ANNUAL EVALUATION REVIEW 2013

European Bank
for Reconstruction and Development

www.ebrd.com

The 2013 *Annual Evaluation Review* includes:

- assessment of the performance of the European Bank for Reconstruction and Development's (EBRD's) mature operations based on a mix of in-depth evaluations by the Evaluation Department (EvD) and validated self-evaluations;
- the main findings from evaluation studies completed in 2013 as EvD's principal contribution to learning;
- a more in-depth study of the evaluability of 2013 project proposals for clarity on expected results and risks;
- a quality check of self-evaluation and transition impact monitoring systems in the EBRD
- a review of EvD's activities and performance during 2013 to highlight achievements and challenges.

Authors

Victoria Millis
Principal Evaluation Manager

Shireen El-Wahab
Principal Evaluation Manager

Saeed Ibrahim
Evaluation Analyst

Keith Leonard
Senior Advisor

Abbreviations

AER	Annual Evaluation Review
BEPS	EBRD's Banking Environment and Performance Survey
CA	Central Asia: Kazakhstan, Kyrgyz Republic, Mongolia, Tajikistan, Turkmenistan, Uzbekistan
CEB	Central Europe & Baltics: Croatia, Estonia, Hungary, Latvia, Lithuania, Poland, Slovak Republic, Slovenia (& formerly the Czech Republic)
CRR-5	The Bank's 5th Capital Resources Review, due to take place in 2015
DAC	The Development Assistance Committee of the Organisation for Economic Co-operation and Development
EBRD	European Bank for Reconstruction and Development
ECG	Evaluation Cooperation Group [of the multilateral development banks]
EEC	Eastern Europe and Caucasus: Armenia, Azerbaijan, Belarus, Georgia, Moldova, Ukraine
ENE	Energy: Natural Resources and Power & Energy sectors
ESRB	European Systemic Risk Board
ETCI	Early Transition Countries Initiative
EU	European Union
EvD	Evaluation department
FIN	Financial Institutions
FX	Foreign exchange
GPS	Good practice standards for private sector evaluation of the ECG
ICA	Industry, Commerce and Agriculture: Agribusiness, Equity Funds, Information & Communication Technologies, Manufacturing & Services and Property & Tourism sectors
ICT	Information & Communication Technologies
IFI	International financial institution
INF	Infrastructure: Municipal & Environmental Infrastructure and Transport sectors
LEF	Western Balkans Local Enterprise Facility
MEI	Municipal & Environmental Infrastructure sector
OCE	Office of the Chief Economist
OE	Operation evaluation
OECD	The Organisation for Economic Co-operation and Development
OEC/DAC	The Development Assistance Committee of the Organisation for Economic Co-operation and Development
OPA	Operations Performance Assessment
OPAV	OPA validation
RUS	Russia
SEE	South Eastern Europe: Albania, Bosnia and Herzegovina, Bulgaria, FYR Macedonia, Montenegro, Romania, Serbia
SME	Small and medium-sized enterprises
TC	Technical cooperation
TC Com	TC Committee
TI	Transition impact
TIMS	Transition Impact Monitoring System

Defined terms

Evaluability	The extent to which the value generated or the expected results of a project are verifiable in a reliable and credible fashion
Impact	The positive or negative long-term effects produced by an intervention, directly or indirectly, intended or unintended; an impact generally results from a series of causal factors of which the project is but one
Indicator	A quantitative or qualitative factor or variable that provides a simple and reliable means to measure achievement, to reflect the changes connected to an intervention, or to help assess the performance of a specified entity
Outcome	The short-term and medium-term effects consequent to delivering the intervention’s outputs
Output	The products, capital goods and services that result from an intervention - its deliverables
Quality-at-entry	A comprehensive check on all aspects of design integrity of an intervention and its alignment with polices and strategies - incorporates evaluability
Result	The output, outcome or impact (intended or unintended, positive or negative) of an activity or intervention


1823

Highlights

Aggregate performance of the EBRD's operations

- Projects approved in the period 2008-10 had a 62 per cent success rate overall, significantly above the average of 56 per cent for the period 1993-2010,
- Transition impact rated *'satisfactory'* or better in 88 per cent of cases and *'good'* or *'excellent'* in 57 per cent of cases, compared with 79 per cent and 54 per cent for the full period 1993-2010.
- The previous under-performance of projects in Central Asia has reduced in recent years, though the numbers are too small for statistical significance.
- Overall performance shows a pattern of improved performance from a low point among projects approved in the immediate pre-crisis period around 2006-07. It is driven by a parallel pattern of improved performance in several component ratings. This mirrors the late 1990s and early 2000s when projects approved shortly after the 1998 Russian financial crisis out-performed those approved before.
- The performance rating methodology is under revision to provide for more transparent and consistently applied criteria and sub-criteria ratings to provide a stronger story for the overall rating.

Main findings from evaluations completed in 2012

- The results agenda remained an important cross-cutting issue from 2013 evaluations and validations, in particular the continued weakness in ex ante results specification despite progress being made. EvD's work has fed into initiatives by Management to develop results frameworks and better define performance metrics across both technical cooperation (TC) and investment operations.
- Evaluations of TC programmes highlighted the importance of embedding policy dialogue activities in the results framework of a donor fund and, where possible, in individual TC operations.
- In-depth evaluations of two crisis-response operations cast further light on the outcomes of projects approved in the immediate aftermath of the financial crisis.
- A study on achieving equity investment objectives made recommendations on enhancing the focus on results, improving monitoring and reporting and modernising Board approval documents. These recommendations were well timed to feed into ongoing Management reviews including on the wider monitoring process.

Evaluability assessments of operations approved in 2013

- In support of the Bank's increased results focus, EvD applied an evaluability checklist to samples from both TC Com submissions and investment operations presented in 2013, to suggest ways to communicate project proposals that are more understandable to wider audiences and to better manage operations to demonstrate achievement of outcomes and impacts.
- The sample of TC results matrices scored very well, though challenges included weaknesses in the presentation and measurement of TC contribution to the overall transition impact of projects, and whether the results framework contains all outputs required to achieve larger objectives.
- The investment operations evaluability assessment suggests that there is still a gap between project evaluability as per the checklist, and comments on quality found in OPA Validations, though it sheds more light than the previous exercise conducted in 2011. The checklist needs further modification to provide clearer guidance to operational staff on how evaluability can be improved.

On self-evaluation and transition impact monitoring

- EvD has developed a new quality review checklist to enable a more systematic review of the quality of Operation Performance Assessments (OPAs) and allow more consistent feedback to portfolio teams on the quality of their OPAs.
- EvD applied the checklist to the 82 OPAs delivered in 2013 and found that on average they were of good quality, scoring 78 per cent across the five dimensions of the checklist. The OPA sections with the lowest average scores were project efficiency (64 per cent) and project impact and sustainability (69 per cent)
- Overall the exercise showed that since last year the Bank has gone some way toward greater ownership and more objective self-assessment. EvD will use the findings of the exercise to improve its OPA guidelines and offer further training to Banking teams on OPA preparation. It gave publicity to the best OPAs and the best parts of selected OPAs to illustrate (and reward) good practice.
- The binary ratings gap between self-evaluation ratings and EvD ratings has fallen, on a five-year rolling average, from 19 per cent of ratings downgraded by EvD in 2007-11 to 15 per cent in 2009-13. Reducing this 'disconnect ratio' to below 10 per cent should be targeted.
- *Ex ante* transition potential was substantially achieved in 77 per cent of cases, and transition risk as assessed at appraisal was a good predictor for the likelihood of achieving satisfactory transition impact at evaluation. EvD's assessment of transition impact at evaluation concurred with that of the Transition Impact Monitoring System (TIMS) in 83 per cent of cases. For completed operations with a final TIMS rating, the figure was 86 per cent.

Aggregate performance

In this Chapter:

- evaluation results for the representative sample of randomly selected operations
- presentation of overall performance of evaluated projects
- review of aspects of performance according to the OECD Development Assistance Committee (DAC) evaluation criteria: relevance, effectiveness, efficiency, and impact and sustainability
- a discussion of patterns emerging in the performance of operations following the post financial crisis
- an update on EvD's review and update of its performance rating methodology.

How operations are evaluated in the EBRD

Rigorous self-evaluation by operations teams is integral to any effective continuous evaluation system. Under the EBRD Evaluation Policy self-evaluations are conducted by operations teams through the preparation of OPAs for all mature projects. The OPA contains Banking's self-assessed performance ratings across key areas and lessons for future application. EvD assists teams preparing the OPA by conducting a review, but the final OPA is exclusively Management's product.

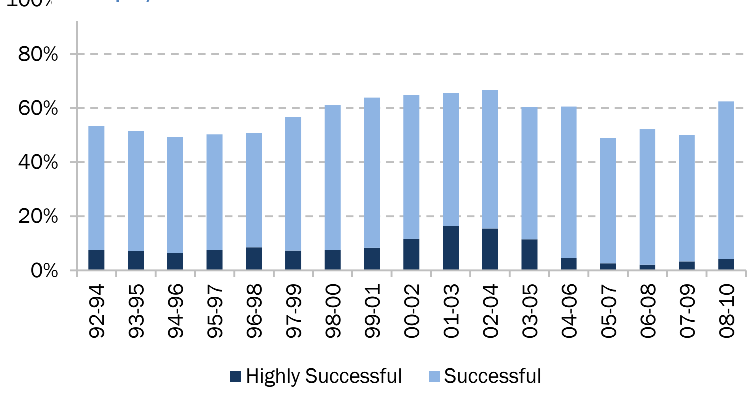
A representative random sample of evaluation-ready projects is selected annually, to be independently evaluated by EvD through either an OPA Validation (OPAV), or a more in depth Operation Evaluation (OE). OPAVs provide independent performance ratings based on EvD analysis, which are compared to OPA self-ratings. They are primarily a desk effort, but may also result in a judgment that additional analysis is necessary, including at the field level. Management is invited to comment on OPAVs, but no endorsement or agreement is requested or given. EvD ratings then feed into the aggregate ratings of institutional performance presented in this chapter. EvD also selects a smaller number of self-evaluated operations for full, independent evaluation through an OE or special study. Operations for full evaluation are chosen deliberately based on their potential to produce operationally useful findings, because particularly important or challenging features have been identified, or for input to a broader study such as a sector strategy evaluation. Annex 1 describes the selection and evaluation process in more detail.

Overall performance

Based on three-year averages (to smooth year-on-year variations), the success of evaluated projects has not varied substantially between the early 1990s and the most recent period (Figure 1). Sixty-two per cent of projects approved in the period 2008-10 were rated successful or better compared with 54 per cent of those approved in the period 1992-94. Aggregate levels of successful (or better) operations peaked at 68 per in the period 2002-04 before decreasing again. There is therefore broad consistency in the institutional success rate of EBRD projects in the range of 50 to 60 per cent, without a sustained long-term trend towards either improved or deteriorating performance.

Because the EBRD is generally intended to go where others do not (the Bank is required to be additional) a very high level of success in absolute terms cannot be expected. However, since success here is measured primarily against expectations at approval the question arises whether Management and shareholders should be satisfied with 60 per cent success and no long-term trend of improvement. Of course, ambitions could have increased over the years and this too may contribute to static performance against expectations.

Figure 1: Overall performance by approval year (rolling three-year sample)



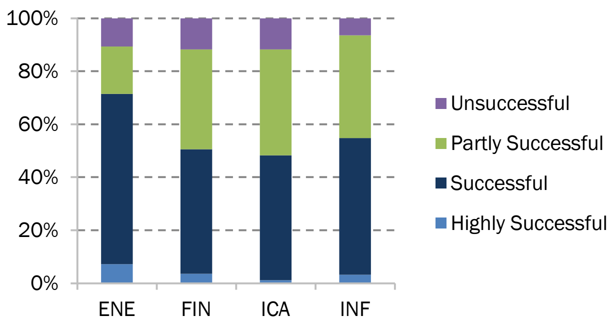
Some of the rise and fall in performance over the last two decades can be explained by the two major financial crises affecting the EBRD region during that time: overall performance was lowest among operations approved in 1996-98 (immediately before the 1998 Russian crisis) and 2005-07 (immediately before the recent financial crisis). The most recent approval years for which results are available are 2008-10 and they have shown a rebound in terms of the success of the projects. This is commendable bearing in mind the challenging environment in which the projects approved in the period immediately after and in the midst of the crisis, were being implemented.

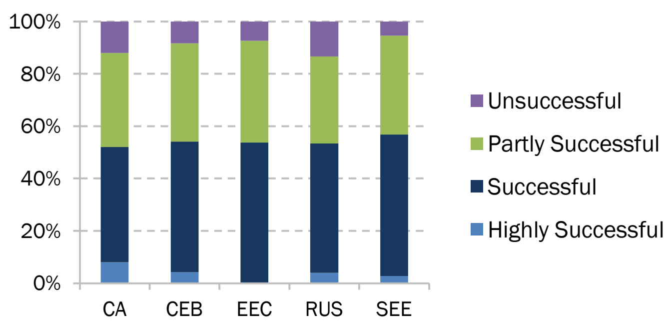
Overall performance by region and sector

In contrast to previous years, a regional comparison of performance shows that the central Europe and the Baltics states (CEB) and south-eastern Europe (SEE) regions are no longer significantly outperforming other regions in terms of overall performance ratings, including Central Asia (CA) and Russia (RUS) who have previously relatively underperformed. The CA region has shown the most improvement in recent years to the extent that it is now performing at roughly the same level as other regions whilst continuing to maintain a relatively high proportion of *'highly successful'* projects compared to other regions.

On the other hand, relative performance by industry sector broadly continues with previous trends, with projects in energy (ENE) performing particularly well between 2005 and 2010. EvD has noted in previous AERs that projects in the financial institutions (FIN) and industry, commerce and agribusiness (ICA) sectors were more likely to have been seriously affected by the recent economic upheavals considering that these sectors are overwhelmingly private sector and open to direct competition, whereas energy and infrastructure sector projects are more likely to be in the public sector or in regulated utilities with a more reliable cash flow. This would go some way to explaining the difference in overall performance ratings among the sectors.

Figure 2: Overall performance by sector and region, projects approved 2005-10



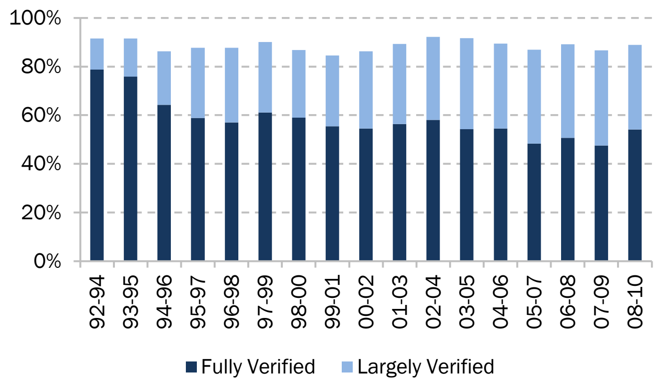


ENE: Energy; FIN: Financial Institutions; ICA: Industry, Commerce and Agriculture; INF: Infrastructure
CA: Central Asia; CEB: Central Europe and Baltics; EEC: Eastern Europe and Caucasus; RUS: Russia; SEE: South-Eastern Europe.

Project relevance and additionality

Evaluation reports include a description of the rationale of the project under the Bank's mandate and rate additionality under this section. Figure 3 below shows the ex-post assessment of additionality in evaluated projects. The results show that while around 90 per cent of projects have consistently been rated '*largely verified*' or better, the proportion achieving the highest rating of '*fully verified*' had been falling progressively to just below 50 per cent from 2007 to 2008. A possible explanation is that prior to the financial crisis the EBRD was becoming less additional in absolute terms but without this being fully reflected in projections at approval. In line with this theory, additionality has improved again since the financial crisis, with the number of projects receiving the highest rating of '*fully verified*' increasing to almost 55 per cent between 2008 and 2010.

Figure 3: Additionality by approval year (three-year rolling sample)



Project effectiveness and efficiency: achievement of objectives, project financial performance and Bank handling

Achievement of operational objectives and project financial performance, which together measure the Bank's effectiveness, are two of the indicators most closely correlated with overall performance. It is not surprising then that both show a performance pattern closely matching that seen in Figure 1, with higher scores for projects approved around the period from 2000 to 2006. Although results then declined modestly, performance has bounced back since 2007.

Projects are achieving more of their objectives after a lull during the crisis

Achievement of objectives

The three year rolling figures show a clear long term improvement in the achievement of operational objectives, levelling off at around 60 per cent of projects having '*good*' or '*excellent*' achievement of objectives up to 2006. Thereafter, there is a considerable fall, with just 46 per cent of projects approved at the height of the financial and economic crisis showing '*good*' or better achievement of objectives. After this nadir, there has been an improvement back towards the long term trend, with 61 per cent of projects approved between 2008 and 2010 realising '*good*' or '*excellent*' achievement of operational objectives.

The achievement of operational objectives generally equates with the production of outputs, the results area most under the control of the project itself. Therefore, one would expect a high level of achievement. That this is not always so does raise the question as to whether the statement of projected results in terms of operational objectives might, in some cases, be over-optimistic.

Figure 5: Achievement of objectives by approval year (three-year rolling sample)

Figure 4: Financial performance by approval year (three-year rolling sample)

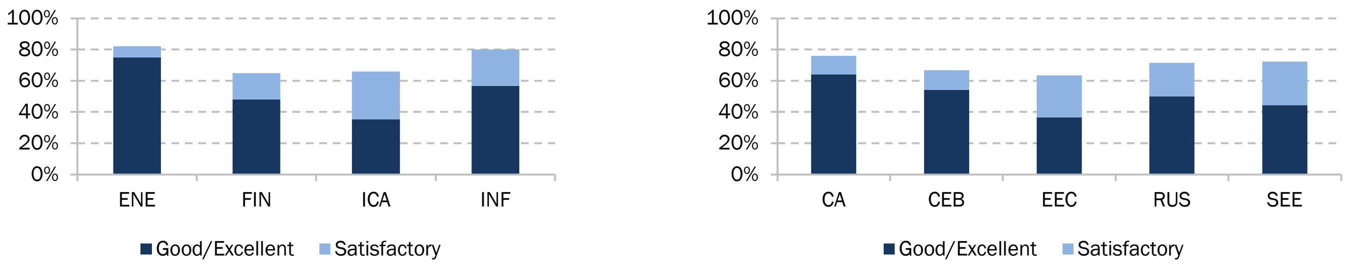
Project financial performance

The financial performance of EBRD projects follows a similar trend. The figures show a steady improvement in financial performance of operations from the low of around 25 per cent rated as '*good*'/'*excellent*' from 1994 to 1996 to a peak of just over 60 per cent rated as such from 2001 to 2003. Understandably, the financial performance of projects was significantly affected by the crisis, and declined to a low point of only 44 per cent of projects approved from 2006 to 2008 attaining '*good*' or '*excellent*' financial performance. Similar to operational objectives, the financial performance of projects started to improve thereafter, with 61 per cent of projects approved from 2008 to 2010 rated as having '*good*'/'*excellent*' financial performance.

However, unlike the operational objective ratings, there have been quite pronounced variations in financial performance by sector and (to a slightly lesser extent) by region. Although, as previously mentioned, the financial institutions and industry, commerce and agribusiness sectors have the greatest exposure to market forces; the financial performance appears to have been particularly weak in the latter, with just 35 per cent achieving '*good*' or '*excellent*' financial performance, whereas for financial institutions projects the figure was 48 per cent. Also unlike overall performance by region, eastern Europe and the Caucasus show significantly weaker financial performance results than other challenging regions such as Central Asia, with just 37 per cent of projects approved from 2005 to 2010 showing '*good*' or better financial performance.

Financial performance in sectors more exposed to the crisis has been weaker.

Figure 6: Financial performance by sector and region, projects approved 2005-10



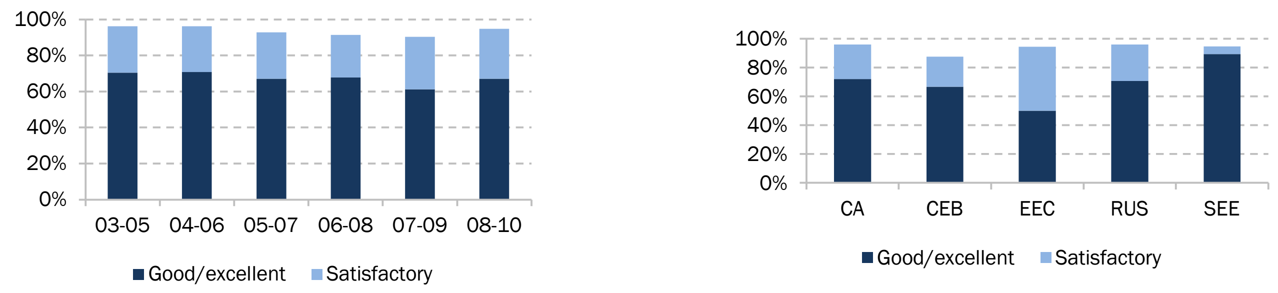
ENE: Energy; FIN: Financial Institutions; ICA: Industry, Commerce and Agriculture; INF: Infrastructure
CA: Central Asia; CEB: Central Europe and Baltics; EEC: Eastern Europe and Caucasus; RUS: Russia; SEE: South-Eastern Europe

Process and resource efficiency: Bank handling and investment performance

The process efficiency of Bank projects is measured through Bank handling, the quality of the Bank's own performance throughout the life of the project while the investment performance of the Bank is used as a measure of efficiency of resource use. This shows a similar pattern over time, but with higher scores across the board. In recent years over 70 per cent of projects have been rated '*good*' for bank handling, and over 90 per cent '*satisfactory*' or better. Evaluations often find the causes of less successful projects to lie outside the Bank's control in the political and economic environment, or in the commitment or performance of partners. That said, as previously noted, sometimes overly-optimistic expectations are a negative in terms of bank handling.

