undertaken, thereby exposing the institution to reputational and operational risks.
- Given major changes in EBRD's results architecture, the self-evaluation system needs repositioning, reform and improvement. Success will depend on Management taking ownership of self-evaluation processes and integrating self-evaluation into its evolving results measurement systems. The required adjustments should distinguish between mandatory self-evaluation and demand-driven evaluation.

4. EvD – main findings and observations

EvD has commented widely and often on self-evaluation in the Bank in contexts ranging from project-level evaluation, to analysis of strategic or thematic issues, to annual reporting. The following points summarise EvD's views as they have developed out of experience in recent years.

4.1 Summary findings

- **Project-level self-evaluation reports (Operation Performance Assessments – OPAs) are produced but not used**. EvD mostly uses only validated reports that it has largely produced itself; there is little sign that Management uses OPAs at all.

- **Self-evaluation coverage is falling**. The system established in the 1990s envisaged 100% coverage of Board-approved operations, but coverage of frameworks has been inconsistent and partial. What was once a minor issue is now significant as investments under frameworks comprised 69% by number and 28% by volume of new EBRD commitments in 2018.
- **The unique value of an OPA is intended to flow from the capture of first-hand knowledge and insights of operational staff**. But in practice it is often produced by new or junior staff unfamiliar with the project who cannot add useful insights beyond what is contained in existing documentation; nor are they given any incentive to do so. Without any meaningful demand-side uptake of the work they produce their time is largely wasted.
- **OPAs are not shared or read by anyone outside the immediate project team** even though it is written as a stand-alone document apparently with sharing in mind. Business confidentiality concerns are cited as the reason, but internal transparency and adequate information security are not mutually exclusive.
- **It is hard to extract useful aggregate information from OPAs**: the most useful material is likely to be found in insights as to the drivers of performance; however, if this is available it is often very difficult to extract.
- **The approach and DAC terminology on which the OPA template has been built are unfamiliar to EBRD staff**. Many do not get the distinction between outputs, outcomes and impacts, and there is confusion where similar words are used differently - "transition impact" (EBRD) versus "impact" (DAC), or project design as distinct from (financial) structure
- **Because OPA scheduling is done individually, project-by-project, it works directly against assessing connected or closely-related projects more systematically** and from that extracting greater insight and value.
- **It is increasingly difficult to determine who is responsible for an operation at Portfolio Manager level**; the scheduling of OPAs is not always agreed with the person responsible for delivering them, since portfolio management was separated from project initiation.
- **Report content has failed to keep up with changes in the Bank's approach**. It assumes a stand-alone, Board approved investment operation and struggles to accommodate frameworks, increased integration with policy dialogue and donor funds, the place of the operation in the developing results architecture, strategic initiatives and new transition qualities.
- **The OPA's dominant focus is accountability with learning mostly an afterthought and no clear links between them.**
- **OPA quality varies widely**, and EvD comments on quality and suggestions for improvements are likewise addressed to variable extent.

- **Engagement of non-Banking departments in producing OPAs is neither consistent nor transparent**.
- **OPAs are often strong on client financial analysis but weaker in most other areas** - the areas where they should bring something distinctive and additional to other Banking reporting systems.
- **Project approval document templates do not facilitate consolidated reporting on project performance**. Despite recent improvements, it remains difficult to identify all the objectives of an operation (physical, financial, transitional, environmental, additionality-related), which are scattered in different parts of the approval document.
- **Self-evaluation reports identify very similar lessons over many years**. Either the authors are drawing on standard responses rather than really thinking about the specific project, or else the same issues keep arising and the lessons are not being learnt.
- **There is no consolidated Management reporting on OPAs at any level**. To EvD's knowledge the necessary resourcing has never been systematically assessed and is nowhere presented on a basis that allows further analysis; there is no capture of any self-evaluation-related metrics in Management Scorecards or work programmes.
- **While a Banking Department survey several years ago indicated some progress on the profile and perceived value of OPAs, EvD finds that little real change is evident.**

4.2 EvD's own contributions to systemic problems

Of course the problems identified above did not arise in a vacuum. EvD too must bear its own share of responsibility for the system as it exists today. A range of such issues was reviewed in the Self-Evaluation prepared by EvD as input to the Kirk Report, which confirmed most of them. There would be value in revisiting them as part of the Working Group process to inform the new work going forward.