One notable pattern is that over the last five years where data is available, projects in eastern Europe and Caucasus have had significantly lower Bank handling ratings than other regions. From this it could be postulated that the Bank has assigned insufficient staff to manage projects in challenging contexts.

Figure 7: Bank handling by approval year (three-year rolling sample) and by region 2005-10

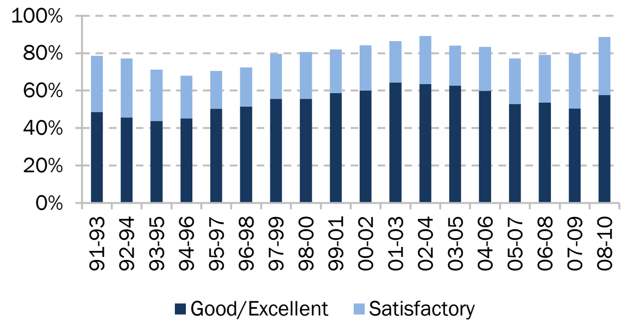


CA: Central Asia; CEB: Central Europe and Baltics; EEC: Eastern Europe and Caucasus; RUS: Russia; SEE: South-Eastern Europe

Project impact and sustainability: transition impact and environmental and social impact

The Bank sets specific transition objectives and benchmarks for all projects at appraisal, and assesses them for transition potential and risk. Figure 8 shows the proportion of projects rated *'good'/'excellent'* and *'satisfactory'* by EvD for transition impact (for three-year periods by year of Board approval). The pattern of performance is similar to that for overall performance (see Figure 1). This is not surprising since transition impact is one of the more important criteria used to derive the overall performance rating and as shown in the 2012 AER, has a high correlation with overall performance. The criteria benefits from a quasi-quantification but it is also EvD's opinion that in many cases the benchmarks actually have limited value as indicators of performance and an excessive number of these are simply recorded as 'on track' or 'not yet due'. EvD also notes that in an unfortunately frequent number of cases an associated TC intended to deliver the transition impact is dropped through lack of client interest or key covenants being waived. EvD recognises that these results occurred prior to improvements made to the transition impact monitoring system and the selection of more meaningful benchmarks. Nonetheless, based on the evidence of recently approved projects, EvD continues to believe further improvement is needed.

Figure 8: Transition impact by approval year (three-year rolling sample)



Projects are also given a transition risk rating to recognise the possibility that they may fail to achieve their transition potential; EvD has conducted analysis of transition impact performance against the transition risk rating, which can be found in the subsection 'Impact monitoring of investment operations' (p. 32) in '

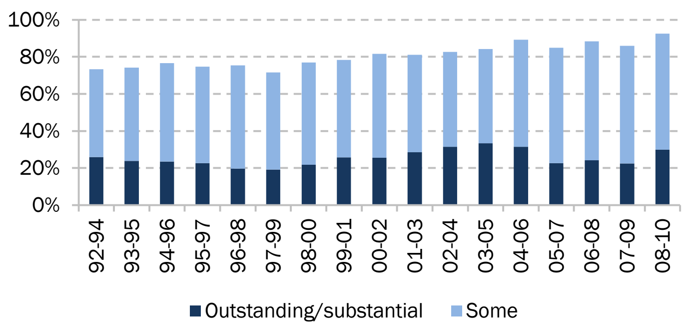
Review of self-evaluation and monitoring'. The findings show that the higher the transition risk rating, the more likely it was that transition impact performance was revised at evaluation, which is reasonable.

Environmental and social impact ratings have markedly improved in the most recent period

The Bank's mandate requires the Bank 'to promote in the full range of its activities environmentally sound and sustainable development' (Basic documents of the EBRD, Article 2 vii). However, it receives a very low weighting in the assessment of overall performance. This is partly because environmental and social impact is a significant factor only in a minority of projects. That said, there has been a marked improvement in the most recent period. Among evaluated projects approved from 2008 to 2010, 69 per cent were rated *'good'* or better for environmental and social performance and a further 25 per cent were rated *'satisfactory'*. Ninety-three per cent also achieved at least *'some'* environmental and social change. This is in contrast to projects approved from 2005 to 2007, of which only 46 per cent were rated as having *'good'* or better for environmental performance. However, given the limitations on the robustness of these ratings, some caution is required in interpreting these results.

Figure 10: Environmental and social performance by approval year (three-year rolling sample)

Figure 9: Environmental and social change by approval year (three-year rolling sample)



Aggregate performance in the post-crisis context

There has been a recurring pattern along several of the OECD DAC evaluation criteria dimensions highlighted above, of deterioration in the results of projects approved immediately before the crisis (from 2005 to 2007) and recovery in performance of projects approved from 2008 to 2010. This was especially the case with ratings for overall performance, additionality, achievement of project objectives, and the financial, environmental and social performance. However, it should be noted that performance in many of these dimensions is assessed against expectations at appraisal for each project, rather than against an objective measure of absolute performance such as a benchmark rate of return. This is also the case for ratings for the achievement of operational objectives, transition impact and environmental and social performance. Along a number of these areas, aggregate performance witnessed a downward trend and low-point for projects approved from 2006 to 2008, (immediately before and during the height of the crisis) and a subsequent improvement in the projects approved from 2008 to 2010. Improved results may indicate either actual improved performance or more realistic appraisal estimates in the context of the difficult economic environment in which the projects were being approved. Either of these scenarios would be welcome, however further investigation would be required to separate the channels of effect. The next section outlines some of the steps proposed to improve the robustness and explanatory power of the performance rating system.

Performance rating methodology

Current basis for project performance ratings

Project ratings continue to be derived on the basis of guidelines previously included in the Evaluation Policy until a new guidance note on performance rating methodology is adopted. The current Evaluation Policy, approved by the Board in January 2013, determined that technical guidance notes would be produced separately. A new guidance note on performance rating is under preparation.

EvD project performance guidance has been derived from the good practice standards of the Evaluation Cooperation Group (ECG) of the international financial institutions. Set up in 1996 to promote harmonisation of evaluation standards and methodology among members, the ECG developed and has successively updated separate good practice standards for the evaluation of private and public sector operations – the private sector good practice standards are now in their 4th edition. The private sector standards were principally developed by EvD and the evaluation department of the International Finance Corporation, the two founder ECG members solely or mainly carrying out private sector evaluations.

Although the EBRD has a public sector portfolio, EvD has never used the ECG good practice standards for the evaluation of public sector operations. Instead, it has applied the same performance rating methodology to both private and public operations. Although EvD's existing performance rating methodology is based on ECG good practice standards for the evaluation of private sector operations, EvD has always opted out of one of the four ECG criteria (project/company economic sustainability).

EvD rates the performance of TC but neither EvD nor ECG has guidance or good practice standards for the evaluation of TC. EvD has tended to apply OEC/DAC standards to the evaluation of TC.

Issues with the current performance rating methodology

Knowing that it needed to prepare a guidance note on performance rating methodology given that this was no longer included in the evaluation policy, EvD undertook an assessment of the current system to identify issues that the new guidance should address in order to strengthen the rating methodology. This analysis was contained in an approach paper prepared for the new guidance note. Some issues had been flagged in previous AERs where the authors struggled to explain trends in overall rating through analysis of criteria ratings. Two of the more important issues include the following:

- A lack of clarity, transparency and consistency in how criteria and even more so, sub-criteria ratings are derived
- A similar lack of clarity, transparency and consistency in how criteria ratings are used to derive an overall rating (the weightings applied)

Meanwhile at the level of ECG a rethink was occurring in terms of the merits of maintaining separate standards for the evaluation of public and private sector evaluations given, among other things, a considerable blurring between the two types of operations and the fact that several members only have one set of guidelines for all evaluations. ECG has launched an initiative to investigate the merging of the two sets of good practice standards.

Preparation of a new guidance note on performance rating

Throughout 2013 EvD worked on developing a new guidance note to cover the performance rating of loan and equity transactions and TC. The work was carried out by a senior EvD staff member and a consultant, with expert input provided by three external peer reviewers – the author of the 4th edition of the ECG private sector good practice standards and consultant to the ECG on harmonising its public and private sector good practice standards, the second consultant assisting ECG with harmonisation of its public and private good practice standards, and an international authority on evaluation.

The draft guidance notes proposes a number of changes to the rating methodology, aiming to achieve

- Closer alignment with OEC/DAC evaluation criteria (relevance, effectiveness, efficiency, sustainability and impact) with some customisation to the EBRD's unique mandate
- A revised set of rating criteria and sub-criteria, with clearer and more consistent guidance on how each should be assessed
- More explicit attention given to unintended results, both positive and negative results, and use of the counterfactual

- Adoption of a numeric scoring and averaging system to derive criterion ratings from sub-criteria ratings, and overall performance ratings from criteria ratings
- Incorporation of guidance for the performance rating of TC.

Annex 3 contains further details of the proposal.

Next steps

EvD aims to have revised guidance in place for the beginning of 2015. The implications of following the revised guidance are currently being explored prior to further consultation, finalisation, approval and adoption.

First, one of the external peer reviewers has been retained to retrofit the new guidelines to four recently completed operations evaluations and to develop a new self-evaluation template consistent with the new guidelines. The retrofitting exercise aims to:

- identify any issues with applying the new guidelines, such as a lack of clarity in the guidelines, data unlikely to be available, duplication (double-counting), lack of applicability and need for customisation for a particular type of transaction;
- determine the implications that will arise from applying the new guidelines in terms of additional data and analysis that may be required; and
- determine the extent to which application of the new guidelines might alter the ratings of projects.

Second, the new guidelines will be applied to two operations evaluations in the 2014 work programme to again identify issues with their application and to demonstrate their utility.

EvD will discuss the draft guidance note with Management, and application of the new guidelines to self-evaluation and a new OPA template designed to incorporate them. Once it has at least an initial reaction from Management, EvD will circulate the proposal to the Board for discussion in the Audit Committee prior to its finalisation before the end of 2014.

Findings from 2013 evaluations

In this Chapter:

- a discussion of cross cutting issues emerging, particularly results measurement, planning and implementation of technical cooperation and crisis response operations
- insights from in depth studies in the areas of equity products, local currency lending and telecommunications
- key factors and lessons from 2013 OPA Validations

Cross-cutting issues

Results frameworks and evaluability

The most prevalent 'results' finding to emerge from a roundup of OPA Validations completed in 2013 was lack of evaluability. There was a frequent inadequacy of verifiable indicators to enable thorough monitoring and evidence-driven conclusions. This issue was discussed in depth in a 2013 EvD discussion brief Evaluability and the EBRD.

Lack of evaluability was a reoccurring theme in Operation Evaluations and Special Studies. An evaluation of the Mongolia Cooperation Fund recommended the use of a results framework for donor funds, similar to the one recently introduced for individual TC operations. This would help the EBRD, donors, clients and other stakeholders to clarify the areas of intervention and the instruments available and act as a monitoring tool. Development of a results framework was also recommended in the study on Local Currency Lending and Borrowing.

> Greater clarity on indicators or targets coupled with improved data collection and monitoring would substantially strengthen the ability to assess performance and extract findings and insights to assist the client and improve future operations.

An evaluation of the Mongolia Financial Sector Framework raised a recurrent issue: how to link results expected from individual transactions to developments seen at a higher sector or country level. It suggested that a more explicit description of the baseline situation and a clearer definition of the objectives would help. In the absence of a neat and coherent definition of objectives and success indicators, valuable 'transition capital' gets lost or is invisible. The evaluation recommended developing the next framework as an integrated approach; making more efforts to describe the ex-ante situation with baseline indicators; setting more suitable transition benchmarks and other metrics to measure progress at framework and sector level; and monitoring the results of all instruments including TC and policy dialogue more systematically.

A 2013 EvD study of 'Performance Metrics - how well do EBRD projects' specify expected results' looked more deeply into this topic with a review of the evaluability of around 60 randomly selected project documents approved by the Board between July 2011 and June 2012. Applying an EvD-developed 'evaluability checklist' consistently to the documents in the sample, the study found that operations scored reasonably well for the definition of operational objectives and their results (mainly outputs), and for transition impact results and their benchmarks. This result seemed counter-intuitive given that EvD reports often comment on the difficulty in establishing the exact operational objectives and benchmarks. As a result, EvD conducted a further review of the evaluability of recently approved projects, the results of which are presented in the chapter of this report 'Evaluability of operations' (p. 24) and Annex 4: Further detail on evaluability assessments of TC and investment operations (p. 54). The Performance Metrics study found that the documents scored less well for the quality of indicators and benchmarks (target values, timelines and baselines), which are necessary conditions to measure performance. EvD recommended that Management should decide on the utility of the checklist and what role it might play in the project appraisal process at the EBRD; and that EvD should develop practical guidance and tools for operational staff as needed (such as templates and good practice examples for a coherent hierarchy of results for individual projects).

More broadly, the study made recommendations to improve performance metrics embedded in projects and presented in project documents:

- Ensure adequate baseline data are presented for all key results and related indicators/TIMS benchmarks;
- Present the key objectives and indicators comprehensively and concisely in an overall results framework in one place of the Board document;
- Strengthen the consistency and coherence of operational results vis-à-vis transition impact;
- Develop an EBRD-specific 'results framework' to improve reporting of institutional achievements through a more structured description of operational results and their links to Bank strategies, policies and corporate scorecard;

In response, Management noted that a number of important reviews on issues related to performance metrics were currently underway and that EvD's analysis and recommendations were very relevant to these initiatives. The basic template for Board project approval documents has been changed, introducing a new section on 'Measuring / Monitoring Success'; TC approval documentation now includes a results framework for all TC, both transactional and non-transactional; and the transition impact methodology has been amended. The Chapter Evaluability of operations (p.24) of this report includes a review of evaluability of TC operations incorporating the new results framework. EvD will conduct a similar review of investment operations in the 2014 Annual Evaluation Review once a sample of projects approved under the new approach is available.

Technical cooperation

Issues covered by Validations within the area of TC included relevance to project objectives, treatment of implementation, adequacy of metrics and reporting, clarity of linkage to the main project and overall effectiveness. In some cases, TC did not meet expectations of reducing risk, strengthening the client or improving sector reform.

In 2013, EvD completed a study TC experience in new countries of operations – the Mongolia Cooperation Fund, in part to assess whether it might provide useful findings for Bank activities in other new countries of operations. The Mongolia Cooperation Fund (MCF) was a multi-donor fund established in 2001 to finance technical cooperation activities in Mongolia prior to its eligibility for regular EBRD operations. Besides recommending the use of a results framework for individual operations and donor funds, the study recommended embedding policy dialogue activities in the results framework of the donor fund and, where possible, in individual TC operations. The fact that such activities cannot be monetised in a fund's budget (if carried out by staff rather than consultants) does not mean that it is not important to report on them. Planning and monitoring policy dialogue activities in a new country of operation is essential to successfully initiating new investment opportunities.

TC funding can benefit from the use of results frameworks including elements on policy dialogue activities

The evaluation also made some recommendations on the efficient management of funds. It noted that major resources were used to support a few public sector clients with whom no investments took place, and that the fund remained open for several years with a very small balance outstanding. It recommended improved monitoring to limit the use of resources once it is clarified that the outcome cannot be achieved, and to review funds regularly upon reaching specific trigger points for utilisation and/or duration. Management commented that the MCF had accomplished all its major activities and there was only one sub-window left open at a donor's request to finance further legal transition related projects.

The evaluation of a packaging project (Summary: An agribusiness company), highlighted the importance of integrating transactional TC with the investment project, particularly where the TC is expected to deliver an important part of the transition impact associated with the project. It also raised the issue of ensuring follow-up on TCs with negative conclusions. In this case the TC was a feasibility study for a recycling scheme, and it identified various obstacles in the legal framework, which did not enable and incentivise such recycling and furthermore blocked exports of scrap metal to countries with suitable recycling facilities. However, the TC was not followed up by any action to try to tackle the obstacles. If the conclusions of a feasibility verification TC are negative, while the potential transition impact of the activity is high, the Bank should explicitly identify how the obstacles will be addressed, including with a follow-on TC.

Crisis response operations

Since EvD's initial 2010 process study The EBRD's response to the 2008/2009 crisis , an increasing number of individual projects and facilities developed in response to the crisis have become ready for final evaluation, two of which was finalised in 2013: the Operation Evaluation: Mid-Sized Corporate Support Facility and a bank. These cast light on the actual outcomes of projects developed in a crisis situation.

The evaluation of a bank operation identified challenges to changing a bank's business culture. It found that insolvent banks offer limited scope for achieving some transition impacts such as changing the business culture or entering new lines of business. Resolving an insolvent bank cannot easily carry the burden of other, standard transition impact objectives. The results matrix for such projects must be calibrated accordingly. A similar issue has been identified in other crisis response projects, such as the 2012 validation of a cosmetic manufacturing company, where the rationale as presented in the Board document and OPA struggled to justify the project in terms of the regular country and sector strategies when a closer focus on the crisis response aspects would have provided a better justification.

Other findings from the evaluation of the bank related to the requirement to disclose the nature and amount of the EBRD's investment, and the importance of transparency to transition impact. Furthermore, the evaluation noted that when a project seeks to improve the integrity performance of a firm or activity, the board report should candidly describe the known level of integrity performance, both to provide a clear baseline for future assessment and to allow a realistic estimation of transition risks and challenges.

While appreciating the substantive exchanges between EvD and operational staff in relation to the report, Management felt that the *'partly successful'* rating did not internalise the crisis-related context reflected in the report, nor reflect the overall impact of the operation. Management felt that the operation and the EBRD's role in it had been central to the country's success in stabilising its financial sector and economy to the extent that it was able to subsequently join the. Management disagreed with the main findings of the report, feeling that EvD failed to take account of the macroeconomic situation and the need to adapt the project to accommodate EU state aids rules.

The Mid-Sized Corporate Support Facility was intended to provide short and medium term loan financing to existing, fundamentally sound EBRD clients in the corporate sector to support them through the crisis and protect transition achievements of the original EBRD investment. EvD's evaluation found that the Bank can play a positive role in a crisis, providing additional financing to its corporate clients; however its eligibility criteria need to be flexible, while expectations related to transition impact and environmental performance should be limited. Although it played a generally positive role, the Facility helped preserve and progress the transition and environmental achievements of only the best companies, which were less likely to require the support in the first place.

The evaluation found:

– the Bank's usual focus on capital investment financing in the corporate sector made it difficult to act as an ad hoc working capital provider, even in a crisis. Clients commented that the need for very speedy approval, the limited amount required and the revolving nature of working capital financing puts it more naturally within the remit of local banks (even if their costs are slightly higher).
– the introduction of new facilities with a new type of financing for the Bank requires a careful approach, with demand for such financing tested with a smaller budget envelope first.
– companies which are financially sound are likely to continue receiving working capital financing from local banks, even during a crisis. Given the Bank's high standards for clients many of them belong to the better-managed corporate elite, and can often continue attracting short-term financing from their house banks, even during a crisis.

New types of financing should be tested on a small scale first to address issues early

The study recommended that some facilities could be consolidated and their eligibility criteria standardised. The introduction of more fundamental changes, such as the universal delegation of approval for all smaller projects (below €20 million) to the Operations Committee or the Small Business Investment Committee would result in substantial decrease in the number of facilities. The issues were acknowledged by the Bank's Finance and Operations Policy Committee, and better impact assessment in respect of financing facilities (also involving EvD) was advised, while decisions on facilities consolidation and minimum thresholds for Board approvals of all projects was deferred. Management did not present formal comments on the evaluation.

Findings and recommendations from special studies

Achieving equity investment objectives

An EvD study reviewed the implementation of various initiatives launched for equity investments since 2007 (Achieving Equity Investment Objectives: a review of initiatives since 2007). Its key findings were as follows:

– The equity approach now employed lacks a clear emphasis on value creation from pre investment to exit and is undermined by the monitoring report structure; whether the new approach to Value Creation Plans introduced in 2012 will correct this is to be seen.
– Monitoring reports no longer fit the purpose of equity monitoring; the current approach is not generating effective monitoring in a coherent and structured way.
– The heavy emphasis on governance underpinning the additionality for EBRD equity investment is not reflected in the approach to due diligence or documentation of issues and opportunities in the final review memorandum, Board Document or monitoring reports.
– There are questions around the effectiveness of engagement with nominee directors, related to a transparent and documented process of setting and monitoring nominee director objectives and performance.

To tackle these issues, the study recommended that Management:

– Review business processes for equity investment to enhance the focus on results. Placing value creation and transition impact at the heart of equity investment should be central;
– Review post investment monitoring and reporting to better define their intended objectives across multiple stakeholders. The review might conclude that the monitoring report needed to be refocused around a smaller number of equity results, value creation drivers and key performance indicators, rather than to try and satisfy multiple users in multiple ways as at present; and
– Modernise Board documents for equity content, incorporating equity-specific elements into a new format. As a minimum, the equity content should reflect current initiatives in Banking to introduce value creation plans for all equity investments and enhanced corporate governance approaches.