- Excessive emphasis on accountability and problem-finding in the early years of independence;
- Reinforcement of the operations-side perception that the main purpose of self-evaluation was essentially to feed EvD's workplan;
- Too many changes/fine-tuning in the self-evaluation template (OPA); partly this reflected pursuit of an elusive gold standard in a system that could not produce it;
- The current format (largely developed by EvD) inadvertently encourages duplication and cut-and-paste from other documents.
- Excessive evaluation-speak with little real meaning for front-line bankers;
- Overly complicated formats; too much cut and paste; insufficient differentiation between the important and the trivial;

- Potential conflict of interest from EvD's preparation of retrospective results frameworks for Management for long-form OPAs and its quality review of OPAs which it will subsequently validate;
- Too little training by EvD on top of none by Management;
- Too little communications/advocacy; not enough getting Management on board;
- Lack of a genuinely useful means to collect and present lessons/insights.

5. Key drivers of current deficiencies

The deficiencies of the existing self-evaluation system both reflect and contribute to a wider range of challenges and concerns relating to performance targeting, reporting and analysis Bank-wide. Many can be considered legacy issues, reflecting the very particular circumstances of the Bank's first 20 years. While its mandate encompassed additionality, transition impact, and sound banking, by far the greatest emphasis for monitoring and reporting was sound banking. Business volume dominated operational targets; management priorities and systems were shaped accordingly; and Boards by and large endorsed this approach.

An important and long-running undercurrent, certainly for the Bank's first decade, was the largely unspoken notion that the Bank might "sunset" after a decade of operations. What all of this meant in practice was that the internal results management systems common to traditional multilateral development banks only began to be developed relatively recently. As a result the body of information and knowledge on wider transition results has been and remains quite limited.

To date most responsibility for reporting results and Bank performance in the wider sense fell to EvD through its Annual Evaluation Review, while the Bank's Annual Report and Financial Report discussed activities and financial results. But it is an essential responsibility for Management to report on all aspects of its performance, particularly results related to its core mandated purposes, and to do so in an way that shareholders find it most useful.

The origin of the practice was EvD's status before its formal independence in 2005. Originally, EvD was part of management with a dedicated Vice President for Evaluation assuring a degree of autonomy from other parts of the Bank. Early evaluation policies describe the learning element of evaluation as contributing to a "quality management" function.[9] As part of Management, EvD de facto acted as both a partial results unit as well as an independent evaluator. Independence in 2005 would have been the appropriate moment for Management to take over the results management function from EvD, but this did not happen. For many years, EBRD was the only MDB without a results unit. Today CSRM (Country Strategy Coordination and Results Management) has a more limited role and is more poorly resourced than the corresponding units at other MDBs. It does not publish an overall Development (or Transition) Effectiveness Report of any kind, as other MDBs do. Importantly, development of a results management framework at

[9] 1997 Evaluation Policy Update (BDS97-096) sections 1 and 2. Evaluation Policy Review of 2004 (BDS04-024) Preface, section 1.2(b)

EBRD has happened entirely separately from its self-evaluation system rather than incorporating it.

From EvD's perspective the roots of many of the problems identified with self-evaluation may be found in three fundamental issues, relating to the ownership, purpose and form of self-evaluation.

5.1 Project self-evaluation is an "orphan"

The external evaluation correctly identified the absence of Management ownership as a key reason for systemic insufficiencies. Lack of ownership explains why the Bank has not taken necessary decisions to ensure the relevance and usability of self-evaluation. This is reflected in:

– The failure to find and explicitly formulate an essential use for self-evaluation that would shape and drive its component parts.

– No overall quality control or Management-level accountability. EvD advises on OPA quality on a case-by-case basis. But there is no consistent process on the Management side for quality assurance, transmission up the decision chain, consolidation or reporting. Management itself has no quality standards in place, or any standard, mandatory review mechanism.

– There is divergence between basic operational approaches for project monitoring within Banking, on the one hand, and those for self-evaluation on the other. EvD is outside of Management processes for virtually all of this, and struggles to keep up with new developments after the event.

– There is no process or responsibility for the essential "downstream stage" of self-evaluation, namely synthesis, analysis and sharing of results and findings that would extract the greatest value from the time and effort put in.

– Responsibility for ex post project level performance evaluation was essentially "outsourced" by Banking to the internal Project Evaluation Department at the time of the Bank's creation; from an "ownership" perspective it has effectively remained there ever since.

5.2 The institutional purpose of self-evaluation has never been articulated

There has never been a serious discussion of what self-evaluation is intended to contribute or achieve.

– The central challenge of balance between accountability and learning (or even the meaning of these terms at EBRD) has not been explored. Results from self-evaluation are not incorporated into Management decision-making or reporting systems. From the project and unit level perspective there is little or no Management-level demand for self-evaluation work or insights.

- Self-evaluation findings are not shared within or across teams or departments; indeed the Bank's culture and confidentiality concerns actively prohibit this. Self-evaluation work is not presented in aggregate to inform strategic decision-making.
- Management views self-evaluation as a cost-centre and administrative burden rather than a useful instrument. It is not specifically resourced even though the most common complaint from project teams is that they lack the resources to do it effectively.
- EvD has pressed hard for OPAs to provide a comprehensive assessment of all elements of project design, including TC, policy dialogue, advisory, and cooperation with other IFIs, in addition to the traditional sole focus on the investment itself. Progress has been made but overall experience has been mixed and non-investment components consistently receive insufficient attention.

Box 1: Purpose of self-evaluation

Self-evaluation system design must flow from institutional agreement on the purposes it is intended to serve and what it is expected to accomplish. These can vary widely. The list of potential purposes provided below is illustrative but not exhaustive.

– Create an instrument useful for more effective operational management, including end-to-end project management and incentives.

– Provide the basis for objective project-level facts/analysis permitting aggregation into a reliable picture of higher-level performance – sector, department, country, instrument, vintage.

– Provide input/means to assess execution of institutional priorities/initiatives – i.e., connect projects to shareholder endorsed strategic objectives.

– Create incentives/pressures for more inclusive, integrated engagement of all relevant Bank units at project design stage and during project implementation.

– Provide integrated consideration of all elements of project design (policy dialogue; TC) as part of a cohesive overall assessment and narrative.

– Capture the real-time operational experience of professionals on the ground while it's still relevant, and use this to build institutional competitive advantage.

– Create an instrument/process to draw together all involved organisational units for ex-post, no regrets review of experience and extraction of lessons (After Action Review format).

– Provide real-time, evidence-based feedback to allow assessment and adjustment of metrics and indicators.

– Establish self-evaluation as a core contributor to a project-focussed component of a wider Knowledge Management system in the EBRD.

– Create a new means to encourage sharing experience and learning across institutional silos, and break them down in the process.

– Create processes/systems focussed explicitly on extracting insights from operational experience that can contribute to institutional competitive advantage.

– Create a quantifiable metric for inclusion in departmental and institutional level scorecards.

– Establish a core set of quality assurance activities that can be managed directly, incentivised, monitored and resourced transparently; create a new core-Management tool.

– Establish an empirical basis for regular reporting to shareholders.

– Establish the basis to reinforce institutional leadership in the IFI system on specific private sector related issues; self-evaluation builds EBRD's institutional brand and niche.

5.3 A single use instrument disconnected from key related processes

Because decisions have not been taken about the primary purpose of self-evaluation, the OPA is a compromise instrument that tries to cover all bases through a single, one-time instrument. It seeks simultaneously to fulfil both accountability and learning requirements. It tries to measure achievement of operational objectives (outputs) as well as longer-term outcomes and impact. It also seeks to capture lessons of experience and make them accessible and useful for the future. It seeks to provide EvD with raw material for validation reports and larger studies, and also to be a project-level case study that can be read on a stand-alone basis. All are valid objectives; but priorities must be established and from that appropriate systems developed and resourced.

Among the issues identified with the current situation are:

- The OPA timing is generally too early to capture transition results and impact but at the same time often too late for operational learning – especially in the case of follow-on projects.
- 100% coverage of operations, desirable for accountability and the aggregation of results and findings, cannot be achieved through a relatively heavy "case study" approach.
- But the "case study" approach only makes sense for a document that is shared and read widely by people unfamiliar with the project. This very rarely happens; and only a minority of projects are sufficiently interesting to justify a case study.
- Combining accountability and learning in one instrument tends to inhibit openness by operational staff about what could have been done differently.