Management found the study highly relevant to the ongoing discussions of the Bank's equity approach and largely concurred with the study's principal observations. A broader review of the Bank's credit monitoring processes, currently underway, would encompass equity monitoring. The revisions to the Board approval document, ongoing at the time of study completion and since finalised, incorporated several enhancements relating to equity operations. The Audit Committee asked EvD to collaborate with Management to produce a progress report in the first half of 2014.

A study on local currency lending and borrowing 2000 to 2010

EvD's evaluation of the EBRD's activities in local currency financing over the 10-year period to 2010 highlighted the limited overall extent of local currency lending over the period and its high degree of concentration in only two countries. These limited outcomes stand in contrast to the Bank's strategic- and country level emphasis on the imperative of developing local capital markets, and the clearly identified fragilities and systemic risks seen widely to have arisen from excessive foreign exchange exposure. The EvD study also raised in this context a number of questions about the balancing of risks associated with alternative choices of instruments. The study made four recommendations:

➢ Local currency lending and borrowing operations would benefit from a clearer market development strategy and results framework; more rigorous, periodic cross-departmental peer review; and evaluable metrics and monitoring;

➢ Operational proposals and risk policies for the fifth credit resource review might include analysis of the trade-off between reducing credit risk through local currency loans and adding market risk - foreign exchange (FX) and interest rate risks - to be able to provide them.

- Simpler credit guidelines could discourage the EBRD from FX lending to unhedged borrowers, which would support the credit quality, transition, and volume objectives of the Bank's local currency lending and borrowing operations. Updated policies could discuss how the Bank would apply the European Systemic Risk Board (ESRB) FX recommendation for its own lending operations. As a first step, the EBRD could host a presentation to the Bank and the Board by the ESRB about its FX recommendation, during the first quarter of 2014.
- The Office of the Chief Economist (OCE) should consider using the next EBRD Banking Environment and Performance Survey (BEPS) to gather detailed information about local currency lending practices, the term structures of interest rates in local currency loan markets, anti-money laundering practices regarding interest rate risk, and other knowledge needed for the EBRD's local currency financial system development work.

Management comments on the study queried the timing of the report, which omitted substantial developments after 2010, and what it saw as the narrowness of the approach, which it felt ignored the economic and political environment of the time and also wider Bank activities in the area beyond local currency lending and borrowing. Management disagreed with the recommendation for the Bank to assume greater FX and/or interest rate risk, believing that this would not address more fundamental problems relating to the wider availability and pricing of local currency lending. They disagreed that a lack of fixed rate local currency lending was a fundamental problem and believed that the Bank already applied ESRB recommendations as far as they were applicable to EBRD activities.

Telecommunications sector study

EvD's 2013 Telecommunications sector review presents a number of findings derived from individual case studies, the Bank's overall performance, and some recommendations specific to the development of a new Information & Communication Technologies (ICT) policy. In addition, it presented some recommendations relevant to project design and execution in the ICT sector:

- When setting transition benchmarks for telecoms projects, the Bank should take account of shifts in demand experienced in the telecoms markets and in consumer preferences elsewhere; in particular, it should note the possibility of mobile telephony substituting for fixed telephony and the expansion of mobile broadband.
- During due diligence, in addition to analysing telecom-specific regulations, due attention should be paid to general laws (particularly anti-monopoly and competition), which can have great effect on telecom sector projects.
- When supporting changes in the regulatory environment, take account of the side effects of regulation experienced in comparable circumstances elsewhere; in particular, consider the relevance of the levels of termination rates and telephony penetration to number portability.
- When assessing the viability of an investment in a telecoms operator, distinguish between the number of active subscribers and the total number of subscribers. An operator can have an impressive subscriber base, but little or no cash flow coming from it.
- When financing operations in highly competitive and risky markets, particularly with smaller telecoms operators, first consider providing quasi-equity or subordinated debt, rather than straight equity, to mitigate the risk of financial underperformance.

Management provided no comments on the study.

Key findings from OPA Validations

Besides the findings already mentioned in the section on Cross-cutting issues (p.16) above, OPA validations also produced findings and lessons on the following issues:

Transition impact

Transition impact metrics and targets sometimes lacked strong enough causal linkages with specific project design elements or loan conditions. In some cases achievement of the projected impacts proved highly vulnerable to exogenous factors outside of Bank or client control, such as the economic and regulatory environment; in others, assumptions about client commitment or consistency of direction proved overly positive. Trade-offs between greater transition impact and acceptance of higher risk were cited; and, missed opportunities were seen for the EBRD to have had wider influence by adding to project scope, such as through training programmes.

Additionality

Issues around additionality emerged with some consistency. Points covered adequacy of the argumentation and evidence, as well as the adequacy of the treatment of counterfactual and contributing factors. Additionality is very difficult to discern in retroactive financing. Not surprisingly, additionality was found to be stronger, and perhaps better argued, across the crisis period.

Policy dialogue

Validations made a number of points in the area of policy dialogue, including with respect to the adequacy of instruments used, realism of objectives, and level of client commitment. Negative findings emerged particularly from projects in Russia, Central Asia and Early Transition Countries, though this may be a product of the greater concentration on policy dialogue in these countries. Legislation was a stumbling block for a financial sector intervention in one case; an attempt to privatise the largest pension fund manager was not achieved in another, and all pension assets were nationalised after the EBRD's exit. Numerous projects did not evince the expected benefits from policy dialogue.

In contrast, policy dialogue associated with a municipal project led to tariff reform in one case, while 'commendable results' were achieved in policy dialogue and implementation of environmental and social standards through an energy efficiency facility in another.

Project structuring

Validations produced a varied range of insights related to project structuring issues and related legal issues. One project encountered problems with land ownership. A failed attempt at privatisation led to the lesson that the Bank's subscription in pre-privatisation equity investments should be conditional on issuance of a privatisation decree, or else the project document should explain why this was not needed. A validation of an ICT project found that a convertible loan can be a useful alternative to a straight equity investment when financing a local company with an unproven track record, particularly in a fast growing sector. It can enable the Bank to better monitor the client's performance over time and take a more informed decision on whether to acquire shares or have the loan repaid.

Risk mitigation

Validations regularly address a variety of risk-related issues across the areas of client, political and market risk, the adequacy of discussion, and mitigation measures employed. Private sector projects in industry and commerce found that the participation of experienced, international market leaders as sponsors does not guarantee the success of a project. In volatile sectors such as the information technology market, even well performing market leaders can be affected by negative sector-wide sentiment. Inherent project and sector risks in volatile and cyclical sectors need to be considered when developing and evaluating expectations, warranting additional caution and additional contingency or safety buffers in the forecasts. Projects involving the introduction of new products, technologies and process should always be carefully assessed on their merits and technical risks properly mitigated. In high tech industries, design and testing periods are integral to product development and can be lengthy.

Board and Audit Committee review of EvD reports

EvD delivered four items to the full Board in 2013:

– EBRD Evaluation Policy
– EvD Work Programme 2013
– 2012 Annual Evaluation Review
– Evaluation Recommendations and Management Follow-Up Report for 2012
– EvD Work Programme 2014-2015 and Budget 2014-15

The main discussion of the new Evaluation Policy had taken place in several sessions of the Audit Committee throughout 2012. At its consideration by the full Board in January 2013, Directors noted that the new Policy established an important milestone in the modernisation of evaluation in the Bank. The Chief Evaluator noted that EvD intended to move to a multi-year work programme to give the Bank a more strategically coherent set of activities to be accomplished over the period.

During the discussion of the 2012 Annual Evaluation Review in September 2013, the President observed that progress in the self-evaluation exercise was uneven across the Bank, noting that more work should be done to make it more consistent over the coming years. He added that he was committed to further improving cooperation between Management and EvD. The importance of stepping up cooperation between Management and EvD was also stressed by some Directors. When the Board approved the EvD Work Programme 2014-2015 and Budget 2014-15 in December 2013, the President, Audit Committee Chair and Chief Evaluator all expressed appreciation of the changes in terms of interaction between EvD and Management.

EvD presented 12 items to nine meetings of the Audit Committee of the Board of Directors in 2013, including the following reports:

– Evaluation Brief: Evaluability - is it relevant for EBRD?
– Synthesis: Findings and insights from TC evaluations
– Corporate Recovery projects: insights from impaired assets
– Special study: EBRD-Italy Western Balkans Local Enterprise Facility (LEF)
– Special study: Early Transition Countries Initiative (ETCI)
– Achieving equity investment objectives - a review of initiatives since 2007
– Operation evaluation: Bank
– Evaluation Department work programme for 2013
– 2012 Annual Evaluation Review
– 2012 Evaluation recommendations and Management follow-up
– Approach paper: Developing an improved system for Management follow-up on EvD recommendations
– Evaluation Department work programme 2014-15 and budget 2014

Selection and timing of evaluation reports

The Audit Committee particularly welcomed reviews of large Bank programmes such as the Special study: EBRD-Italy Western Balkans Local Enterprise Facility (LEF), the Early Transition Countries Initiative and the Synthesis: Findings and insights from TC evaluations and encouraged EvD to produce more reports like these. In some cases, Management found evaluations well timed to feed into the preparation of new strategies or the renewal of existing frameworks and programmes. However, Directors expressed concern that their discussion at Audit Committee was sometimes delayed beyond the point at which it would have been most useful for Directors. In the discussion of the 2013 work programme, the Audit Committee asked EvD to provide a timeline giving expected delivery dates for major studies to ensure that Directors were able to consider EvD's findings at the most appropriate time. For 2014, it was agreed that EvD would improve the flow of information to the Committee and at least one Audit session per quarter would be focused on evaluation work.

Technical cooperation

The Synthesis: Findings and insights from TC evaluations prompted an in-depth discussion of TC at the Bank, which took place in February 2013. Directors emphasised that TC funds should not be taken for granted. The opportunity cost for other activities in donor countries was high and national authorities were very keen to see results achieved through TC funds. Committee members noted that the Secretary of TC Com now attended Operations Committee meetings but only as an observer, which limited the influence of donor representation. It noted that the weight within the Bank and the skills of the specialist TC units needed to be enhanced to better protect donor funds. While recognising that there had been significant headway and reforms taking place since July 2012, the Committee stressed three major challenges lying ahead:

- the need for a much more integrated approach to TC operations, policy dialogue and projects;
- the need to ensure effective coordination between TC Com, the Operations Committee and the Strategy and Policy Committee;
- the need to narrow the gap between investment and TC components.

Small and medium-sized enterprises

The Committee's review of the studies on the ETCI, EBRD-Italy Western Balkans Local Enterprise Facility (LEF) and Corporate Recovery projects all led to discussions of how to handle small and medium-sized enterprises (SMEs). The Committee expressed strong support for the Bank's activities with SMEs - described as the 'backbone of a market economy' - while recognising the challenges they posed.

The Corporate Recovery study noted the relatively high proportion of smaller projects among those entering recovery, while bearing in mind these still constitute only a very small proportion of the Bank's total portfolio. It questioned whether some projects were too large relative to the size of the SME, resulting from the Bank's difficulty in down-scaling the size of its investment projects to fit the financial capacity of smaller firms. Management noted improvements in due diligence introduced by the Small Business Investment Committee but recognised remaining challenges, especially regarding business practices, financial reporting and accounting standards. The discussions of the EBRD-Italy Western Balkans Local Enterprise Facility (LEF) and the Corporate Recovery studies considered the advantages and disadvantages of financing through local banks, which could provide additional services and were well placed to monitor but in some areas - including the western Balkans - were capital constrained and not expanding SME financing.

Evaluability of operations

In this Chapter:

- the rationale for review of the evaluability of Bank investment and TC project proposals
- an assessment of 38% of TC plan results matrices approved by TC Com in the 2H 2013
- a review of a random sample of 46 investment operation plans approved by the Board in 2013

Rationale for review

'Evaluability' is understood as the extent to which the value generated or the expected results of a project are verifiable in a reliable and credible manner. This was elaborated in EvD's 2012 brief Evaluability – is it relevant for the EBRD? And a subsequent 2013 in depth study Performance Metrics: how well do EBRD projects specify expected results?.

The EBRD's increased focus on evaluability is reflected within the EBRD's internal 'One Bank' programme which emphasises measurement of the Bank's impact, the Grant Co-financing Strategic Review in 2012 and the introduction in July 2013 of a mandatory results framework matrix for all new TC project proposals. EvD's Synopsis of OPA Validations completed in 2013 found concerns over OPA quality which often stem from poor evaluablity. It is therefore timely at the end of 2013 for EvD to review the Bank's progress to date in this area.

The review covers a random selection of investment operation plans approved by the Board and technical cooperation results matrices approved by TC Com in 2013 to understand their strengths and weaknesses. The next two sections summarise the results of each evaluability check, with additional information, including checklists, in Annex 4: Further detail on evaluability assessments of TC and investment operations (p. 54).

Preliminary quality assessment of TC submissions under the new results framework

Together with the Bank's TC Team, EvD developed a TC submission evaluability checklist (Table 8: The evaluability checklists (transaction and non-transactional) for TC operations) (p. 55) that was used as part of the results framework rollout, and to support TC Com's broader quality at entry assessment of TC Com submissions. The assessment covered a random sample of 51 results matrices from TC Com submissions, covering 38 per cent of the population submitted in the second half of 2013.

The sample scored very well against the evaluability checklist. A matrix meeting all the requirements of the checklist would score 100 per cent. The average score was 79 per cent for non-transactional TCs, and 82 per cent for transactional TCs. Despite the overall high scores, there were also areas of continued weakness and variance across different sections.

For non-transactional TCs, the two sections within the results matrix which scored least were:

- 'TC outcomes contribute to transition impact,' which in almost a quarter of all cases, was assessed as 'No';
- 'Transition impact indicators are specified,' where over 70 per cent the sample specified either 'few' or 'none'. [1]

This outcome is noteworthy because transition impact in these cases cannot rely on transition impact from a linked investment operation. The TC must therefore set out its own linkages between actions and anticipated results, and the indicators needed to assess performance. The finding suggests there is considerable scope for improvement.

For transactional TC, the sections scoring least were:

- *'All outputs necessary to achieve the outcome(s) have been stated'*, where 35 per cent failed to score any points;
- 'Transition indicators are specified,' where a third of the sample scored 'few' or 'none'.

This is of concern, given that operations typically make transactional TCs integral and explicit to the operation and therefore the design of the TC must be fit for purpose. Further, in a few transactional TC cases, the stated outcome was rated as not possible to be achieved by the TC, as submitted in the results matrix. On a deeper look, it is often the expression of the TC design rather than the failing of the design itself, but this must be captured accurately to fully optimise the results matrix.

Evaluability assessment of 2013 Board approved investment projects

In supporting the progressive improvement of evaluability of EBRD projects, EvD has developed a checklist to assess the evaluability of investment operations using the Board approved operations report as a base. The checklist treats the evaluability of investment operations along three dimensions:

- expression of a results framework;
- expression of a monitoring framework;

[1] The possible ratings for this section in the Evaluability Checklist were: 'All', 'Mostly', 'Few' or 'None'. See Annex 4.

– treatment of risks.

These dimensions are stratified into:

– a 'basic test' which includes current requirements and what EvD considers the most important elements of good practice that are broadly in line with current guidance even if not fully specified by it; and
– an 'advanced test' which is in line with accepted best practice.

OPA validations and evaluations have in recent years included comments that project design at appraisal lacked baseline information and monitoring indicators, resulting in poor monitorability and low evaluability. This is the second evaluability check that EvD has undertaken,[2] to try to more fully understand these identified areas of weakness. The assessment reviewed a stratified random sample of 46 operations reports approved by the Board from January 1 to December 31 2013. It aimed to provide immediate feedback on the feasibility of applying the checklist, which had been revised since the previous assessment, and to provide an initial baseline of experience for future assessments (depending on the robustness of the revised checklist).

The exercise indicated that EBRD operations have just a fair minimum level of evaluability, hovering around an average of 73 per cent on the basic test of evaluability. However, it also revealed a number of ongoing flaws in the checklist which cast doubt on the validity and applicability of the findings:

– The EvD 2013 study on 'Performance Metrics: how well do EBRD projects specify expected results?' observed that key information on results (other than transition) is often not presented clearly and consistently within the Board approval document but is scattered throughout it. The current version of the checklist does not assess this issue.
– The checklist does not adequately assess how well the approval document describes the contribution of the operation to higher level impacts, including those identified in country and sector strategies and other Bank policies.
– The qualitative assessment of transition (and other) benchmarks does not penalise weaknesses so long as a bare majority of indicators are valid; this is a weak test of the adequacy. Nor does it penalise overuse of indicators, though it is clear that a project which lists seven indicators for one transition impact benchmark is not best practice.
– The Performance Metrics study noted some concerns over the 'risk' section of the checklist which led to inconsistent rating between the two consultants conducting the assessment. Consequently, the assessment of risk has moved to a separate part of the checklist but the questions over its inclusion have not been fully resolved.
– The checklist does not reflect the diversity of operations approved at Board level, for example framework operations versus standalone operations.
– Because of efforts to keep the checklist short, it does not sufficiently tease out important dimensions of evaluability, such as separating individual elements in the basic test of indicators.

Given the deficiencies of the checklist, EvD has chosen not to present the detailed findings. Instead, in 2014 it will continue to work on refining the checklist to address the issues identified above. The target will be to develop a revised checklist by the end of 2014, allowing a similar evaluability exercise to be conducted for the 2014 Annual Evaluation Report. The checklist must consistently identify examples of good and bad practice in order to be useful as a tool to improve the evaluability of Bank projects presented for approval.

[2] Following on from the assessment conducted for the 2012 Special Study, 'Performance metrics - how well do EBRD project specify expected results?' (PE12-560S), published May 2013.

Review of self-evaluation and monitoring

In this chapter:

- A quality assessment of 2013 self-evaluation (OPAs)
- Illustrations from a sample of OPAs
- Characteristics of good quality OPAs
- Discussion of the ratings gap between self - an d independently evaluated operations
- Analysis of impact monitoring and evaluation of investment operations

Quality assessment of self-evaluation

EvD introduced a new self-evaluation template in 2012, the Operation Performance Assessment (OPA), intended among other things to increase the depth and rigour of the mandatory self-evaluations done for all completed projects. Following the introduction of this new approach, this year, EvD has developed a quality review checklist[3] which will enable a more systematic review of OPA quality to provide more consistent feedback to portfolio teams on the quality of their submissions.

EvD developed two variants of the OPA checklist for application to long-form and short-form OPAs.[4] The long form checklist includes questions along the following five dimensions of quality, each of equal weighting: (i) project relevance; (ii) project effectiveness; (iii) project efficiency; (iv) project impact and sustainability; (v) completeness and sufficiency. The short form OPA checklist was developed to take into account the differences in the short-form OPA, notably that it contains no assessment of project rationale.

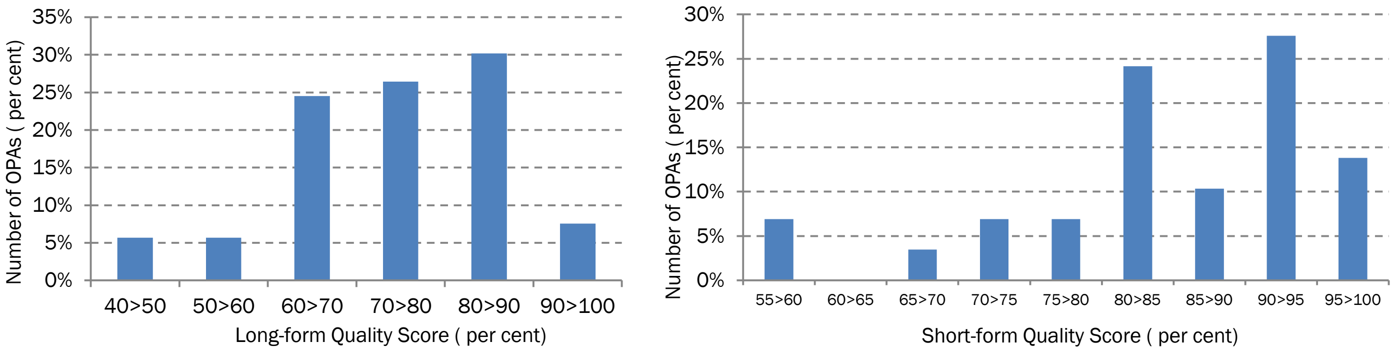
EvD has applied this quality review framework to the population of OPAs signed off and submitted to EvD in 2013. This preliminary review of 82 OPAs from 2013 will provide a reference point for annual review, and more immediately, helps identify areas of strength and weakness. Importantly, the checklist is also used to identify those OPAs with particular merit with a view to recognising the extra efforts of specific operations teams in carrying out self-evaluations, and set examples for future self-evaluations.

Overall the checklist proved an effective tool. There was an average of 10-12 per cent divergence between the conclusions of the two different evaluation team members with a cross reference of over 50 per cent the population, which can easily be reduced by a few clarifications in the checklist. Similarly, the OPAs with most merit were picked up by both the checklist and EvD managers' comments during validations and review.