6. Features of a modified system

The Bank needs a comprehensive project self-evaluation system that

- is owned by Management, useful to Management and used by Management
- is consonant with the Bank's systems and approach to operations
- enables Management to satisfy shareholder and public accountability and performance reporting demands
- provides evidence-based learning at both strategic and operational level
- serves a distinct, complementary purpose, taking account of the potential strengths and weaknesses of project self-evaluation within a broader monitoring and evaluation system.

A comprehensive system is one that captures all results (positive and negative) and learning from all projects (large and small, stand-alone and framework operations) and from all the elements of a project (investment, transactional TC and non-TC donor funds, and transaction-related policy dialogue).

The design of instrument(s) of self-evaluation follows on from decisions on its primary purpose.

Box 2: Observations regarding desirable changes

The mostly well-established EvD findings described in sections 3 and 4 are the key drivers of the systemic features and concerns identified in the Kirk Report. They also underpin a number of observations and suggestions for improvement that EvD has made – both formally and informally – over the years. These are directly relevant for the efforts of the Working Groups and may be summarised as follows

- The primary measure of success is that the system is relevant to EBRD and yields something that is useful to Management and to the Board. In EvD's view, this is more important than conformity to the detail of international standards of evaluation.
- Management responsibility should reside at a sufficiently senior level to allow resourcing decisions, coordinate work across units and provide credibility within the Bank and to the Board.
- A strengthened Results Unit is needed with the capacity to manage the system, assure quality control and make use of the outputs through aggregation and reporting of results, and synthesis and dissemination of findings.
- As much attention should be given to the use of self-evaluation as to the quality of the inputs, particularly the use of learning derived from self-evaluation.
- A single, one-time reporting instrument as currently used is unlikely to be the best means of addressing all requirements.
- Developments under OE&E, if pursued further, may ultimately allow the Bank to report on performance and results more efficiently than through a project report, but this is some way off.
- Credible reporting and analysis of performance and results should be a clear priority for Management in its own right. There are multiple ways for Management to increase quality assurance, objectivity, and constructive feedback without any role for an independent evaluator.
- However some means of independent validation by EvD is both necessary and valuable
- Numerous essential preconditions for value-adding ex post self-assessment lie solely in Management's hands. Several are built into the Evaluation Policy, such as availability of baseline data, performance metrics that are clearly articulated and monitorable.
- Improved reporting on results will also benefit from increased monitoring of actual results on the ground. This is something for which substantial additional resources are required if the Bank wishes to produce credible results commensurate with its move towards a more typical development bank model and increasing reliance on concessional funding.

6.1 Owned by Management, useful to Management and used by Management

Self-evaluation should in the first instance be a tool to meet the needs of Management in order to deliver on shareholder expectations. This is likely in some respects to come at the expense of achieving a theoretically superior evaluation system. The most important characteristic of a self-evaluation system is that it is useful to, and used by, Management.

Ownership means Management takes responsibility for:

- definition of the primary focus
- design of instrument(s)
- scheduling of project self-evaluation or its various elements, development of processes, administration
- training staff in preparation of project reports
- quality control of project-level reporting
- providing sufficient resources and integrating the activity fully into incentive and accountability systems
- aggregation, synthesis and analysis of the findings and data
- use of outputs in reporting on institutional performance and results, strategic decision-making, operational performance management and staff training programmes
- annual reporting to the Board on "its scope and operations" as required by the Evaluation Policy.

Management taking ownership also means that EvD must relinquish ownership and not try to define precisely the content or form of self-evaluation. It should contribute from an independent position only. While EvD has direct and substantial interests in the outcome, the OPA should not be designed around providing information that EvD needs for special studies.

It is Management's choice where to position a self-evaluation system within Management. EvD sees the following requirements:

- a responsible unit/person at VP level, able to make resourcing decisions, support specific initiatives that would strengthen or streamline self-evaluation (see Boxes 6 and 7), and provide credibility within the Bank and to the Board
- a strengthened results unit with the resources to manage the system and to make use of the outputs through aggregation and reporting of results, and synthesis and dissemination of findings.