Summary findings

Taken on average, the 82 OPAs submitted in 2013 were of good quality, scoring 78 per cent across the five dimensions. Encouragingly, given the novelty of the format, only one submission scored 0 per cent on any individual question. In terms of the difference in average overall scores, the short form was slightly higher than the long form score at 84 per cent.

Figure 11: Quality scores of long-form and short-form OPAs prepared in 2013



Having said this, there was much dispersion around the average, as the graphs above show. With respect to long form, the vast majority (81 per cent) scored between 60-90 per cent on the quality review. The lowest score was 42 per cent and the highest scoring OPA reached 95 per cent, with four OPAs (out of 53) scoring above 90 per cent overall. Because of the smaller number of questions on the

[3] Attached at Annex 5.
[4] Operation teams prepare long-form OPAs for operations that will be subject to independent validation or evaluation by EvD; they prepare short-form OPAs on all other operations.

quality review of the short form, the scores were more bunched; hence it was necessary to disaggregate the scores into vigintiles. Despite a more diffused distribution overall, 76 per cent of the short form OPAs scored at least 80 per cent for overall quality. For short forms, the lowest score was 55 per cent. Once more, four OPAs scored above 92 per cent and can be seen as particularly strong models of self-assessment.

Overall the OPA quality exercise showed that, since last year and with the introduction of the new reporting structure, the Bank has gone some way to achieving the objectives of greater ownership and better critical self-assessment. Not only are some teams showing greater ownership by incorporating the timing and satisfactory submission of OPAs within the performance appraisal objectives for OLs, but within the OPAs themselves there is a greater focus on dimensions beyond targeted financials. Some of the best quality OPA examples in 2013 are taking longer term views of the projects/programmes within a broader transitional context. In contrast to last year, although the financial performance section remains strong, OPAs also scored highly on project relevance and almost all used clear and consistent project objectives derived from the board documents, treating operational objectives separately from transition impact objectives. Long form OPAs were scored on project relevance and here too, were found to reach an excellent standard of reviewing the case for additionality, with 96 per cent percent presenting a good retrospective assessment. This allows for a more coherent storyline, and an OPA that tries to chronicle all major events in a project life, whether discussing reasons for dropped TC, or chronicling examples of successful environmental and social change, leads to better assessments.

Since the introduction of the new self-assessment the Bank has made substantial progress towards taking ownership of the process and engaging in sharper critique of operational achievements

Figure 12: Illustrations from example OPAs

- An excellent example in the Telecoms sector not only provided compelling evidence in support of self-evaluation judgments, but also insightful detail on how its specific structure helped to achieve its transition objectives.
- Another in Power and Energy succinctly tied together the project story into a tight synthesis within the summary section, whilst also providing additional information and specific examples of project performance in both the transition impact and environmental and social sections, which are traditionally areas of skeletal detail. This OPA also benefitted from a good communication between EvD and banking, which helped adjust the initial report structure so that all project elements were captured logically ad consistently with the Board documents.
- A further excellent example from MEI took a complicated and multifaceted project which had undergone significant changes to its design since board approval, presenting challenges to evaluation, and submitted a detailed and frank report nonetheless.

Within the OPA population there were several excellent examples; Figure 12: Illustrations from example OPAs and Figure 13: Characteristics of good quality OPAs and provide some examples of common characteristics of good quality OPA's.

Although the average quality scores were high and progress has been made, several substantial shortfalls remain in crucial areas where the EBRD must be able to assess its project performance. Across both long and short form OPAs, the sections with the lowest scores were project efficiency and, more worrying given the EBRD's mandate, project impact and sustainability, scoring 64 per cent and 69 per cent in this area respectively; importantly, long form OPAs fell short of 60 per cent on impact and sustainability taken alone. It should also be noted that the checklist follows current guidance, which is not highly demanding in some areas such as relevance and rationale. This can be made stronger in future iterations of the checklist.

Bank handling scores

Unsurprisingly, quality of the self- assessment of bank handling was the most challenging section since those doing the self-evaluation are reflecting on the performance of themselves and their team, among others. Self-assessment scores were often substantially diluted by poor assessments in the area. The bank handling section is intended to draw out lessons for improving bank efficiency; however, analysis often tended to concentrate instead on justifying why bank handling was good, irrespective of the operational performance. Further, self-assessments often lacked a retrospective assessment of project risk identification and mitigation measures, or a discussion of whether the project was designed fit for purpose. Only 49 per cent of long form OPAs captured more than the half the points available for bank handling. Finally, many self-assessments failed to explain the project origination context, which is important to understand the spirit in which the project was approved and undertaken.

Figure 13: Characteristics of good quality OPAs

All of the highest scoring self-assessments went further than the confines of the OPA template and exhibited similar characteristics, such as:

- discussed project origination;
- placed the project in the context of ongoing policy dialogue and client relationships within the Bank, referring to related or pipeline operations;
- described how the project was structured to address the most critical transition challenges in the country, region or sector;
- offered a balanced retrospective assessment of the pluses and minuses of decisions surrounding due diligence and design of the operation; especially useful for future programming in more variable sectors or unstable political/economic climates;
- provided detail on policy dialogue and technical cooperation which took place around the operation, whether by EBRD or others;
- included discussion on ‘efficiency’ and ‘sustainability’ underpinnings of the operational achievements enriched these OPAs.
- Provided insightful lessons that were measured, easy to understand and useful to their peers.

Project impact and sustainability scores

Project impact and sustainability scores were often reduced because teams failed to adequately convey the potential transition impact, including its risk, beyond cursory phrases. Assessment quality also suffered from a lack of any analysis beyond reproduction of the latest TIMS report, which in several cases was outdated and suffered from benchmarks that were not good indicators of performance. Also some benchmarks had already been achieved at approval while in other cases they were assessed as *'on track'* or *'not yet due'* with little or no assessment of progress towards their achievement. The effect was greatest amongst long form OPAs, averaging 60 per cent, with the lowest OPAs failing to score 20 per cent in terms of assessment quality here. In the worst cases, OPAs failed to score in this section, where several gave a superficial overview and failed to convey anything of value to an outside reviewer, nor justify their ratings as a result. EvD observes that on several occasions TIMS monitoring had stopped before completion of project implementation because once a benchmark has considered to have been achieved or it is judged no longer achievable it is no longer monitored.

Suggestions to improve self-evaluation

The OPA process is a critically important part of the project cycle. Done effectively, it can contribute directly to a more rigorous and results-focused operational perspective across the whole project cycle, including beyond its financial closing; and it can provide a highly valuable experience for new project design and execution. The OPA is designed therefore to enhance the EBRD’s accountability and learning at the project team level, to provide insights for future operations bank-wide and increasing the efficiency of evaluation. In this vein, the OPA quality checklist exercise resulted in three general observations where banking teams can concentrate their efforts to improve future assessment quality, Management can consider improving future evaluability of operations, or where EvD can take certain actions to support OPA quality:

- Consistency across Bank’s project documents and clear measurement standards for performance benchmark would improve ex ante evaluability for later assessment. At the stage of self-evaluation, ex ante evaluability weaknesses and overly-optimistic projections of expected results become problems. These turn on the:
    - definition and appropriateness of project objectives;
    - suitability of relevant indicators for the objectives; and,
    - identification of baseline and benchmarks are common themes.
    - ➢ As good practice, EvD would support the introduction of a results framework into the pre-signing documents, which collates in one place all outputs, outcomes, impacts and a measurement framework before board approval. This would also include the strategic or wider transitional context as well as those specific to the operation. The new template for Board approval documents takes some steps towards this, which may help to improve project evaluability if applied, something EvD will monitor through observations around future OPAs.
- Critical assessment underlies evidence-based ratings. Those OPAs which did not shy away from taking a stance either way on project handling and effectiveness gave the richest insights to operations and identified team strengths and challenges projects face in reality. Moreover, concrete examples serve best as evidence for justifying project ratings, and the lack thereof impeded OPA quality. Internal recognition and incentives for such an approach could be strengthened and applied more widely.
- Imprecise guidelines and flexible templates lead to inconsistent interpretation allowing variable depth of assessment and therefore quality. An example of where EvD can support OPA quality is through enhancing guidelines, particularly around project efficiency, which was the area that showed the greatest range (between 0 per cent and 100 per cent in short form and 13 per cent and 100 per cent in long form). In areas where the OPA form was more instructive, such as with

Self-evaluations which did not shy away from taking a stance either way on a project handling and effectiveness gave the richest insights to operations project effectiveness, there was less variance (between 50 per cent and 100 per cent for long form, and 70 per cent and 100 per cent for short form). EvD should further revise the template project efficiency sub-sections and accompanying guidance notes, and potentially assign different sub-ratings.

The ratings gap in evaluated investment operations

One of the indicators of weaknesses in the self-evaluation system is the difference between ratings assigned by EvD and operation team self-evaluated OPA ratings. This gap is analysed here using the binary ratings gap – which looks at the proportion of ratings that EvD changed from an overall positive rating (*'successful'* or *'highly successful'*) to an overall negative rating (*'partly successful'* or *'unsuccessful'*) or vice versa when it reviewed the project. EvD would expect that the greater attention given to self-evaluation and the increased communication between EvD and bankers should result in less divergent views on project performance.

Table 1: Overall performance binary ratings gap between self-assessment and evaluation 2009-13

Type of review by EvD	Ratings raised substantively by EvD	Ratings substantively unchanged	Ratings lowered substantively by EvD
Validations	1%	86%	13%
Evaluations	2%	76%	20%
All reports	1%	84%	15%

There are two key messages from Table 1:

– There is an upward bias in self-evaluation ratings, with 15 per cent of ratings downgraded by EvD and only 1 per cent upgraded;

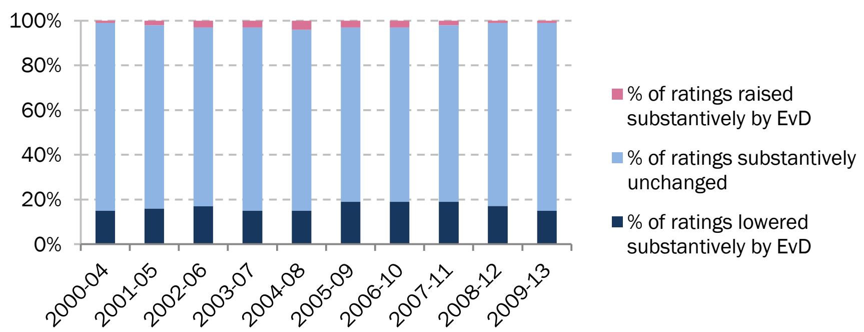
– The ratings gap is substantially higher in full evaluations than in validations.[5]

This is broadly in line with findings from previous years. However, it is worth noting that there has been an improvement in the binary ratings gap from the 2008-12 period. Figure 14 shows the development of the ratings gap over a longer period of time. It shows that the five year rolling average of the gap continued to grow until 2011, before improving in 2012 and (from the preliminary results) and a further reduction in the ratings gap in 2013, although the number of reports involved is not yet large enough to make a definitive judgement. Nonetheless, for the 2009-13 five-year period, the percentage of overall performance ratings substantively unchanged was 84 per cent, the highest figure since 2004-08. A snapshot comparison of the rolling 5-year averages, Table 2 shows a slight increase in the number of ratings that have been left unchanged and that the ratings gap in full evaluations for example has come down significantly.

Table 2: Overall performance, binary ratings gap comparison between 2008-12 and 2009-13

Type of review by EvD	Ratings raised substantively by EvD	Ratings substantively unchanged	Ratings lowered substantively by EvD
Validations	-	-2%	+3%
Evaluations	+1%	+8%	-10%
All reports	-	+3%	-2%

Figure 14: Overall performance binary ratings gap between self-assessment and evaluation - change over time



That said, there is considerable room for improvement. Among evaluations and validations conducted as part of the 2013 work programme alone, EvD concurred with the overall performance rating in the self-evaluation in only 78 per cent of cases, lowered the rating substantively in 20 per cent of cases and raised the rating substantively on one report. These figures are considerably worse than the five year averages, and will need close monitoring in 2014, if the positive longer term trend on the ratings gap is not to be reversed.

The binary ratings gap between Operations self-evaluated ratings and EvD's ratings of projects has decreased slightly since the last evaluation period

A review of 2013 Validations showed that the single biggest reason for EvD revising the overall performance rating downwards was a lowering of the transition impact rating, which accounted for half the downgrades. Reasons for this included reduced scope of the operation when part of it was cancelled; EvD's focus on transition impact it considered actually achieved or realisable, where transition impact monitoring and the OPA itself focused on the future transition potential; and a gap between the achievement of specific transition benchmarks and the achievement of the project's wider

[5] The difference in the rating gap between in-depth evaluations and validations is not a 'like for like' comparison since projects evaluated in-depth are purposefully selected for their learning potential while those validated are randomly selected.

transition aims. Reductions in the OPA ratings for financial performance and achievement of operational objectives also contributed to some downgrades.

EvD has observed increased attention given to self-evaluation, including recently from more senior Banking staff than previously. In a welcome development, at least two teams have incorporated OPA quality into scorecard and appraisal criteria for staff responsible for preparing OPAs in a given year.

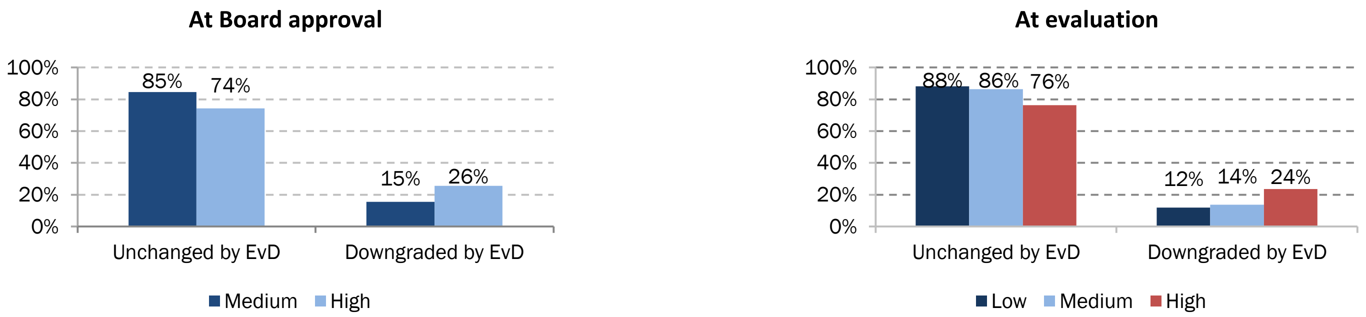
Impact monitoring of investment operations

The Bank sets specific transition objectives and benchmarks for all projects at appraisal, and assesses them for transition potential and risk. Only operations rated at least satisfactory for transition potential proceed to approval. Benchmarks (indicators) are identified to be measures of achievement of the identified sources of transition impact. During the lifetime of the project, progress on achieving the benchmarks is monitored by the operation leader in cooperation with the Office of the Chief Economist (OCE) through the Transition Impact Monitoring System (TIMS) although, as revealed most recently by the TIMS working group under the results taskforce, the actual rating is carried out by OCE with little or no input by the banking team, which consequently does not have much ownership of the transition impact assessment. The possibility that operations may fail to achieve their transition potential is also recognised at appraisal through the transition risk rating.

Analysis of binary changes to transition impact during evaluation was carried out on all projects evaluated between 2009 and 2013.[6] Most of the operations in the sample had '*high'* to '*medium'* transition risk, which reflects the difficulties faced by the Bank's projects and the Bank's willingness to accept risk. From Figure 15 it can be seen that overall, projects with a higher transition risk were more likely to have their transition impact rating substantially changed by EvD and those with lower risk were more likely to be left unchanged. For example, 26 per cent of projects with high transition risk at approval had their transition impact rating downgraded by EvD, whereas only 15 per cent of those with a medium risk rating were later downgraded by EvD[7]. Figure 15 also shows that this finding remains true whether the transition risk rating was at approval or at evaluation, with EvD downgrading the transition impact rating of 26 per cent of projects with '*high'* risk at approval and 24 per cent of those with '*high'* risk rating in TIMS at the time of evaluation.

This analysis only considers binary changes, that is, substantive changes from *'satisfactory'* or above to *'marginal'* or below. It does not capture less significant changes such as a downgrade from *'excellent'* to *'good'* or from *'marginal'* to *'unsatisfactory'*. Including those changes would increase the disconnect ratio for projects evaluated in the last five years to 47 per cent downgraded from their original transition potential at Board approval, and 38 per cent downgraded compared with the TIMS transition potential at time of evaluation. The corresponding upgrade rates are seven per cent and 10 per cent respectively.[8]

Figure 15: EvD ratings change vs transition potential at Board approval and TIMS rating at time of evaluation (binary)



The analysis also showed that overall there was a strong correspondence between EvD transition impact ratings and those in TIMS. Figure 15 shows that 88 per cent, 86 per cent and 76 per cent of projects with '*low',* '*medium'* and '*high'* transition risk ratings respectively, in TIMS at the time of evaluation, were not substantially changed by EVD. Moreover from Table 3, it can be seen that this was also the case for projects where TIMS monitoring was complete by time of evaluation and the remaining risk was classed as negligible. In the evaluation period there were 65 such projects, for which 86 per cent the final TIMS rating was left unchanged by EvD.

Table 3: EvD ratings change vs final TIMS rating (for completed projects) 2009-13

	Binary change in transition impact at evaluation			
Risk	Upgraded by EvD	Unchanged by EvD	Downgraded by EvD	Total
Negligible	2%	86%	12%	65
No. of reports	1	56	8	65

6 A small number of evaluated projects have been excluded from the comparison because they have not been monitored in TIMS. For example, frameworks are not fully covered, and some operations are not individually monitored if there are several facilities with a single client. Furthermore, TCs are not (yet) systematically captured in TIMS.
7 Results for those with low risk were not included as the number of observations in the sample, did not provide a meaningful comparison.
8 In the OPA Validations synopsis paper, EvD compared the transition impact ratings in OPA validations with those in OPAs for validations completed in 2013 only. It considered all changes rather than making a binary comparison. The numbers differ here because both the sample and the comparison are different


1817

Achievements and challenges

In this Chapter:

- An update on implementation of the 2013 Evaluation Policy
- Discussion of progress on improving results specification in project proposals
- Elaboration of new evaluation tools including new guidance notes and information systems
- A proposal for tracking Management follow up on EvD recommendations
- A note about EvD staffing in 2013
- A report on activities EvD has undertaken with international evaluation groups
- Presentation of evaluation work at the EBRD Annual Meeting

Update on strategic initiatives

Evaluation policy

The EBRD evaluation policy approved by the Board in January 2013 incorporated significant changes to policy designed to ensure that evaluation contributes more effectively to better institutional performance by enhancing both its accountability function and its learning function. Substantial changes have followed across the full range of the Bank's evaluation work, with new processes, products and tools well in place by the close of the year. Key developments flowing directly from the policy's new or clarified provisions include:

Evaluation as a tripartite function of the Board, Management and EvD

The policy states that 'evaluation at the EBRD is a Bank-wide effort, incorporating multiple and interconnected roles and responsibilities for EvD, the Board of Directors, and Management.'

Previously, evaluation was seen largely as the function of EvD even though Management had always been required to produce self-evaluations for all completed transactions. However, these self-evaluations were widely perceived as an administrative burden to meet EvD needs rather than as something of value to operations teams or the Bank more broadly. While there is still some way to go to achieve universal embrace of evaluation and its attendant responsibilities on the operational side of the Bank, significant gains have been made. As reported elsewhere in the AER, Management now has full ownership of self-evaluations, while EvD has full ownership of the independent validation of this work. Similarly EvD no longer signs off on the inclusions of lessons in the Board document as responsibility for ensuring this now clearly rests with Management. As described elsewhere, this clearer demarcation of responsibilities has contributed to a more collegial and positive working relationship between EvD and Management though not in any way compromising EvD's independent status.

Management focal points for evaluation matters

The new policy provides that Management 'maintains a focal point(s) for coordination with EvD and the Board (and/or subordinate bodies) on all evaluation-related matters,' where previously there had not been any such clearly assigned responsibility.

In compliance with this provision, Management promptly appointed two focal points – one for project-level evaluations (the Managing Director of the Portfolio Business Group) and another for broader evaluations (Director of Strategy and Policy Coordination). The arrangement is a big advance on what existed before and it is working well. The new arrangements have required some fine-tuning and efforts on all sides, which have been approached in a climate of collegiality and cooperation. There are regular meetings between the Vice President Policy and Director of Strategy and Policy Coordination on the one hand and the Chief Evaluator and Senior Adviser on the other. These meetings have improved communication and the early identification of issues requiring agreement on purpose and joint effort to accomplish it.

Management comments on EvD reports

The policy states 'Management will comment in writing on evaluations as a matter of general practice, indicating areas of agreement and disagreement, unresolved issues, prospective follow-up actions, and potential resource considerations.'

Management is fully complying with this provision. A failure of Management to comment on some EvD reports (often but not always because it agreed with them) was a longstanding frustration to the Audit committee, which wished to have Management's position clearly stated for all evaluations. This has now been resolved. That said, there is still scope to more closely align Management comments with the provisions of the policy.