Assigning the results unit responsibility to a strengthened CSRM, in the Policy & Partnerships Vice Presidency (VP3), would contribute to the integration of self-evaluation with developing results reporting processes and facilitate the integration of support unit perspectives with the dominant Banking perspective of the current system. However, it risks allowing self-evaluation to remain a secondary consideration to Banking, an add-on or afterthought to the main project cycle; this has been a critically important driver of existing systemic deficiencies.

Placing the responsibility within Banking could lead to a more integrated and intrinsic system, with self-evaluation as a standard and mandatory element within the project monitoring and management process. It would be critical to ensure that is did not exclude or marginalise the valuable perspectives of departments such as Credit, Procurement, ESD. It is an unmistakable feature of the Bank that “non-core” functions tend to be “outsourced” internally; they generally orbit around Banking and often have to fight for voice. And there is no single unit managing portfolio monitoring in Banking: Banking Portfolio Group covers corporate debt and sustainable infrastructure debt, plus an analytics team. Direct equity, equity funds and Financial Institutions debt operations are each managed by separate units. A single point of responsibility is crucial.

6.2 Consonant with the Bank's systems and approach to operations

In order to be useful to Management, the system should in the first instance be tailored to EBRD rather than impose international standards. While EvD’s independent evaluation work is expected to meet international standards, operational staff assessing individual project performance cannot be expected to have the same detailed understanding of evaluation approaches and should not be asked to conform to unfamiliar standards. Specifically, the self-evaluation system should:

– assess performance and results against the Bank’s mandate, strategies and policies, and individual operation approval documents

– use EBRD terminology that Bankers understand: transition impact, sound banking, additionality; this doesn’t mean banishment of DAC terminology (relevance, effectiveness, efficiency, outputs, outcomes, impacts) but rather ensuring that it doesn’t become the sole vocabulary

– be aligned to developments in results architecture and reporting, transition qualities, Public Information Policy, and the increased use of operations approved under delegated authority

– link all the elements of a project: investment, donor-funded assignments and policy dialogue, as foreseen in Project Monarch.

6.3 Enables Management to satisfy accountability demands

The Bank faces multiple legitimate demands for accountability, both formal and informal.

– its Board of Directors, which has dual obligations as country representatives and as officers of the Bank

– its donors, whose contribution is becoming larger and more central to the Bank’s operations

– external bodies with standing on specific issues (G20; UN; EC) and in relation to the EBRD’s commitments to global agendas such as the Sustainable Development Goals.

The recent G20 comparability exercise highlighted an issue raised in the external evaluation and in the comments of the Independent Evaluation Adviser: the lack of “comprehensive, independently validated reporting on institutional performance” (chapter 4) or of “the accountability for results dimension” (annex 6). Providing such accountability requires contributions from

Management through self-evaluation and from EvD as the independent validator. A review of Development Effectiveness reports of other MDBs reveals that they all base assessments of performance on self-evaluations which have been validated by independent Evaluation Departments.

Box 3: Ratings

The issue of project performance ratings is at the heart of decisions about the purpose of self-evaluation and will need to be resolved as part of the redesign process.

Ratings can provide a structure and consistency to project evaluation; allow aggregation for an overall performance assessment; help to identify project elements that did not go as planned; and provide comparability between projects and across time, project type and against other institutions. International standards (see section 2.1) generally regard the use of ratings as good practice.

Ratings can also become a straitjacket, particularly in self-evaluation, transforming what should be an objective and reflective process into a litigious and defensive one. A predictable effort to present the project in the most favourable light possible will come to the detriment of real accountability and learning. Views of evaluators are split between those who find performance rating incompatible with learning and those who consider it the necessary starting point to identifying what could be done better.

EvD's revision of the self-evaluation process in 2016 replaced EBRD-specific ratings with internationally comparable OECD-DAC ratings. With hindsight, EvD regards this as a negative development. The process became more detached from normal Bank activities and less meaningful to staff preparing OPAs. DAC terminology is only useful if the Board and Management determine that the principal purpose of self-evaluation is to provide comparability across IFIs. If the primary purpose is internal accountability, learning and performance improvement, EBRD terminology is clearly preferable.