Management comments on EvD special studies bring additional challenges as these more frequently look at issues cutting across the responsibilities or interests of different units within the Bank. In some cases there can be a lack of consensus on the Management side. This can make it time-consuming for EvD to get comments and sometimes difficult to identify clearly a unified Management position. But this can also be useful as it reflects the reality of the diversity of views. This is not a problem if this diversity is set out clearly and in a way that contributes to discussion and understanding.

Tracking actions on agreed recommendations

The policy provides that Management 'tracks actions on agreed recommendations and periodically reports to the Board on implementation…'

EvD introduced an approach paper to the Audit Committee in 2013 on a new system for tracking Management's implementation of EvD recommendations. This involves production of clear actionable and practical recommendations and agreement from Management to produce a detailed action plan and provide follow up on the implementation status every six months. A draft proposal shared and discussed with EvD's Management focal points met with agreement. Guidelines, templates and a tracking system are scheduled for completion in 2014.

Ensuring an effective system for self-evaluation

The policy commits Management to ensure 'an effective system of self-evaluation and [that it] reports periodically to the Board on its scope and operations.'

Important gains have been made in this area but more remains to be done – an assessment of the quality of self-evaluation reports is contained in the section Quality assessment of self-evaluation (p.28).

To date, Management has not reported to the Board on the scope and operation of the self-evaluation system. Rather, this has been carried out by EvD and reported in the AER, as is the case in the current AER. As it considers appropriate, Management provides comments on EvD's assessment of the self-evaluation system. The intent of the evaluation policy provision is being met though not as originally envisaged. In fact, it provides for a more objective assessment of the self-evaluation system by EvD doing it.

Improved results specification

The policy indicates that Management will ensure 'that proposed operations clearly specify expected results and related performance indicators with sufficient clarity so as to allow effective evaluation' and 'that programmes, policies and strategies identify their expected results with sufficient specificity as to allow effective evaluation.'

Management has made a major advance in improving the results specification for TC and is about to put in place an online system for monitoring and reporting on results achievement. The TC completion report (self-evaluation for TC) has also been completely revamped to focus on the achievement of results in terms of outputs and outcomes. In bringing about these changes Management invited the participation of EvD on a regular basis. In this AER, EvD has taken the TC team's evaluability checklist (developed in cooperation with EvD) and applied it to a sample of TCs approved since the introduction of the new TC submission format with a results framework incorporated as a special topic for this AER (see the section 'Preliminary quality assessment of TC submissions under the new results framework' p. 24).

EvD also participated in the TIMS working group established under the results task force set up by the President of the Bank in 2012. Some important advances were made though progress was rather less than EvD had anticipated. There remains a major deficiency, in EvD's view, in the absence of a requirement to have a results framework based on a theory of change with a hierarchy of results (outputs, outcomes and impacts) logically flowing one from the other and plausibly achievable from the inputs provided and activities to be carried out. Such a hierarchy of results linked by explicitly identified cause and effect relationships sharpens the analytical context and makes monitoring and reporting more useful since a failure to achieve lower level results (outputs) will almost certainly limit the achievement of higher level results; 'early warning' allows corrective action to be taken. As it is, TIMS tracks the achievement of a set of unlinked benchmarks (at least unlinked explicitly) that may be any mix of outputs, outcomes, impacts (rarely) or even activities; thus TIMS provides little useful information for taking corrective action if needed and it is hard to derive a robust conclusion about the achievement of results. Some progress has been made to introduce stronger results mechanisms in specific areas. However, Management has not embraced improvements that would accomplish the fuller and more systematic approach described above.

A joint Board and Management group designed a revamped standard format for Board documents for project proposals. Although the new format contains a section entitled 'measuring/monitoring success,' in EvD's view this falls well short of a robust basis for accomplishing either, for the reasons outlined in the preceding paragraph. A 2012 EvD study ('Performance Metrics: how well do EBRD projects specify expected results?') developed and applied an evaluability checklist to a random sample of projects approved between July 2011 and June 2012. Based on a sample application of the checklist developed by the study, the report recommended that it needed refinement to make the evaluability assessment more robust. As a special topic for this AER the checklist has been revised and applied to a random sample of 2013 project approvals (see the chapter on 'Evaluability of operations' (p. 24) for the results of this exercise).

Management has also been working on making country (but not yet sector strategies) more results oriented. It has piloted a 'prototype country strategy results framework' for the Serbia country strategy. This incorporates most of the features that evaluation good practice considers necessary for a more effective results-based approach.

Policy documents currently do not provide an explicit results framework, or any indication of the results expected. For example, although the revised Public Information Policy which is currently on public consultation, provides for monitoring of its implementation there is no mention of monitoring the results of its implementation nor any indication of what will constitute *'success'* or even any identified indicators for tracking and reporting on results flowing from the policy.

Without clear ex ante specification of expected results and their arrangement in a cause and effect results framework there will always be major limits to the ability to make evidence-based statements about the outcomes and impacts of the EBRD's activities and to manage actively for intended results. It also constrains EvD's ability, as it has done for many years, to evaluate achievements against expectations and to communicate the achievements and results of the EBRD's operations in terms that outsiders can understand.

Using evaluation findings and lessons

The 2013 Evaluation Policy states:

> 'Lessons and findings from evaluation will demonstrably be taken into account by Management in the design and approval process for new operations, programmes, policies, strategies and processes. Management develops and refines processes and instruments to take account of lessons, in consultation with EvD, and reflects these in the Operations Manual, other guidance documents and Learning and Development courses.'

The new Board document format noted above has eliminated an explicit requirement to incorporate lessons from evaluation or other experience with the removal of the 'past experience' section. Instead, in the section on strategic fit and key issues banking teams are expected to 'Where particularly relevant [emphasis added], please refer to the Bank's experience in similar transactions, either by explaining how successful features are replicated or adjustments have been made.' This means the incorporation of lessons from experience is discretionary; and relevance is not defined. On the positive side, where teams do deem it relevant to include evaluation lessons, they are now expected to explain how prior experience has influenced the choices made.

EvD therefore has some concerns about the extent to which the provision of the Evaluation Policy that 'lessons and findings will demonstrably be taken into account' is being met. Still, the previous system was judged by most, EvD included, to be a largely ineffective exercise yielding little of value. Time will tell if lessons and prior experience – which are in effect the Bank's embedded knowledge -- get adequate treatment through the new document format. EvD plans to assess the treatment of lessons and findings in Board documents as a special topic in the 2014 AER.

Use of approach papers

The Evaluation Policy provides that:

> 'each EvD study will commence with an approach paper setting out purposes and methods. These will be discussed with and input sought from Management in advance. Staff and client participation in the evaluation process will be encouraged.'

Each evaluation carried out by EvD now has an approach paper. Well prepared approach papers have proven to be a good way to identify issues and evaluation questions that should be explored by the evaluation. Good approach papers have also proven their worth in terms of helping ensure that the methodology adopted will in fact provide robust answers to the questions to be answered. Finally, they have also been well-received by Management in terms of providing an early indication of the direction in which the evaluation is planning to go, and an opportunity to comment. While some variability in these is to be expected in early days, EvD is working to achieve greater consistency in terms of scope and structure on the basis of a formal Guidance Note. Approach papers were published on evaluation internet pages for the first time in 2013. One, for a special study on EBRD's experience with policy dialogue in Ukraine was posted in English and Ukrainian along with a two-page information sheet about the evaluation.

Particularly for special studies, but also for operations evaluations, it is clear that preparation of a good approach paper requires some preliminary work in terms of key document review, first round interviews with the main stakeholders and primary audience for the evaluation, portfolio analysis and literature review. Essentially the study starts and the approach paper is the first deliverable. This has now been accepted and it should lead to better approach papers across the board.

Guidance notes

A major change introduced in the 2013 Evaluation Policy was to separate out evaluation technical guidance from the policy. This ensured that the policy is considerably shorter as its focus is now solely strategic. This change commits EvD to prepare a series of guidance notes on matter previously contained in the policy as well as other areas where no formal guidance exists.

Since approval of the policy EvD has produced seven guidance notes covering:

- Preparation of approach papers
- Support for preparation of Operations Performance Assessment reports
- Guidance on core data collection
- Preparation of Operations Performance Assessment Validations
- Quality standards for evaluation
- Sampling methodology
- Procedures for using the external advisory panel

A draft guidance note on performance rating methodology has been developed with input provided by three external peer reviewers. It is a significant revision to the guidance for the performance rating of transactions (addressing a number of issues in current guidance) and fills a guidance gap by covering performance rating of TC. The guidance is being discussed by the TC team for the ex-post evaluation of TC. Given that the changes proposed in performance rating of transactions appear quite radical, the new guidance is being retro-tested on four recently-completed operations evaluations and pilot-tested on two operations in the EBRD Evaluation Work Programme 2014-15 with a view to the new guidance being rolled out in 2015.

Discussion on draft evaluation reports

Another new feature of the 2013 policy is a requirement that 'EvD and Management will meet to discuss draft final evaluation reports to ensure accuracy and, to the greatest extent possible, agree on findings, lessons and recommendations.'

While written peer review comments have always been sought, this provision in the policy is aimed at stimulating a discussion between EvD and Management on key issues rather than a one-way communication in writing. Implementation of this provision has been gradual and where it has occurred it has been welcomed by Management. EvD will seek to fully implement this in 2014.

Distribution of evaluation reports

The policy differentiates between internal distribution and external disclosure. On the former it says 'Evaluation reports will be made available to internal users consistent with confidentiality safeguards. The decision on internal disclosure rests with the Chief Evaluator taking account of advice from Management.'

On external disclosure the evaluation policy says 'The provisions of the Public Information Policy regarding external disclosure will apply.'

Internal distribution has been considerably widened as a result of the policy. EvD is now distributing all project evaluations to the Board in full. Reports are also circulated to key audiences in banking, including the sector team, the regional offices and the full project team. EvD produced a first synopsis of validation reports for the Board containing a summary of key findings both by project and theme. This synopsis was designed to relay key messages for people who may not have time to read the full reports, and to raise awareness of issues with Banking that may need to be addressed. This initiative will continue in 2014, with summary points sent to key audiences within the Bank.

A revised draft of the Public Information Policy is undergoing public consultation as of March 2014. This provides from some modest increase in transparency but falls short of full disclosure of evaluations.

Multi-year work programme

The policy directs EvD to prepare 'multi-year work programmes setting out proposed evaluation priorities, products and activities in detail.'

The EBRD Evaluation Work Programme 2014-15 was a first step towards a multi-year work programme as it identified a number of topics that could be the subject of special evaluations that might be initiated as part of the 2015 work programme. However, it remains challenging to get enough advance notice of future decision points to plan for evaluation to make a contribution to these. In the last six months Management has twice approached EvD with specific requests for evaluative work to contribute to decision-making. However, while these overtures were welcome, a lead time of just a few months makes it highly unlikely that EvD will be able to respond with substantial new evaluation work.

Access to information

The Evaluation Policy is clear that EvD is to have 'access to all internal information required to execute its Board-approved work programme. The fundamental operating presumption is one of full information sharing coupled with full respect for confidentiality.' The section goes on to state 'Management will invite EvD to observe or contribute to Management meetings where this may enhance communication and understanding of issues of common concern.'

The provision on access to information continues something already provided for in the previous policy although in that policy the issue had been much more contentious and it received substantial coverage to spell out the terms of that access, which had been subject to separate Board approval.

That Management would invite EvD to observe or contribute to Management meetings was a new provision in the 2013 policy. The background to this was the fact that implementation of the 2005 decision to make EvD independent of Management with a direct reporting relationship to the Board had the unintended and undesirable consequence of isolating EvD from some internal discussions critical to its ability to contribute. Assurance of EvD's objectivity and independence while avoiding a self-defeating isolation is acknowledged by Board and Management, to be kept by EvD under constant review.

Management has responded positively. EvD now receives the agenda of the Strategy and Policy Committee and can indicate its interest in attending any item. There are periodic invitations to the Chief Evaluator to attend Executive Committee meetings. The Chief Evaluator has regular meetings with the President and semi-annual meetings with banking senior management. EvD is now regularly invited to present to various team meetings.

Establishing a more equal balance between accountability and learning

The Evaluation Policy provides that 'Evaluation plays a critical role … by contributing in two equally-important and mutually-reinforcing ways: by reinforcing institutional accountability for the achievement of results; and, by providing objective analysis and relevant findings to inform operational choices and to improve performance over time.'

This move was already well underway by the time the current evaluation policy was approved in January 2013 as it had been reflected in the EvD Work Programmes for 2012 and preceding EvD strategy document. This has seen EvD direct a greater share of its efforts to producing strategically relevant special studies and somewhat less effort to deep-drilling individual evaluations. Generally, this shift has been welcomed by Management though as noted above, it does pose some challenges in terms of Management comments.

Staffing EvD

In 2013, the Chief Evaluator approved a new structure for EvD to better align this with the new strategic priorities for the department. Key features of the new structure include:

- A move away from having specialist evaluation manager positions such as senior evaluation manager (environment), less suited to delivering the Department's changed product mix;
- Reprogramming two administrative staff positions to positions performing more front-line work, such as analytical support;

– Extension of the contract of the senior adviser has allowed him to take some management responsibilities as provided for under the original contract – the two evaluation analysts and two administrative officers now report to the senior adviser.

Whilst awaiting formal confirmation of the departmental structure and staff movements, the department operated for much of 2013 without a full complement of staff members, with several vacancies and first-ever secondments to other parts of the Bank. All vacancies have now been filled with four of the five new recruits on board. The new staffing provides an opportunity to increase the output of the department, and the resources to improve report quality and the timeliness of their completion. The additional staff will also provide greater depth and flexibility to react to opportunities.

Facilitating use of evaluation reports

Library of reports

In 2013, EvD introduced a new electronic library of past evaluation reports for the Bank. This replaced a previous version which had limited capacity. The new library allows EvD to easily share reports with key people, including Board members, through embedded links, both within evaluation reports and correspondence. It draws on archives and has resulted in complete and consistent records accessible by all within the EBRD.

Lessons

The new Lessons Investigation Application, also released in 2013 for banking staff, contains lessons from Operation Performance Assessments and project evaluations, and has direct linkages to reports in the Library for ease of reference. Revised content includes up to date and relevant material edited for better comprehension. Lessons in the new application are searchable by numerous fields such as product type, sector, country, region and key topics. In 2013, the Operations Committee developed a new template for Bankers submitting project proposals for approval, with a new requirement to include past experience throughout the document, rather than having one section devoted to it, which often led to a superficial cut and paste of lessons. It is hoped that the new template will lead to project proposals that consider lessons in all areas; this will be subject to review and follow up in 2014 as noted above. A version for external disclosure is envisaged.

Workflow tracking tool

To enhance learning and improve accountability in the course of conducting evaluations, EvD has introduced a new workflow tracking system for all OPA validation reports to capture key information during the draft review and approval stages. This platform enables more uniform and accessible overview of evaluation quality to take place, capturing comments from operations teams, and information on timing and status for reporting. It enables capture of key data (at the point of entry) and has the capacity to be further tailored to match specific reporting needs. The workflow system ensures each evaluation is produced according to the same review procedures and will allow for identification of gaps and weaknesses which may need to be addressed, which should lead to more even quality standards. This tool will be expanded in 2014 to include the OPA reports.

Engagement with the international evaluation community

EvD benefits greatly by engaging with the international evaluation community and in turn it contributes to that debate, particularly in the area of the evaluation of private sector operations. There are two main vehicles for international engagement by EvD and the nature of the engagement with both in 2013 follows.

Evaluation Cooperation Group of the International Finance Institutions

The Evaluation Cooperation Group (ECG) is main vehicle used for international engagement. This is a grouping of the heads of evaluation and other senior staff members of nine international finance institutions along with three permanent observers and two aspiring members. The three permanent observers provide the members with a formal link to the Evaluation Network of the OECD/DAC and the United Nations Evaluation Group. The independent evaluation office of the Global Environment Facility is also a permanent observer.

Formed in 1995 at the behest of the Development Committee, ECG's purpose was to establish a more harmonised methodology for evaluation of projects. Given this, the ECG has expended considerable effort on development separate good practice standards for the evaluation of private and public sector operations and country programme evaluations. The private sector good practice standards are now in their fourth edition. Periodic benchmarking exercises, some independently carried out, check members' compliance with the standards.

When formed in 1995, most members' evaluation departments were not independent and ECG has played an important role in promoting and protecting independence, which is now a condition of membership. It has produced good practice standards for the independence of the evaluation function in international finance institutions.

ECG was perhaps a little slow to react to the changes taking place in the evaluation departments of its member institutions – in particular, the growing importance of learning and contribution to improved institutional performance and consequent reduced emphasis on the evaluation of individual transactions; and the increased proportion of sector, country, thematic and corporate evaluations in work programmes. However, ECG has recently decisively changed direction to be aligned with these changes. Moves are underway to explore an integrated set of good practice standards covering all types of project and TC evaluation, particularly given the blurring of the boundaries between public and private sector operations. Also, the good practice standards are no longer the centrepiece of ECG meetings. These now focus much more on shared learning – both in terms of the findings of major evaluations carried out by members and the methodologies employed.

Evaluation Network of the OECD Development Assistance Committee

The Evaluation Network of the OECD/DAC also meets twice a year. It brings together 25 bilateral (including the European Commission) development partners, seven multilateral organisations and a number of observer members. Through its participation EvD gains valuable insights on methodological approaches, means of knowledge sharing and bilateral thinking on evaluation matters and multilateral performance. Regular participation by a number of ECG members (EvD included) has contributed greatly to reducing the high degree of suspicion that existed between many bilateral agency evaluation units and those of the IFIs. Most bilateral agencies contract out evaluations to (supposedly) independent consultants whereas IFIs have in-house evaluation units, which many bilateral evaluation departments considered to be not genuinely independent of their managements. At the most recent meeting of the network, the strengths and weaknesses of the various models were openly and constructively debated to the great interest of all.

Annual meeting

EvD participated in the 2013 Annual Meeting in Istanbul, with the Chief Evaluator hosting a special session for Civil Society Organisations. The Chief Evaluator presented the new EvD Policy and key initiatives for improving evaluation within the Bank.

For the first time, EvD also had an institutional stand to present significant studies conducted over the last year. All participants, including shareholders and other stakeholders including Civil Society Organisations had the opportunity to take away fliers with key evaluation messages or compact reports of interest. Access was provided to a new EvD Annual Meeting website landing page which showcased key pieces from 2012.

As a result of the successful presence of EvD in 2013 and this excellent opportunity to further disseminate EvD work, it was hoped, with Board support, that EvD's Annual Evaluation Review could be launched at future Annual Meetings to reach a key audience.

Annexes 1 and 2 are included in order to comply with the Good Practice Standards for Private Sector Evaluation (GPS), which have been developed jointly by the evaluation departments of major multilateral financial institutions under the auspices of the Evaluation Cooperation Group.

Annex 1: Selection of investment projects for evaluation

This Annex contains:

- How the sample of projects for evaluation are selected;
- Information about projects selected in 2013, including:
    - Information about the size and representation of the sample
    - A note about the standard error in the sample
    - A description of the sample
    - A comparison of the evaluation database with the Bank's portfolio.

Identification of the population of projects ready for evaluation

The process for selecting projects for evaluation is based on the Good Practice Standards for Private Sector Evaluation (GPS) of the ECG. Each year, unevaluated operations are reviewed to identify those that have reached early operating maturity. According to the GPS, this is achieved when:

i) The project financed has been substantially completed

ii) The project financed has generated at least 18 months of operating revenues for the company

iii) The EBRD has received at least one set of audited annual financial statements covering at least 12 months of operating revenues generated by the project.

In practice, EvD does not have this information readily available for all projects. It therefore sets as a working assumption that loan operations can be ready for evaluation 18 months after last disbursement, and equity operations two years after last disbursement. It sends operation teams a list of projects in their area that will reach this status within the evaluation year. Each operation team then identifies the projects expected to meet all three criteria for early operating maturity in the course of the year.

Excluded from the population are:

- dropped and cancelled investments where no disbursement has been made ;
- very small investments made under large frameworks (which are generally evaluated on a programme basis through a Special Study);
- certain follow-on operations, such as minor capital increases or investments undertaken to help finance further expansion or cost overruns on projects previously financed by the EBRD, especially where such follow-on operations did not have separate objectives against which performance could be evaluated.

The GPS also allow the exclusion of 'jeopardy' cases, which in the EBRD's case means projects that have been transferred to the Corporate Recovery Unit for special handling. EvD follows the advice of the Director for Corporate Recovery on the timing of evaluations of these projects.

Subject to these exclusions, the population includes all investments that have reached early operating maturity, plus any unevaluated investments that have already been closed, even if they never reached early operating maturity (for example, prepaid operations).

Projects not expected to reach early operating maturity during the year are excluded from the population and rolled forward for inclusion in a future year. Investments are included in the population only once (that is, only for the year in which they will have reached early operating maturity).

Selection of the sample of projects for evaluation

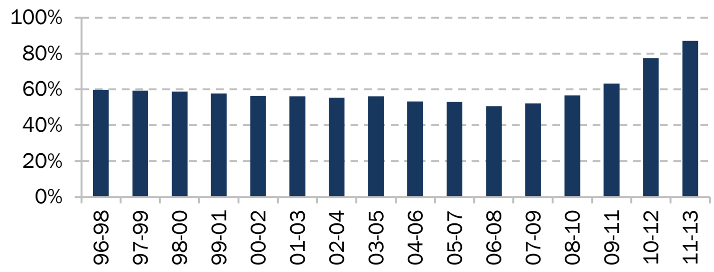
Once the population of projects ready for evaluation has been identified, EvD takes a random, representative sample of sufficient size to establish, for a combined three-year rolling sample, performance rates at the 95 per cent confidence level, with sampling error not exceeding ±5 percentage points, for key performance indicators. This procedure has been followed for the last five years (starting 2009) to ensure EBRD compliance with the GPS.