It is primarily for the Board to define what constitutes success in the EBRD context, the reporting it requires to assure itself of the Bank's performance, and the resources it is prepared to allocate to this function. Resources in this context means resources provided to Management and EvD, and also the degree of scrutiny the Board is able to provide.

While substantial progress has been made in corporate reporting of results in recent years, Box 4 below describes some remaining weaknesses.

Box 4: Critique of existing reporting at EBRD

This topic has been touched upon by several EvD reports in recent years, including sector studies, project-specific studies and the review of country strategies. Key issues include:

- The focus of reporting is on inputs not outputs: portfolio allocation not results.
- Reporting on outcomes and impacts is almost completely lacking and no resources are allocated for follow-up or validation of expected impacts. Even reporting on "green" projects simply aggregates ex ante calculated carbon savings rather than what may actually have been accomplished.
- EBRD reports specific successes, not the overall rate of success: coverage is partial and selective and there is no sense of what proportion of projects or total investment are contributing meaningfully to reported achievements.
- Separate elements are reported separately: Bank financial performance, green performance, TIMS reviews, TC: there is no overview of entire performance of a project.
- EBRD is not able to provide comprehensive responses to inter-MDB comparison exercises.
- There is no measure of the contribution made by donor-funded activities or policy dialogue to overall performance.

There are two distinct elements of accountability: performance and results.

Reporting on **aggregate performance** would allow the Bank to answer questions such as:

- What proportion of projects performs satisfactorily in terms of additionality, sound banking, transition impact and sustainability (mandate pillars)?
- What proportion of operations achieves the objectives established at approval?
- What is the aggregate performance of Bank operations approved (or signed, or evaluated) in a given year compared with previous years, or with other IFIs?

Reporting on **aggregate results** would allow the Bank to answer questions such as:

- What is the Bank's development/transition effectiveness overall? What has it got to show for the money and guarantees contributed by shareholders and donors?
- Which strategic initiatives or transition qualities are being successfully targeted by the Bank's projects?
- What is the total effect of the Bank's projects in a particular country, sector or target area over time?

The two approaches need not be mutually exclusive, but they are different and require different questions, different data and different timing and coverage.

6.4 Provides evidence-based learning at both strategic and operational level

In contrast to accountability, the demand for learning is likely to be largely internal to the EBRD and Board. The main audiences for learning are

- The Board and senior management as strategic decision-makers
- Front-line staff looking for operational learning and experience
- Training providers seeking case studies as input to staff training programmes.

Learning at the strategic level is closely related to performance and results measurement. It will be concerned with the comparative performance or results of different types of operation. It should address questions such as

- What changes should we make when next replenishing this programme/initiative?
- What kinds of project generally perform most satisfactorily or show the greatest problems?
- What are the most common reasons for projects meeting or failing to meet expectations?
- As an approver, what should I look out for when a new project (of a specific type, or in a specific country) comes for approval?
- In case of a macroeconomic or political event (devaluation, a coup d'état), what proportion of projects have failed and what were the characteristics of those that survived?

Learning at the operational level is concerned with the practicalities of project implementation. It addresses questions such as

- What innovative approaches or products have worked well or badly?
- How often has political risk been realised and under what conditions has the Bank been able to intervene successfully?
- What practical approaches to policy dialogue have worked in the past?

Both strategic and operational learning could be used as inputs to a more developed quality-at-entry mechanism. This could be inspired by – but need not slavishly follow - the former Quality Assurance Group (QAG) at the World Bank. It might be better tailored to EBRD as an informal, pre-OpsCom, no-fault consultation process. It would consider issues such as design, evaluability and incorporation of past experience into project design. By bringing together or making available views and findings from across the Bank, outside the individual operation team, it would contribute to breaking down silos and developing a "One Bank" approach.[10]

[10] Two EvD papers -- Evaluability – is it relevant for EBRD? (PE11-540S/SGS12-182); and, Performance metrics – how well do EBRD projects specify expected results? (PE12-560S/SGS13-113) -- discussed some of these points in more detail and presented a draft evaluability checklist which would need to be revised and updated if considered for adoption today.

Until now, self-evaluation has addressed learning through a section at the end of the report requesting individual, stand-alone lessons. EvD stores these lesson "nuggets" in a database for future reference. This is not the only possible approach to learning and is probably not the most useful. Some key points:

– Learning does not have to be limited to individual lessons as nuggets for filing and quoting. It could also be factors and patterns of performance.

– Specific lessons can be very context-specific and difficult to translate to a different set of circumstances.

– Learning is not about correcting the course of the current project - that is a function of monitoring, not of evaluation. It is about learning for future projects or strategic decisions.

– A single lesson, provided by an operation team and then fed back to that team, is not useful. The value is in aggregation and synthesis.

A more imaginative approach requires an additional stage of analysis beyond the production of the project report. This has to be done by a central results unit, which can transform multiple individual findings into a digestible form for wider sharing, avoiding the problems of confidentiality or concerns over a "blame culture" that inhibit the sharing of individual project reports. It will tailor learning products to the audience: Board and senior management may require syntheses of findings, operational staff may find specific "nuggets" of information useful, while training providers will need case studies of a small number of interesting projects. This report leaves further discussion of Knowledge Management to the forthcoming EvD report on the subject.

6.5 Serves a distinct purpose within a broader Monitoring & Evaluation system

Project self-evaluation should play its role within a broader monitoring and evaluation system. It does not necessarily need to cover all the elements of evaluation in itself. It is better suited to some aspects of evaluation than to others.

What is special about project self-evaluation? To state the obvious, it is **evaluation** (in contrast to monitoring) conducted at **individual project level** by **staff closely involved in the implementation** of the project.

Monitoring versus evaluation

The Kirk Report highlighted Management confusion about the difference between monitoring and evaluation.

Monitoring	Evaluation
• Takes place periodically during implementation	• Takes place one-time when operational
• Can be used to clarify/modify expected results and associated targets	• Takes Board-approved expected results and targets as given – assesses against all stated (or inferred) results whether monitored or not
• Translates objectives into a limited set of performance indicators that should show movement during implementation	• Uses a wider range of indicators and all available evidence to assess realised results
• Tracks delivery of inputs, conduct of activities and whether expected results achievement is on- or off-track	• Assesses achievement of expected and unanticipated results and the causal contribution of activities to such
• Reports progress to managers and alerts them to problems so that corrective action can be taken	• Provides findings, lessons and recommendations to improve future operations • Among its performance attributes, goes beyond results to assess relevance, process and allocative efficiency, and sustainability • To the extent possible, incorporates a counterfactual into performance assessment

Source: EvD Guidance Note 6 on Evaluation Performance Rating

– The confusion between monitoring starts at project design and approval stage, where the approval document establishes one set of targets for "measuring/monitoring success".

– Self-evaluation adds an overview of whether the project achieved what was expected of it in a broad sense, not just whether it hit monitoring benchmarks.

Implications of **project-level quality** (versus thematic, strategic or programmatic evaluation):

- EBRD is a project bank and its impact is the aggregate of project-level effects; it is not possible to account for contribution/attribution without project-level effects.
- It is unlikely that broad developments, patterns of performance or transition impact will be visible at individual project level (other than very large projects or in very small countries) and individual bankers should not be asked to try to speculate on such effects in a project-level report.
- Some project detail may not be suitable for sharing widely, or not of interest to other teams; raw findings often not directly transferable to other projects.

Implications of **self-evaluation quality** (versus independent evaluation):

- Likely positive bias in qualitative assessment (hence not reliable without validation).
- Staff have greater knowledge of the project and the reasons for its success/failure than an independent evaluator can have.
- Staff are more likely to learn from their own or their peers' experience than from an outside evaluator.

6.6 The self-evaluation instrument(s)

The choice of instrument depends on the choices discussed above: what kind of performance or results reporting the Board requires, what kind of learning is useful to Management at a strategic or operational level, how the outputs will be used.

The single, one-time project report is the standard instrument for self-evaluation in MDBs and other development institutions. Section 5.3 above identified some problems this approach creates at EBRD. EvD believes that such **an instrument alone is unlikely to be the best means of addressing all the requirements of self-evaluation**, especially in the longer term.