The Chapter 'Aggregate performance' (p. 8) is based on findings from the randomly selected operations.

The sampled projects may be evaluated through Operation Evaluations or lighter OPA Validations. EvD elects to prepare Operation Evaluations for a subset of sample projects with the aim of maximising the potential for learning lessons. Some additional projects may also be selected purposively for evaluation through Operation Evaluations, again with an exercise of judgement as to prospective insights and lessons; these remain outside the sample and have not been included in the results presented in the Chapter on 'Aggregate performance' (p.8).

Figure 16 below shows the proportion of the results derived from Validations (or their predecessor, XMR Assessments) rather than from more in-depth operation evaluations. The proportion has risen in recent years with a greater focus on self-evaluations validated by EvD.

Figure 16: Proportion of results derived from OPA validations, 1996-2013



Projects selected in 2013

In 2013, 62 projects were randomly sampled for addition to the evaluation database. During the year some projects dropped out of the sample (see Annex 3 p.53) and others were not completed in time to be included in this year's AER. In total, 49 operations from the 2013 work programme were added to the evaluation database. Of these, four were evaluated through Operation Evaluations and 45 through OPA Validations.

Size and representation of the sample

The random sample is intended to achieve statistical significance over a three-year rolling period. This section therefore considers the latest such period, projects randomly selected for evaluation from 2011 to 2013.

Standard error of the sample

The GPS specify that the sample should be of sufficient size to establish, for a combined three-year rolling sample, success rates at the 95 per cent confidence level, with sampling error not exceeding ±5 percentage points. In the three years 2011 to 2013, there was a combined population of 264 individual operations ready for evaluation, excluding the sub-operations of large frameworks. Of these, 132 were evaluated by EvD. Thus the overall coverage ratio was 50 per cent. At a confidence level of 95 per cent, the standard error of the sample was 6.04 per cent, outside the limit set by the ECG. This shortfall occurred because some operations originally selected for evaluation over the period turned out not to be ready for evaluation, while some more are still scheduled for an evaluation but have not been completed in time for reporting in this AER. From 2013, EvD started selecting a slightly higher starting coverage ratio to allow for drop-outs. It is expected that this will bring the standard error down to the target level by 2014, assuming that the work programme is completed in good time.

Not all projects are rated for every indicator. Table 4 shows the standard error for each indicator at the binary level.

Table 4: Summary performance and sample errors for projects evaluated 2011-13

Indicator	Binary success rate	No. of rated operations	Population size	Standard error of the sample
Overall performance	54%	132	264	6.04
Transition impact	79%	131	264	6.09
Environmental and social performance	92%	126	264	6.32
Extent of environmental change	35%	119	264	6.67
Additionality	86%	132	264	6.04
Financial performance	72%	127	264	6.28
Achievement of operational objectives	82%	131	264	6.09
Bank handling	92%	132	264	6.04

Description of the sample

Turkey, a relatively new country of EBRD operations is under-represented because few Turkish projects have yet reached evaluation maturity

The sample of projects evaluated over the three-year period from 2011 to 13 comprises 132 operations, of which 17 are covered by Operation Evaluations and 115 by OPA Validations. They total €3.997 million in business volume. Table 5 below compares the sample with the Bank's active portfolio of projects as at the end of December 2013, with reference to instrument type, sovereign risk type, industry sector and geographic region. There are some differences between the sample and the portfolio. The most obvious of these is the over-representation of Russia at the expense of other regions. This is mainly because of the relatively larger size of projects in Russia. Turkey is under-represented because it is a relatively recent addition to the Bank's countries of operations and few Turkish projects have yet reached evaluation maturity. The sample also appears to under-represent sovereign operations slightly. This may be because sovereign operations tend to have longer total lifespans and are therefore over-represented in the active portfolio compared to the number and volume of projects originally undertaken. The sampling process attempts to match the population of projects ready for evaluation (rather than the Bank's total active portfolio) in terms of the number of operations in each category (rather than volume). Given the different target of the sampling process, some differences of this kind can be expected and are not a cause for concern.

Table 5: Comparison of the evaluation database 2011-13 with the Bank's portfolio

	Evaluation database 2011-2013		**EBRD portfolio December 2013**	
	MEUR	%	MEUR	%
	3,997		37,840	100
Instrument type				
Debt	3,413	85	29,585	78
Equity	584	15	8,255	22
	3,997	100	37,840	100
Sovereign risk				
Non-sovereign	3,461	87	31,429	83
Sovereign	536	13	6,411	17
	3,997	100	37,840	100
Sector				
Energy	921	23	8,146	22
Financial Institutions	912	23	9,235	24
Industry, Commerce and Agribusiness	1,286	32	11,273	30
Infrastructure	878	22	9,186	24
	3,997	100	37,840	100
Region				
Central Asia	269	7	2,736	7
Central Europe and the Baltic states	566	14	6,375	17
Eastern Europe and Caucasus	584	15	7,020	19
Russia	1,582	40	8,936	24
South-eastern Europe	863	22	8,993	24
Southern & Eastern Med.	0	0	579	2
Turkey	133	3	3,202	8
	3,997	100	37,841	100

Annex 2: Evaluated project performance by year of approval

This Annex contains:

- The evaluated project performance statistics behind each criteria by year of project approval.
- The criteria include: overall performance, transition impact, financial performance , environmental and social performance, extent of environmental and social change, additionality, achievement of operational objectives and bank handling

Overall performance

Figure 18: Chart of overall performance by approval year

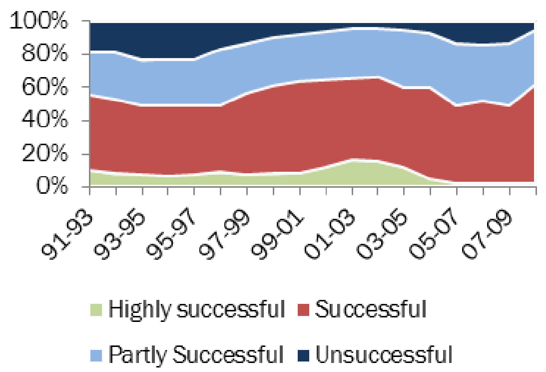


Figure 17: Overall performance by approval year

Year	Highly successful		Successful		Partly Successful		Unsuccessful		Total Reports
	#	%	#	%	#	%	#	%	#
'91-'93	7	10	32	46	18	26	13	19	70
'92-'94	9	8	54	46	33	28	22	19	118
'93-'95	11	7	68	44	38	25	36	24	153
'94-'96	11	7	72	43	47	28	38	23	168
'95-'97	14	7	80	43	53	28	40	21	187
'96-'98	14	9	69	42	52	32	28	17	163
'97-'99	12	7	80	49	48	30	22	14	162
'98-'00	11	8	77	53	42	29	14	10	144
'99-'01	13	8	86	55	44	28	12	8	155
'00-'02	17	12	78	53	43	29	9	6	147
'01-'03	23	16	70	49	43	30	6	4	142
'02-'04	20	15	68	52	38	29	6	5	132
'03-'05	15	11	65	49	45	34	7	5	132
'04-'06	6	4	75	56	43	32	10	7	134
'05-'07	4	3	72	46	58	37	21	14	155
'06-'08	3	2	70	50	47	34	20	1	140
'07-'09	4	3	58	47	46	37	16	13	124
'08-'10	3	4	44	58	25	33	4	5	76
'09-'11	5	22	24	22	19	44	3	11	51

Figure 19: Overall performance by sector (projects approved 2005 to 2010)

Sector	Highly successful		Successful		Partly Successful		Unsuccessful		Total reports
	#	%	#	%	#	%	#	%	#
ENE	2	7	18	64	5	18	3	11	28
FIN	3	3	42	48	32	37	10	11	87
ICA	1	1	40	47	34	40	10	12	85
INF	1	3	16	52	12	39	2	6	31
All	7	3	116	50	83	36	25	11	231

Figure 20: Overall performance by region (projects approved 2005 to 2010)

Region	Highly successful		Successful		Partly Successful		Unsuccessful		Reports
	#	%	#	%	#	%	#	%	#
CA	2	8	11	44	9	36	3	12	25
CEB	1	4	14	54	9	35	2	8	26
EEC	0	0	29	54	21	39	4	7	54
RUS	3	4	37	49	25	33	10	13	75
SEE	1	3	20	54	14	38	2	5	37
TUR	0	0	1	33	2	67	0	0	3
Regional	0	0	4	36	3	27	4	36	11
All	7	3	116	50	83	36	25	11	231

Transition impact

Figure 22: Transition impact by approval year

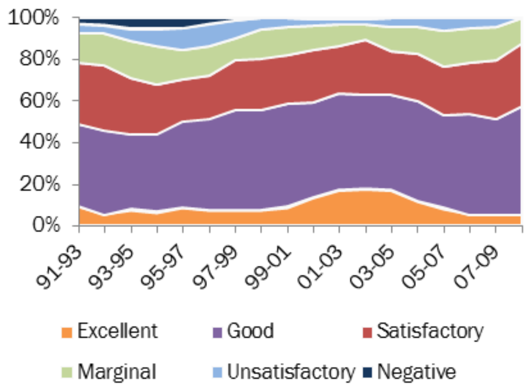


Figure 21: Transition impact by approval year

Year	Excellent		Good		Satisfactory		Marginal		Unsatisfactory		Negative		Total reports
	#	%	#	%	#	%	#	%	#	%	#	%	#
'91-'93	6	9	28	40	21	30	10	14	3	4	2	3	70
'92-'94	6	5	48	41	37	31	18	15	5	4	4	3	118
'93-'95	11	7	56	37	42	27	27	18	9	6	8	5	153
'94-'96	10	6	66	39	38	23	31	18	14	8	9	5	168
'95-'97	15	8	79	42	38	20	26	14	20	11	9	5	187
'96-'98	11	7	73	45	34	21	23	14	17	10	5	3	163
'97-'99	11	7	79	49	39	24	17	10	14	9	2	1	162
'98-'00	10	7	70	49	36	25	20	14	8	6	0	0	144
'99-'01	13	8	78	50	36	23	21	14	7	5	0	0	155
'00-'02	19	13	68	46	37	25	18	12	4	3	1	1	147
'01-'03	24	17	66	46	33	23	15	11	3	2	1	1	142
'02-'04	23	17	60	45	35	27	10	8	3	2	1	1	132
'03-'05	22	17	61	46	28	21	15	11	6	5	0	0	132
'04-'06	15	11	65	49	31	23	17	13	6	4	0	0	134
'05-'07	12	8	70	45	37	24	26	17	10	6	0	0	155
'06-'08	7	5	68	49	35	25	23	16	7	5	0	0	140
'07-'09	6	5	57	46	35	28	20	16	5	4	0	0	123
'08-'10	4	5	39	52	23	31	9	12	0	0	0	0	75
'09-'11	4	8	20	40	19	38	7	14	0	0	0	0	50

Figure 23: Transition impact by sector (projects approved 2005 to 2010)

Sector	Excellent		Good		Satisfactory		Marginal		Unsatisfactory		Negative		Reports
	#	%	#	%	#	%	#	%	#	%	#	%	#
ENE	4	15	12	44	5	19	6	22	0	0	0	0	27
FIN	5	6	45	52	14	16	18	21	5	6	0	0	87
ICA	5	6	39	46	31	36	6	7	4	5	0	0	85
INF	2	6	13	42	10	32	5	16	1	3	0	0	31
All	16	7	109	47	60	26	35	15	10	4	0	0	230

Figure 24: Transition impact by region (projects approved 2005 to 2010)

Region	Excellent		Good		Satisfactory		Marginal		Unsatisfactory		Negative		Total reports
	#	%	#	%	#	%	#	%	#	%	#	%	#
CA	2	8	10	40	9	36	2	8	2	8	0	0	25
CEB	3	12	12	46	9	35	1	4	1	4	0	0	26
EEC	0	0	32	59	9	17	12	22	1	2	0	0	54
RUS	6	8	35	47	20	27	11	15	3	4	0	0	75
SEE	5	14	15	42	10	28	5	14	1	3	0	0	36
TUR	0	0	1	33	2	67	0	0	0	0	1	0	3
Regional	0	0	4	36	1	9	4	36	2	18	2	0	11
All	16	7	109	47	60	26	35	15	10	4	3	0	230

Financial performance

Figure 26: Chart of financial performance by approval year

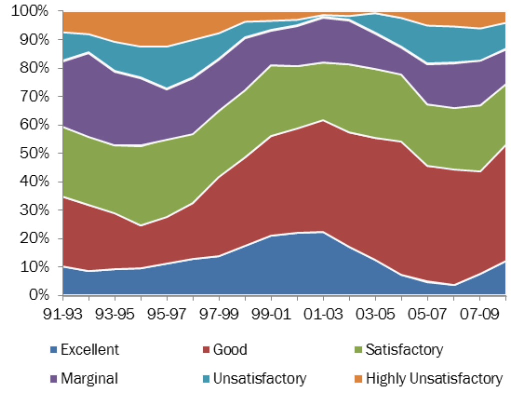


Figure 25: Financial performance by year of approval

Year	Excellent		Good		Satisfactory		Marginal		Unsatisfactory		Highly Unsatisfactory		Total reports
	#	%	#	%	#	%	#	%	#	%	#	%	#
'91-'93	7	10	17	25	17	25	16	23	7	10	5	7	69
'92-'94	10	9	27	23	28	24	34	29	8	7	9	8	116
'93-'95	14	9	30	20	36	24	39	26	16	11	16	11	151
'94-'96	16	10	25	15	47	28	40	24	20	12	19	11	167
'95-'97	21	11	31	17	51	27	33	18	28	15	23	12	187
'96-'98	21	13	32	20	40	25	32	20	22	13	16	10	163
'97-'99	22	14	45	28	37	23	29	18	15	9	12	8	160
'98-'00	24	17	43	31	33	24	25	18	8	6	5	4	138
'99-'01	31	21	52	35	37	25	18	12	5	3	5	3	148
'00-'02	31	22	52	37	31	22	20	14	3	2	4	3	141
'01-'03	31	22	55	40	28	20	22	16	1	1	2	1	139
'02-'04	22	17	52	40	31	24	20	16	2	2	2	2	129
'03-'05	16	13	55	43	31	24	16	13	9	7	1	1	128
'04-'06	9	7	60	47	30	24	12	9	13	10	3	2	127
'05-'07	7	5	60	41	32	22	21	14	20	14	7	5	147
'06-'08	5	4	54	41	29	22	21	16	17	13	7	5	133
'07-'09	9	7	44	36	28	23	19	16	14	12	7	6	121
'08-'10	9	12	31	41	16	21	9	12	7	9	3	4	75
'09-'11	9	18	19	33	13	44	3	0	3	0	4	11	51

Figure 27: Financial performance by sector (projects approved 2005 to 2010)

Sector	Excellent		Good		Satisfactory		Marginal		Unsatisfactory		Highly Unsatisfactory		Total reports
	#	%	#	%	#	%	#	%	#	%	#	%	#
ENE	6	21	15	54	2	7	0	0	4	14	1	4	28
FIN	2	3	37	47	13	16	13	16	12	15	2	3	79
ICA	7	8	23	27	26	31	13	15	10	12	6	7	85
INF	1	3	16	53	7	23	4	13	1	3	1	3	30
All	16	7	91	41	48	22	30	14	27	12	10	5	222

Figure 28: Financial performance by region (projects approved 2005 to 2010)

Region	Excellent		Good		Satisfactory		Marginal		Unsatisfactory		Highly Unsatisfactory		Total reports
	#	%	#	%	#	%	#	%	#	%	#	%	#
CA	2	8	14	56	3	12	6	24	0	0	0	0	25
CEB	0	0	15	58	3	12	3	12	2	8	3	12	26
EEC	6	12	13	25	14	27	8	15	10	19	1	2	52
RUS	3	4	32	46	15	21	10	14	7	10	3	4	70
SEE	4	11	12	33	10	28	3	8	7	19	0	0	36
TUR	1	33	1	33	1	33	0	0	0	0	0	0	3
Regional	0	0	4	40	2	20	0	0	1	10	3	30	10
All	16	7	91	41	48	22	30	14	27	12	10	5	222

Environmental and social performance

Figure 30: Chart - Environmental and social performance by approval year

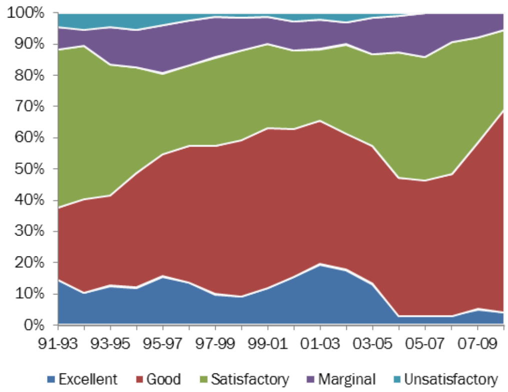


Figure 29: Environmental and social performance by approval year

Year	Excellent		Good		Satisfactory		Marginal		Unsatisfactory		Highly Unsatisfactory		Total Reports
	#	%	#	%	#	%	#	%	#	%	#	%	#
'91-'93	10	14	16	23	35	51	5	7	3	4	0	0	69
'92-'94	12	10	35	30	57	49	6	5	6	5	0	0	116
'93-'95	19	13	44	29	63	42	18	12	7	5	0	0	151
'94-'96	20	12	61	37	56	34	20	12	9	5	0	0	166
'95-'97	29	16	73	39	48	26	29	16	7	4	0	0	186
'96-'98	22	14	71	44	42	26	23	14	4	2	0	0	162
'97-'99	16	10	77	48	46	28	21	13	2	1	0	0	162
'98-'00	13	9	72	50	41	29	15	10	2	1	0	0	143
'99-'01	18	12	78	51	41	27	13	9	2	1	0	0	152
'00-'02	22	15	68	48	36	25	13	9	4	3	0	0	143
'01-'03	27	19	64	46	32	23	13	9	3	2	0	0	139
'02-'04	23	18	57	44	37	28	9	7	4	3	0	0	130
'03-'05	17	13	57	44	38	29	15	12	2	2	0	0	129
'04-'06	4	3	56	44	51	40	15	12	1	1	0	0	127
'05-'07	4	3	62	44	56	39	20	14	0	0	0	0	142
'06-'08	4	3	58	45	54	42	12	9	0	0	0	0	128
'07-'09	6	5	62	53	39	34	9	8	0	0	0	0	116
'08-'10	3	4	49	65	19	25	4	5	0	0	0	0	75
'09-'11	2	4	41	80	6	12	2	4	0	0	0	0	51

Figure 31: Environmental and social performance by sector (projects approved 2005 to 2010)

Sector	Excellent		Good		Satisfactory		Marginal		Unsatisfactory		Highly Unsatisfactory		Total reports
	#	%	#	%	#	%	#	%	#	%	#	%	#
ENE	3	11	16	57	8	29	1	4	0	0	0	0	28
FIN	0	0	46	58	24	30	9	11	0	0	0	0	79
ICA	1	1	38	48	33	41	8	10	0	0	0	0	80
INF	3	10	11	37	10	33	6	20	0	0	0	0	30
All	7	3	111	51	75	35	24	11	0	0	0	0	217

Figure 32: Environmental and social performance by region (projects approved 2005 to 2010)

Region	Excellent		Good		Satisfactory		Marginal		Unsatisfactory		Highly Unsatisfactory		Total reports
	#	%	#	%	#	%	#	%	#	%	#	%	#
CA	0	0	14	58	8	33	2	8	0	0	0	0	24
CEB	0	0	17	68	7	28	1	4	0	0	0	0	25
EEC	1	2	25	47	18	34	9	17	0	0	0	0	53
RUS	4	6	30	45	24	36	9	13	0	0	0	0	67
SEE	1	3	19	51	14	38	3	8	0	0	0	0	37
TUR	1	33	2	67	0	0	0	0	0	0	0	0	3
Regional	0	0	4	50	4	50	0	0	0	0	0	0	8
All	7	3	111	51	75	35	24	11	0	0	0	0	217

Extent of environmental change

Figure 34: Extent of environmental change by year of approval

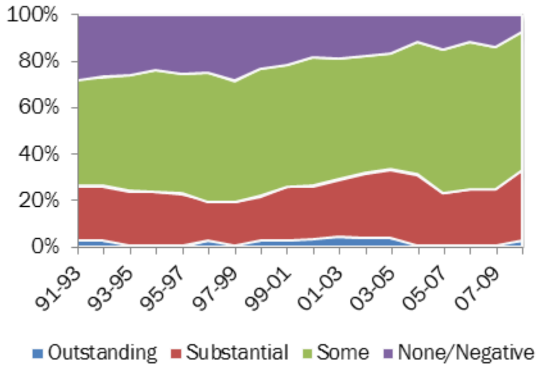


Figure 33: Extent of environmental change by approval year

Year	Outstanding		Substantial		Some		None/Negative		Total reports
	#	%	#	%	#	%	#	%	#
'91-'93	2	3	16	23	32	46	19	28	69
'92-'94	3	3	27	23	55	47	31	27	116
'93-'95	2	1	34	23	76	50	39	26	151
'94-'96	2	1	37	22	88	53	39	23	166
'95-'97	2	1	40	22	97	52	47	25	186
'96-'98	4	2	28	17	90	56	40	25	162
'97-'99	3	2	28	17	85	52	46	28	162
'98-'00	4	3	27	19	79	55	33	23	143
'99-'01	4	3	35	23	80	53	33	22	152
'00-'02	5	3	32	22	80	56	26	18	143
'01-'03	6	4	34	24	73	53	26	19	139
'02-'04	5	4	36	28	66	51	23	18	130
'03-'05	5	4	37	29	64	50	21	17	127
'04-'06	2	2	36	29	71	58	14	11	123
'05-'07	1	1	30	22	84	62	20	15	135
'06-'08	1	1	29	24	78	64	14	11	122
'07-'09	2	2	25	23	68	62	15	14	110
'08-'10	2	3	21	30	42	60	5	7	70
'09-'11	1	2	15	32	27	57	4	9	47

Figure 35: Extent of environmental change by sector (projects approved 2005 to 2010)

Sector	Outstanding		Substantial		Some		None/Negative		Reports
	#	%	#	%	#	%	#	%	#
ENE	0	0	16	53	12	40	2	7	30
FIN	0	0	6	7	65	78	12	14	83
ICA	2	3	24	30	43	54	11	14	80
INF	2	5	15	38	19	48	4	10	40
All	4	2	61	26	139	60	29	12	233

Figure 36: Extent of environmental change by region (projects approved 2005 to 2010)

Region	Outstanding		Substantial		Some		None/Negative		Reports
	#	%	#	%	#	%	#	%	#
CA	0	0	6	23	15	58	5	19	26
CEB	0	0	11	38	16	55	2	7	29
EEC	1	2	12	22	38	70	3	6	54
RUS	2	3	13	18	43	61	13	18	71
SEE	1	2	17	41	20	49	3	7	41
TUR	0	0	1	50	1	50	0	0	2
Regional	0	0	1	10	6	60	3	30	10
All	4	2	61	26	139	60	29	12	233

Additionality

Figure 38: Chart- additionality by year of Board approval

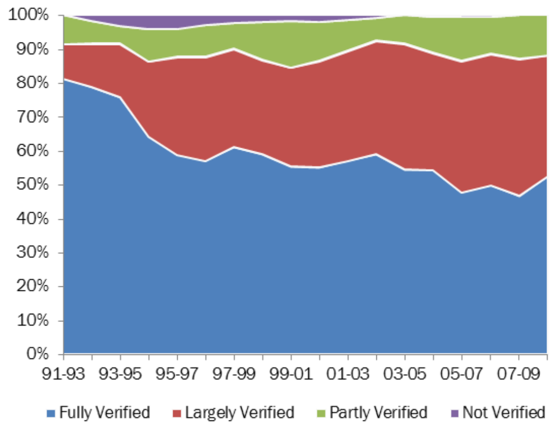


Figure 37: Additionality by year of Board approval

Year	Fully Verified		Largely Verified		Partly Verified		Not Verified		Total reports
	#	%	#	%	#	%	#	%	#
'91-'93	57	81	7	10	6	9	0	0	70
'92-'94	93	79	15	13	8	7	2	2	118
'93-'95	116	76	24	16	8	5	5	3	153
'94-'96	108	64	37	22	16	10	7	4	168
'95-'97	110	59	54	29	16	9	7	4	187
'96-'98	93	57	50	31	15	9	5	3	163
'97-'99	99	61	47	29	12	7	4	2	162
'98-'00	85	59	40	28	16	11	3	2	144
'99-'01	86	55	45	29	21	14	3	2	155
'00-'02	81	55	46	31	17	12	3	2	147
'01-'03	81	57	46	32	13	9	2	1	142
'02-'04	78	59	44	33	9	7	1	1	132
'03-'05	72	55	49	37	11	8	0	0	132
'04-'06	73	54	46	34	14	10	1	1	134
'05-'07	74	48	60	39	20	13	1	1	155
'06-'08	70	50	54	39	15	11	1	1	140
'07-'09	58	47	50	40	16	13	0	0	124
'08-'10	40	53	27	36	9	12	0	0	76
'09-'11	22	43	19	37	10	20	0	0	51

Figure 39: Additionality by sector (projects approved 2005 to 2010)

Sector	Fully Verified		Largely Verified		Partly Verified		Not Verified		Total reports
	#	%	#	%	#	%	#	%	#
ENE	20	71	5	18	3	11	0	0	28
FIN	34	39	40	46	13	15	0	0	87
ICA	38	45	35	41	11	13	1	1	85
INF	22	71	7	23	2	6	0	0	31
All	114	49	87	38	29	13	1	0	231

Figure 40: Additionality by region (projects approved 2005 to 2010)

Region	Fully Verified		Largely Verified		Partly Verified		Not Verified		Total reports
	#	%	#	%	#	%	#	%	#
CA	12	46	10	38	4	15	0	0	26
CEB	33	61	9	17	12	22	0	0	54
EEC	32	43	39	52	4	5	0	0	75
RUS	20	54	15	41	2	5	0	0	37
SEE	0	0	2	67	1	33	0	0	3
TUR	2	18	5	45	3	27	1	9	11
Regional	114	49	87	38	29	13	1	0	231
All	15	60	7	28	3	12	0	0	25

Achievement of operational objectives

Figure 42: Chart - Achievement of operational objectives

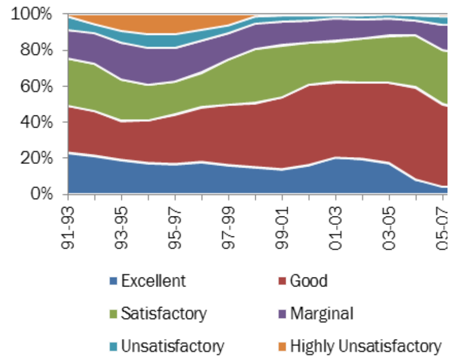


Figure 41 Achievement of operational objectives by approval year

Year	Excellent		Good		Satisfactory		Marginal		Unsatisfactory		Highly Unsatisfactory		Total reports
	#	%	#	%	#	%	#	%	#	%	#	%	#
'91-'93	16	23	18	26	18	26	11	16	5	7	1	1	69
'92-'94	25	22	29	25	30	26	20	17	6	5	6	5	116
'93-'95	29	19	33	22	34	23	31	21	11	7	13	9	151
'94-'96	29	17	40	24	33	20	35	21	12	7	18	11	167
'95-'97	31	17	52	28	34	18	35	19	15	8	20	11	187
'96-'98	29	18	50	31	31	19	29	18	11	7	13	8	163
'97-'99	26	16	54	34	40	25	23	14	8	5	9	6	160
'98-'00	21	15	50	36	42	30	20	14	5	4	2	1	140
'99-'01	21	14	60	40	43	29	20	13	5	3	1	1	150
'00-'02	23	16	65	45	33	23	18	13	4	3	1	1	144
'01-'03	29	21	59	42	32	23	18	13	2	1	1	1	141
'02-'04	26	20	56	42	32	24	14	11	3	2	1	1	132
'03-'05	23	17	59	45	34	26	13	10	3	2	0	0	132
'04-'06	11	8	69	51	38	28	11	8	4	3	1	1	134
'05-'07	8	5	70	45	46	30	22	14	7	5	2	1	155
'06-'08	6	4	59	42	44	31	23	16	6	4	2	1	140
'07-'09	9	7	51	41	37	30	20	16	5	4	1	1	123
'08-'10	7	9	38	51	23	31	6	8	0	0	1	1	75
'09-'11	7	14	21	42	19	38	2	4	0	0	1	2	50

Figure 43: Achievement of operational objectives by sector (projects approved 2005 to 2010)

Sector	Excellent		Good		Satisfactory		Marginal		Unsatisfactory		Highly Unsatisfactory		Total reports
	#	%	#	%	#	%	#	%	#	%	#	%	#
ENE	4	0	14	50	6	21	3	11	0	0	1	4	28
FIN	2	0	44	51	26	30	9	10	5	6	0	0	86
ICA	5	0	39	46	25	29	13	15	1	1	2	2	85
INF	4	0	11	35	12	39	3	10	1	3	0	0	31
All	15	0	108	47	69	30	28	12	7	3	3	1	230

Figure 44: Achievement of operational objectives by region (projects approved 2005 to 2010)

Region	Excellent		Good		Satisfactory		Marginal		Unsatisfactory		Highly Unsatisfactory		Total reports
	#	%	#	%	#	%	#	%	#	%	#	%	#
CA	2	0	12	48	7	28	2	8	2	8	0	0	25
CEB	2	0	11	44	9	36	1	4	1	4	1	4	25
EEC	2	0	29	54	15	28	7	13	1	2	0	0	54
RUS	2	0	34	45	26	35	10	13	3	4	0	0	75
SEE	6	0	17	46	10	27	4	11	0	0	0	0	37
TUR	1	0	1	33	1	33	0	0	0	0	0	0	3
Regional	0	0	4	36	1	9	4	36	0	0	2	18	11
All	15	0	108	47	69	30	28	12	7	3	3	1	230

Bank handling

Figure 46: Bank handling by year of approval

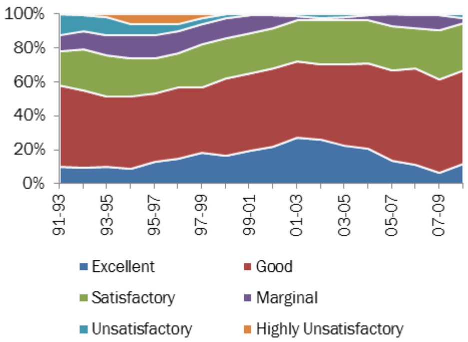


Figure 45: Bank handling by approval year

Year	Excellent		Good		Satisfactory		Marginal		Unsatisfactory		Highly Unsatisfactory		Total reports
	#	%	#	%	#	%	#	%	#	%	#	%	#
'91-'93	7	10	33	48	14	20	7	10	8	12	0	0	69
'92-'94	11	9	53	46	28	24	12	10	11	9	1	1	116
'93-'95	15	10	63	42	36	24	18	12	16	11	3	2	151
'94-'96	15	9	73	44	35	21	23	14	13	8	8	5	167
'95-'97	24	13	76	41	38	20	27	14	11	6	11	6	187
'96-'98	24	15	69	42	32	20	22	13	7	4	9	6	163
'97-'99	29	18	62	39	41	26	19	12	5	3	4	3	160
'98-'00	23	16	64	46	33	24	17	12	3	2	0	0	140
'99-'01	29	19	69	46	35	23	16	11	1	1	0	0	150
'00-'02	31	22	66	46	34	24	11	8	1	1	0	0	143
'01-'03	38	27	63	45	34	24	3	2	2	1	0	0	140
'02-'04	34	26	58	44	34	26	2	2	3	2	0	0	131
'03-'05	30	23	63	48	34	26	3	2	2	2	0	0	132
'04-'06	28	21	67	50	34	25	4	3	1	1	0	0	134
'05-'07	21	14	83	54	40	26	11	7	0	0	0	0	155
'06-'08	16	11	79	56	33	24	11	8	1	1	0	0	140
'07-'09	8	6	68	55	36	29	11	9	1	1	0	0	124
'08-'10	9	12	42	55	21	28	2	3	2	3	0	0	76
'09-'11	6	12	26	51	17	33	1	2	1	2	0	0	51

Figure 47: Bank handling by sector (projects approved 2005 to 2010)

Sector	Excellent		Good		Satisfactory		Marginal		Unsatisfactory		Highly Unsatisfactory		Total reports
	#	%	#	%	#	%	#	%	#	%	#	%	#
ENE	5	18	17	61	5	18	0	0	1	4	0	0	28
FIN	8	9	53	61	23	26	3	3	0	0	0	0	87
ICA	13	15	36	42	27	32	8	9	1	1	0	0	85
INF	4	13	19	61	6	19	2	6	0	0	0	0	31
All	30	13	125	54	61	26	13	6	2	1	0	0	231

Figure 48: Bank handling by region (projects approved 2005 to 2010)

Region	Excellent		Good		Satisfactory		Marginal		Unsatisfactory		Highly Unsatisfactory		Reports
	#	%	#	%	#	%	#	%	#	%	#	%	#
CA	1	4	17	68	6	24	1	4	0	0	0	0	25
CEB	5	19	13	50	5	19	2	8	1	4	0	0	26
EEC	8	15	19	35	24	44	2	4	1	2	0	0	54
RUS	7	9	46	61	19	25	3	4	0	0	0	0	75
SEE	5	14	28	76	2	5	2	5	0	0	0	0	37
TUR	2	67	0	0	1	33	0	0	0	0	1	0	3
Regional	2	18	2	18	4	36	3	27	0	0	2	0	11
All	30	13	125	54	61	26	13	6	2	1	3	0	231

Annex 3: Proposed changes to the performance rating methodology

As described in ‘Preparation of a new guidance note on performance rating’ (p. 12) of the main report, EvD has prepared a draft guidance note on the evaluation performance rating system for discussion with Management and Board. The draft note proposes a number of significant changes to the rating methodology, including:

- Closer alignment with OEC/DAC evaluation criteria (relevance, effectiveness [termed results in the guidance note], efficiency, sustainability and impact) as the basic structure of performance assessment (with some customisation to the EBRD’s unique mandate);
- A revised set of criteria and a revised, more explicit, set of sub-criteria along with clearer guidance on how each should be assessed;
- Adoption of a results framework approach consisting of a hierarchy of inputs (and activities), outputs, outcomes and impacts (with definitions based on OEC/DAC) and the logical connection between them as a more robust basis for assessing performance in situations (commonly occurring) where results (particularly impacts and sometimes outcomes) are not fully observable at evaluation, or they are not measurable because data is unavailable or incomplete;
- Eliminating the separate consideration of the achievement of operational objectives, transition impact, environmental and social objectives and the results of attached TC, policy dialogue and staff contribution to capacity development (some of which were not previously routinely considered in performance rating) with all results considered under the criterion of ‘results’ and identified as outputs, outcomes or impacts;
- Creating a clear separation between (a) results that can be observed at evaluation or can plausibly be inferred and (b) those where their future achievement is largely speculative – the former are rated and included in the overall performance rating while the latter are assessed but not rated and so not included in the overall performance rating;
- Giving more explicit attention to unintended results, positive and negative results and use of the counterfactual;
- Assessing the plausible contribution of the EBRD’s operations to the achievement of outcomes and impacts rather than trying to establish attribution, which may not be possible within the time and resource limitations of evaluation in the Bank
- Some changes to what is included in the overall performance rating: rating criteria are proposed as relevance, results and efficiency (of resource use and process efficiency) while non-rating criteria are forward looking assessment and client and co-financier contribution; an optional criterion exists for innovation and merit features, which is not included in the overall rating;
- Adoption of a numeric scoring and averaging system to derive criterion ratings from sub-criteria ratings and overall performance ratings from criteria ratings;
- Establishment of clear boundaries to separate sub-criteria, criteria and overall performance ratings with provision for the legitimate exercise of evaluator discretion within identified 'grey zones' that straddle the boundaries;
- Incorporation of guidance for the performance rating of TC (discussions are already underway for these to be adopted by the TC team);
- Adoption of six rather than four categories for overall performance rating to allow for a more differentiated performance (previously most performance ratings fell into the categories of successful and partly successful with few operations being rated highly successful or unsuccessful)
- Dropping the word 'successful' from rating category descriptors and its replacement with 'satisfactory' as a less value-laden term.

Annex 4: Further detail on evaluability assessments of TC and investment operations

Preliminary quality assessment of TC submissions under the new results framework

Introduction

Following the Grant Co-financing Strategic Review in 2012, the EBRD's management introduced a results framework, aiming to create a single unified system that meets internal management needs and donor accountability requirements. As part of this framework, it is mandatory for a completed results matrix to be attached to the approval documentation submitted to TC Com for new TC operations. This requirement has been in force since July 1 2013, with the only exceptions being in cases where the TC is for an extension and the original date of submission predates July 1 2013; and as might be expected for non TC, such as for use of non-TC grants used alongside TCs as part of an investment package.

The rollout of the results matrix process was accompanied by training presentations and a set of user centred guidelines available on the intranet. The guidelines provide full instructions and explain how the matrix aims to define measureable results that TCs are intended to achieve prior to completion at output and outcome levels. Within the matrix, outcome indicators must specify target values, timeframes and baselines, and output statements must specify target values and timeframes. Further, a transactional TC[9] results matrix will be aligned with the investment operation's Transition Impact Monitoring System. In those cases where TC involves more than one separate consultancy assignment under a framework or programme package, one integrated results matrix should be submitted.

Together with the TC Team, EvD has developed a TC submission evaluability[10] checklist that has been used to support teams as part of the results framework rollout, and to support a broader quality at entry assessment of TC submissions by the TC Committee itself. The checklist is attached as Table 8 below. Given that the introductory phase has now passed, and that the checklist has broader relevance as a quality assurance and learning tool, as part of the 2013 AER, EvD has undertaken a preliminary assessment of the quality of TC submissions under the new results framework. In addition to the checklist used by Management, the EvD separately checked TC submissions to ascertain if risks were captured, in a way that reflects best practice in evaluability.

Of the 176 TC requests submitted since 1 July 2013, 147 had results matrices. EvD tested the checklist on a stratified random sample of 51 TC requests (with risk matrices) approved by the Bank between 1 July 2013 and 31 December 2013. The process provided immediate feedback on the feasibility of applying the checklist, as well as an initial baseline of experience for future assessments. The sample contained transactional and non-transactional TC, including individual, frameworks and integrated packages across all EBRD sectors, representing different sizes, durations and approval dates. This exercise looked at general evaluability trends, differences between TC types, and other observations to make an initial assessment of and further improve evaluability in the TC submission process.

Table 6: Detail of the sample for the TC evaluability exercise

Period of project approval	Size of population	Size of sample	Coverage
July 2013 – December 2013	133[11]	51	38%

Assessment methodology

The checklists for transactional and non-transaction TCs are shown in Table 8 overleaf. Some evaluability criteria are binary; they are assessed by answering 'yes' or 'no', that is, the statement is true or false when applied to the results matrix.

Other criteria need to be applied to 'multiple instances', that is, the criteria has to be applied to several instances in the same results matrix. In this case a four point rating system is applied with two positive and two negative ratings similar to the overall performance rating applied by EvD (ranging from 'highly successful' to 'unsuccessful').

Table 7: Rating system for the evaluability criteria

All	= criterion was true in every instance – 100% = Yes
Mostly	= criterion was true in >50% and <99% of instances
Few	= criterion was true in <49% and >1% of instances
None	= criterion was not true in any instance – 0% = No

It is important to note that the 'range' on the top and the bottom of this rating scale is very narrowly defined, by 100 and zero per cent respectively, whereas the two 'medium' ranges encompass everything between 50 and 99 per cent or one and 49 per cent.

Taking an analogy with academic exams, one could pass with either 55 per cent or 90 per cent, both of which would be rated 'Mostly'. However, this rating does not reflect the very different quality across these two results.

[9] A 'transactional' TC is one that is directly linked to an investment operation.
[10] 'Evaluability' is understood as 'the extent to which the value generated or the expected results of a project are verifiable in a reliable and credible manner.'
[11] 147 is actual corrected population taken after the analysis was conducted.

From comparable experience in other multi-lateral development banks, ratings of over 90 per cent 'All' for each checklist criterion is considered a high quality benchmark, while ratings below 70 per cent 'All' would be a concern. However, in the initial years of applying the checklist, if the two positive categories together account for at least 90 per cent of each criterion, this is an acceptable quality level to build towards the above high quality benchmark of 90 per cent 'All'.