EBRD is not unique in using a single reporting instrument for self-evaluation: this is common practice at IFIs and underlies international standards such as the ECG Good Practice Standards. But all the IFIs are grappling with similar problems to the EBRD and none has yet found its self-evaluation system entirely satisfactory. The EBRD has the opportunity to develop a tailored approach which can do better.

The boxes overleaf address three areas of interest for developing a more effective and innovative approach to self-evaluation: Box 5 considers the timing of different elements of self-evaluation and Boxes 6 and 7 highlight some current Bank initiatives on reporting and learning. All are worth exploring but EvD detects a lack of coordination and ambition to achieve 100% coverage and full integration into the project cycle – largely because many of these initiatives are being developed at an operational level without adequate resourcing or senior Management buy-in.

Box 5: Timing and coverage of self-evaluation

Section 5.3 identified the timing of self-evaluation as an unsatisfactory compromise. This is one of the key arguments against a single, one-time self-evaluation report.

- **Learning** is an ongoing process that needs to be captured in real time. The problem of loss of institutional knowledge when a project has had multiple operation leaders really shows that self-evaluation is not capturing learning as it should: self-evaluation should contribute to mitigating this problem, not suffer from it.
- **Accountability for performance** can be captured through a one-time review as now (but more briefly), perhaps in the form of a project completion report with 100% coverage and scheduled shortly after the last milestone comes due. A brief PCR for every operation – standalone and sub-operations - designed to extract aggregable data would allow performance to be reported at the level of countries, sectors, frameworks, initiatives and project types.
- **Accountability for results** requires a later follow-up to assess longer-term effects and sustainability of a project's results. Coverage can be partial, either on a random sample basis or more likely through a deliberately selective process, grouping related projects with a common objective or in a specific country or sector. It would be focused heavily on transition, assessing the soundness of the original design, accuracy of the theory of change, role of policy dialogue and donor-funded activities, and collaboration with other IFIs. It requires projects to have trackable transition metrics, and a resourced monitoring plan to track them. It is likely to contribute strongly to the learning aspect of evaluation.

Box 6: Bank initiatives on reporting

At present, several parts of the Bank are pursuing worthwhile initiatives on reporting, formally or informally, to address many of the issues raised.

- CSRM has made major changes to the results architecture of the Bank and developed the compendium of standardised indicators which will enable improved aggregate reporting
- EPG is working on Project Christopher to build on the compendium to allow a degree of automation in setting and monitoring TI indicators and to expand TIMS coverage
- Donor Co-Financing is piloting templates that would bring together reporting on TCs and investments
- Project Monarch is working on a "whole project" approach, linking investments to transactional TC and policy dialogue
- OE&E is developing a project dashboard intended to bring together the performance of a project in several areas: financial, transition, environmental.

Box 7: Bank initiatives on learning

- Portfolio Management Group has developed "financials capture", entering client financials into spreadsheets allowing performance analysis across regions, sectors and project types
- PMG and CSRM are trying to build on financials capture to incorporate other factors of performance and possibly achievement of transition benchmarks to allow factor analysis
- Business Development COO facilitates sessions where operation teams share best practice case studies as a way of sharing experience of successful and novel operations
- The Communities of Practice initiative brings together experience from different parts of the Bank on specific themes (e.g. state-owned enterprises, disruptive technology)
- EPG manages the Policy Academy which brings together EBRD and external experience in areas of policy dialogue
- Corporate Recovery shares some of its findings with Banking on weaknesses in project structuring and handling that have contributed to projects being transferred to them

7. Some next steps for the working groups

Many of the points raised in this document require more detailed analysis and review of options, followed by senior level decision-making. The Working Groups should:

- survey Board and Management views on the desired features and purposes of self-evaluation and review how other institutions answer this question. Important issues include: scale of incremental resources needed for an effective end-to-end project monitoring and self evaluation system; perceived importance of project-specific ratings and capacity to produce internationally comparable performance reporting;
- survey budget and strategic planning instruments in other IFIs to see how they address the matter of resourcing for project monitoring and self-assessment; and,
- consider the inclusion of self-evaluation metrics in scorecards and the specification of both short-term targets and long-term results in project approval documents.