Table 8: The evaluability checklists (transaction and non-transactional) for TC operations

Results matrix evaluability checklist – transactional TC			
#	*Evaluability criteria*	*Definition applied*	*Assessment type*
Result framework			
1.1	Purpose is found in the investment operation	*The problem should be found in Operation Report (OR) (Sections 1.3 and/or 1.4)*	Yes/No
1.2	TC outcome statement is found in the investment operation	*TC outcome is an output, covenant, activity etc. found in the OR in Section 1 and/or 3*	Yes/No
1.3	Transition impact is found in the investment operation	*TI criteria in OR Section 3.1 repeated in TCR*	Yes/No
1.4	Outcome(s) directly addresses the purpose	*Achieving the outcome will solve the stated problem*	Yes/No
1.4	Outputs are all necessary to achieve the outcome	*All outputs necessary to achieve the outcome(s) have been stated*	Yes/No
1.5	Outcome(s) can be achieved by TC completion	*If all outputs are achieved and risks are mitigated, the outcome will be achieved*	Yes/No
Performance measures			
2.1	Transition impact's indicators specified	*Indicators specified in the TI statement are sufficient to measure the TC's contribution to the TI*	Multiple
2.2	Outcome(s) indicators specified	*Indicators specified are sufficient to measure achievement of the outcome(s)*	Multiple
2.3	Each outcome indicator specifies target value, timeframe and baseline	*Quantitative indicators specify targets, an agreed timeframe & pre-approval baseline;* *Qualitative indicators specify an agreed timeframe & baseline*	Multiple
2.4	Data source specified for each indicator	*Primary or secondary data source for each indicator is specified*	Multiple
2.5	Output statements specify target values and timeframe	*Each output statement specifies a target value & an agreed timeframe for completion*	Multiple
Risks and mitigation measures			
3.1	Key risks to achieving outputs and outcome(s) are specified	*Risks for each output and outcome as appropriate are specified*	Multiple
3.2	Risk mitigation measures are realistic	*Mitigation measures either under control of the TC or can reasonably be expected to be implemented*	Multiple
3.3	No killer risks are included	*If not mitigated a killer risk will stop the outcome being achieved*	Multiple

Results matrix evaluability checklist – non-transactional TC			
#	*Evaluability criteria*	*Definition applied*	*Assessment*
Result framework			
1.1	TC outcome(s) contributes to TI	*Reasonable link between the outcome & TI*	Yes/No
1.2	Outcome(s) directly addresses the purpose	*Achieving the outcome will solve the stated problem*	Yes/No
1.3	Outputs are all necessary to achieve outcome(s)	*All outputs necessary to achieve the outcome(s) have been stated*	Yes/No
1.4	Outcome(s) can be achieved by TC completion	*If all outputs are achieved and risks are mitigated, the outcome will be achieved*	Yes/No
Performance measures			
2.1	Transition impact(s) indicators are specified	*Indicators specified in the TI statement are sufficient to measure the TC's contribution to the TI*	Multiple
2.2	Outcome(s) indicators specified	*Indicators specified are sufficient to measure achievement of the outcome(s)*	Multiple
2.3	Each outcome indicator specifies target value, timeframe and baseline	*Quantitative indicators specify targets an agreed timeframe & pre-approval baseline;* *Qualitative indicators specify an agreed timeframe & baseline*	Multiple
2.4	Data source specified for each indicator	*Primary or secondary data source for each indicator is specified*	Multiple
2.5	Output statements specify target values and timeframe	*Each output statement specifies a target value & an agreed timeframe for completion*	Multiple

Risks and mitigation measures			
3.1	Key risks to achieving outputs and outcome(s) are specified	*Risks for each output and outcome as appropriate are specified*	Multiple
3.2	Risk mitigation measures are realistic	*Mitigation measures either under control of the TC or can reasonably be expected to be implemented*	Multiple
3.3	No killer risks are included	*If not mitigated a killer risk will stop the outcome being achieved*	Multiple

Summary of findings

EBRD TC submissions generally filled out the results matrix adequately. In only one case did the submission fail to fill out the 'data source' section. Similarly, only one of the sample matrices failed to complete the 'output' section. Going beyond the evaluability matrix requirements, the picture is similarly positive. Only one TC submission of the sample failed to score on the 'risk' section. Conversely, a number of TC submissions failed to complete the 'reporting schedule' section of the results matrix. No TC is exempt from producing a completion report, yet three of the sample failed to check that they would carry out either a progress report or a completion report. Separately, there were several cases where the details in the results matrix did not match the TC request form (TCR) or the attached terms of reference, and sometimes it fell short of the TCR's narrative quality on change theory or comprehensiveness in describing the Bank's support package.

The sample of TC results matrices scored very well against the evaluability checklist. Out of a possible score of 100 per cent, including the risk section, the average score was 79 per cent for non-transactional TCs, and 82 per cent for transactional TC. Excluding the risk section, reflecting the bare minimum required by the Bank, sample submissions averaged 76 per cent and 81 per cent respectively. This did not vary much by sector, size or duration, though longer term TCs tended to score very slightly higher on results framework and performance measures on average. Further, overall evaluability scores improved towards the end of the year, suggesting that the ongoing TC Team support and increasing team familiarity with the tool resulted in improved evaluability, and bodes well for future TC Com submissions. Finally, very few of the sample had multiple risk matrices attached, and where this was the case, there were often substantial improvements in the quality of the final risk matrix produced in terms of definition of purpose and outcome indicator descriptions in particular.

Despite the overall high scores, there were also displays of continued weakness. Some risk matrices still confused outcomes and outputs, or outcomes and purpose, and in a few cases the outcome did not directly address the purpose, or perhaps was overly ambitious with the purpose; for example where assuming receipt of a deliverable would contribute to an action alone. Some risk matrices in the sample provided non-specific output targets or unclear outcomes, and many could have been even more specific about the transition impact indicators for the TC. Some risk matrices used a plethora of indicators (outcome and output) and diluted the matrix clarity overall, and some only stated one or two outputs, where the TCR stated the full set which would together be expected to achieve the purpose of the TC.

Further, the evaluability of results matrices varied according to the different sections. For non-transactional TC, the two sections within the risk matrix which scored least were:

– 'TC outcomes contribute to transition impact,' which in almost a quarter of all cases, was assessed as 'No';
– 'Transition indicators are specified,' where over 70 per cent the sample specified either 'few' or 'none'. [12]

This outcome is noteworthy because transition impact in non-transactional TC is not linked to an operation, and so the TC must express its own logic chain within the Bank's larger sectoral or country strategies, and suggests there is considerable scope for improvement.

For transactional TC projects, the sections scoring least were:

– 'Outputs are all that is necessary,' where 35 per cent failed to score any points;
– 'Transition indicators are specified,' where a third of the sample scored 'few' or 'none'.

This is important to reflect on as the EBRD makes efforts to improve the evaluability of TC, because with transactional TCs, often this is an integral and explicit part of the operation and therefore its design must be fit for purpose. Further, in a few transactional TC cases, the stated outcome was rated as not possible to be achieved by the TC, as submitted in the risk matrix. On a deeper look, it is often a failing in the way the TC design is expressed rather than a failing of the design itself, but nevertheless an optimal risk matrix must captured the design fully. A further observation is that several of the transactional TC sample were not referred to in their board documents; in four cases, neither the purpose nor the TC outcome was mentioned in the investment operation. In eight occurrences either one or the other was missing. Even though this number is small, it is an important gap, since these TCs are justified by the fact that they are specifically designed to support the success of investment operations. These findings therefore suggest that there is continued scope for improvement in general.

Interestingly, within the individual segments, there was some variance according to type, size and duration:

– Results framework: Of the four lowest scoring submission samples (< 50 per cent), three came from TCs with long-term duration;
– Performance measures: Of the five lowest scores (< 50 per cent) four were medium-sized, and four were also of long duration;
– Risks: Four of the six lowest scores (<50 per cent) came from transactional TC; five of the six were long term and five of the six were large.

The evaluability exercise also identified a number of good practice cases with regard to results frameworks, performance measures and risks.

[12] The possible ratings for the section in the Evaluability Checklist were: 'All', 'Mostly', 'Few' or 'None'.

Table 9: Three good practice examples

Name of TCR	Type
Cairo Bus Project	Integrated package (of several individual TC numbers)
CIL-Strengthening Capacities for MSME Lending	Transactional TC
Second Extension for Regional (TCS 33133) Joint Vienna Institute Project - Training Provider - Extension	Non transactional TC

Finally, the evaluability checklist, (Table 8 above) proved to be a somewhat effective assessment tool overall. It was found to fulfil the key requirements of such a tool: it contains the necessary criteria, it is straightforward to interpret, and was easy to apply to the risk matrices for the selected sample. Having said this, the tool could be further improved. The checklist lacked an ability to evaluate risk, which is a key qualitative dimension of a TC risk matrix. It was also difficult to 'check', and grade in a nuanced way, other qualitative aspects of the matrix, such as the adequacy of indicator and target choice. In some cases, where the TC is for due diligence, there is little link to transition impact, and the checklist seems to include several inapplicable questions for these cases. EvD will discuss with the TC Team how best to refine the checklist to address these issues.

Evaluability assessment of Board approved projects in 2013

The EBRD's increased focus on evaluability is reflected within the One Bank programme's emphasis on measuring our impact, the recent Strategic Grants review and last year's introduction of a results framework into all new TC project proposals. It is therefore important to review the Bank's investment operations to understand their strengths and weaknesses in terms of evaluability. The review is particularly timely since EvD's latest 'Synopsis of OPA Validations' discusses concerns over OPA quality which stem from an inadequate evaluability base. Conclusions from EvD validations have included comments that project design at appraisal lacks baseline information and monitoring indicators, resulting in poor monitorability and low evaluability, and especially so when it comes to the adequacy of metrics. Moreover, there are indications that in some cases an inadequate results matrix and insufficient performance monitoring against objectively verifiable indicators is one of the reasons for ratings downgrades between banking team OPA submissions and EvD reviews; essential information and data are sometimes lacking as evidence and this opens the scope for a difference of ratings. Ultimately, a review of OPA validations suggests that a closer consideration of evaluability and project monitoring ex ante could improve the quality and integrity of project management. It is therefore an important topic for EvD to look at, which is why the 2013 AER focuses part of its special theme on the evaluability of investment operations.

As part of its support to the progressive improvement of evaluability of EBRD projects, EvD has developed a checklist that specifically assesses the evaluability of investment operations using the Board approved operations report as a base. The checklist treats evaluability of investment operations along three dimensions:

i) expression of a results framework;
ii) expression of a monitoring framework;
iii) treatment of risks.

Notably, the development of such a checklist follows on from the 2012 insights paper 'Performance Metrics - how well do EBRD projects' specify expected results' which included an initial pilot checklist. Taking the pilot checklist as a base, the new checklist further stratifies evaluability sections into:

i) a basic test which includes current requirements and what EvD considers the most important elements of good practice that are broadly in line with current guidance even if not fully specified by it; and,
ii) an 'advanced test' which is in line with accepted best practice.

From comparable experience from other multi-lateral development banks, a score of 75 per cent or more would represent acceptable evaluability, meaning that on most levels a project expressed its results, monitoring framework or treated its risks to a satisfactory degree or more.

EvD tested the checklist on a stratified random sample of 46 operations reports for operations approved by the Bank between 1 January and 31 December 2013. The sample contained projects across all EBRD sectors, representing different sizes, packages (standard/framework) and products (debt/equity/both). This exercise looked at general evaluability trends, differences between operation type, and other observations to provide immediate feedback on the feasibility of applying the checklist, as well as an initial baseline of experience for future assessments.

Table 10: Detail of the sample for the investment operation evaluability exercise

Period of project approval	Size of population	Size of sample	Coverage
January 1 2013 – 31 December 2013	181	50[13]	28%

[13] The sample size went down to 46 and coverage to 25 per cent because four items turned out not to be full Board approval documents.

The evaluability checklist

		Excellent	Satisfactory	Poor	Absent	Notes
No.	**A. Expression of Expected Results**					
	Basic Test					
A1.1	Operational Objectives are clearly described and consistently stated throughout OR					
A1.2	All immediate results flowing from the achievement of operational objectives are specified in terms of outputs and/or outcomes, whether explicitly identified as such or not					
A1.3	Expected transition impact results are clearly specified in terms of outcomes and impacts, whether explicitly identified as such or not					
A1.4	Where part of a project, any policy oriented, environmental and social inclusion results are clearly specified as any or all of outputs, outcomes and impacts, whether specifically identified as such or not					
	BASIC TOTAL	**0.00%**				
	Advanced Test					
A2.1	Existence of plausible logic linking the various levels of results (outputs, outcomes and impacts) to each other and the inputs to be provided by EBRD and others					
A2.2	Expected results (both O and TI) are realistic- neither overly optimistic nor insufficiently challenging- given the inputs to be provided by EBRD and others					
	ADVANCED TOTAL	**0.00%**				
	Basic and Advanced Total	**0.00%**				
	Optional Measure					
A3.1	The ultimate target beneficiary group for project's transition impact is explicitly considered and quantified if possible					
	B. Expression of Monitoring Framework					
	Basic Test					
B1.1	At least (preferably no more than 3) measurable and appropriate indicator is assigned to each expected result					
B1.2	Each indicator has clearly defined and adequate baseline and time bound target					
	BASIC TOTAL	**0.00%**				
	Advanced Test					
B2.1	Monitoring responsibilities are assigned, and a data collection plan is described, responsibility assigned and resourced					
	ADVANCED TOTAL	**0.00%**				
	BASIC and ADVANCED TOTAL	**0.00%**				
	Pre RIsk BASIC Total	**0.00%**				
	Pre Risk BASIC AND ADVANCED TOTAL	**0.00%**				
	C. Treatment of Risks					
	Basic Test					
C1.1	Any potential 'killer' risks are identified, with the likelihood of occurrence indicated, and a justification for proceeding with the project provided					
C1.2	Mitigating actions to be taken (esp including by EBRD), to reduce the possibility and potential impact, are identified for each risk					
	BASIC TOTAL	**0.00%**				
	After risk basic total	**0.00%**				
	After risk basic and advanced total	**0.00%**				

Findings and next steps

The exercise indicated that EBRD operations have just a fair minimum level of evaluability, hovering around an average of 73 per cent on the basic test of evaluability. However, it also revealed a number of ongoing flaws in the checklist which cast doubt on the validity and applicability of the findings:

- The study on Performance Metrics has previously observed that key information on results (other than transition) is often not presented clearly and consistently within the Board approval document but is scattered throughout the document. The current version of the checklist does not assess this issue.
- The checklist does not adequately assess how well the approval document describes the contribution of the operation to higher level impacts, including those identified in country and sector strategies and other Bank policies.
- The qualitative assessment of transition (and other) benchmarks does not penalise weaknesses so long as a bare majority of indicators are valid; this is a weak test of the adequacy of indicators. Nor does it penalise overuse of indicators, though it is clear that a project which lists seven indicators for one transition impact benchmark is not best practice.
- The Performance Metrics study noted some concerns over the 'risk' section of the checklist which led to inconsistent rating between the two consultants conducting the assessment. Consequently, the assessment of risk has moved to a separate part of the checklist but the questions over its inclusion have not been fully resolved.
- The checklist does not reflect the diversity of operations approved at Board level, for example framework operations versus standalone operations.
- Because of efforts to keep the checklist short, it does not sufficiently tease out important dimensions of evaluability, such as separating individual elements in the basic test of indicators.

Given the deficiencies of the checklist, EvD has chosen not to present the detailed findings. Instead, in 2014 it will continue to work on refining the checklist to address the issues identified above. The target will be to develop a revised checklist by the end of 2014, allowing a similar evaluability exercise to be conducted for the 2014 Annual Evaluation Report. The checklist must consistently identify examples of good and bad practice in order to be useful as a tool to improve the evaluability of Bank projects presented for approval.

Annex 5: OPA quality framework checklist

	[Project Name] OPA Quality Questions	Y	N	Comment
1	Explains project consistency with the Bank's policies, strategies and mandate and includes retrospective assessment?			
2	Reviews the appropriateness and accuracy of the original case for additionality, in hindsight (including but not only pulling the argument from board document)			
	Score for Self-assessment of Project Relevance	**0.0%**		
3	Project objectives used in the OPA are clear and consistent?			
4	Project objectives are derived from the Project Description / Business Purpose section of the Board Report Summary Fact Sheet?			
5	Appropriate Project sub-objectives are utilised for the purposes of self-assessment?			
6	Specific Measurement Targets / Indicators are clearly specified? These indicators are consistent with those (where appropriate) presented at approval or have been used for monitoring purposes?			
7	Project operational objectives have been addressed separately from transition impact objectives?			
8	Project/Company financial performance has been analysed appropriately and fully, including any variance analysis as required?			
	Score for Self-assessment of Project Effectiveness	**0.0%**		
9	Full description and analysis has been provided of project origination (who instigated and why)?			
10	Full description and retrospective analysis has been provided of the quality and adequacy of the project due diligence and structuring process? (Critical analysis of decision making? Any subsequent failings or regrets in hindsight have been highlighted?)			
11	Full description and analysis has been provided of the project design to meet objectives in hindsight?			
12	Full description and retrospective analysis has been provided of project risk identification and mitigation?			
13	Full description and analysis has been provided of other project quality at entry aspects such as client relationships, syndication, legal etc.			
14	Full description and retrospective analysis has been provided of the quality and adequacy of the project implementation and monitoring process? (Critical analysis of decision making? Any subsequent failings or regrets in hindsight have been highlighted?)			
16	A new estimate of projected profitability for the Bank has been calculated?			
17	Explanation/justification has been provided of the differences between the new estimates of cumulative project contribution and internal rate of return before and after risk adjustment and the figures from the Board approval document?			
	Score for Self-assessment of Project Efficiency	**0.0%**		
18	Realised/potential transition impact and development outcomes not identified in the TIMS Report have been identified?			
19	Full description has been provided of the remaining transition potential?			
20	Estimation has been provided of the risks to the remaining transition potential? (not just pulled from TIMS as a rating)			
21	Retrospective analysis of the adequacy of the TIMS for the project has been provided?			
22	Assessment of the compliance of the client and sponsor with the Bank's ESP, national and EU standards and EBRD's PRs have been provided?			
23	Implementation of the ESAP has been described and assessed against the agreed benchmarks and milestones?			
	Score for Self-assessment of Project Impact and Sustainability	**0.0%**		
24	Insightful and specific lessons were identified through the self-assessment with little interpretation or rewording necessary			
25	The report is complete, with all sections and tables populated?			
26	All questions in the OPA, incl justifications fully answered? Where appropriate full analysis has been provided of any associated TC, policy Dialogue, project structural changes, covenant waivers, Cooperation with other IFIs etc.			
27	The "storyline" flows logically and persuasively? (All major events in project lifecycle have been explained comprehensively and chronologically)			
29	Takes a coherent approach to results. (IE definition and appropriateness of project objectives, the suitability of relevant indicators for the objectives, and the identification of baseline and benchmarks are common themes)			
30	Sufficient evidence provided throughout the self-assessment in support of given ratings? (relevant for both positive and negative bias)			
31	Do the ratings "add up"?			
	Score for Completeness and Sufficiency of Self-assessment	**0.0%**		
	Score for Overall Quality of Self-assessment	**0.0%**		

Annex 6: Work programme completion report for 2013

Approval of the work programme

The Evaluation Department's work programme for 2013 was set out in Board document and considered by the Audit Committee on 12 June 2013.

Evaluations and special studies

The work programme comprised 17 special studies, operation evaluations and corporate reports. These are the reports that are circulated to the Board of Directors. In addition, work remained to be done on several reports carried over from previous years. Figure 49 summarises the status of reports at the end of March 2014.

Figure 49: Status of evaluation reports, March 2014

Type of report	Carried over from 2012	2013 WP	Changes during 2013	Reports completed	Reports pending
Special studies	6 pending + 2 under review	7	-1 cancelled /postponed	7	4 pending + 3 under review
Operation evaluations	4 pending + 5 under review	7	1>OPAVs -3 cancelled /postponed	6	3 pending + 3 under review
Corporate reports	0	3	0	3	0
Total	10 pending + 7 under review	17	- 5	16	7 pending + 6 under review
OPA validations	12 pending + 8 under review	51	-1 cancelled + 1>OE	51	6 pending + 14 under review
OPA reviews	1 pending +2 under review	34	0	37	0

During the course of the year, some projects turned out to be more or less interesting than expected, and therefore were upgraded from self-evaluations to Operation Evaluations or vice versa. Some planned reports were cancelled. Table 2 below summarises the reasons for cancellations of special studies and operation evaluations.

Figure 50: Reports cancelled in 2013

Report name	Report type	Reason for cancellation
Long-term transition impact of past EBRD projects	Special study	Folded into the 'Equity Projects – Longer Term Performance' special study - to which it is related
North Expedition	Operation evaluation	Turned into OPA Validation
Burgas Water Company	Operation evaluation	Folded into the Bulgaria WWF special study
Mobiasbanca (three operations)	Operation evaluation	Turned into OPA Validation

Tools and methods

The work programme for 2013 specified seven guidance notes under the heading 'Tools and methods'. Their status is shown individually below.

Figure 51 Status of work on tools, resources and processes scheduled for 2013

Tools, resources and processes	Detail as presented in work programme document	Status at end of March 2014
1. Performance Rating Methodology	EvD will produce a clear articulation of the project performance rating methodology in order to improve consistency and comparability across evaluations and between EvD and Operations.	Ongoing: Approach paper has been finalised and a consultant has been engaged.
2. Operations performance assessments (OPAs)	The new self-assessment template being used by Banking may undergo some modifications as a result of its pilot testing. EvD will prepare guidance covering both content and process.	Complete: New OPA templates including guidance notes are currently being used by Banking teams.
3. OPA validations (OPAVs)	Follows from the update to the OPA and will involve consultation on a new model	Complete: Launched October 2013
4. Operation evaluations (ORs)	Follows from the update to the OPA and will involve consultation on a new model.	Yet to be completed.
5. Peer review (Internal and external)	The objective is to establish consistent EvD practice.	Yet to be completed.
6. Sampling methodology	The previous Evaluation Policy contained substantial technical detail on the methodology used to select operations for evaluation; a guidance note would be a preferred vehicle for this information.	Complete: Guidance note on the methodology for selection and sampling of projects for evaluation is now on the intranet.
7. Communicating evaluation results	Internal guidance for integrating communications elements more consistently and systematically into individual EvD pieces of work.	Yet to be completed: Preparatory work has been undertaken, expect this to be completed in 2014.
* (additional) - Quality Standards for Evaluation	A quality framework to provide assist in the aim of producing evaluations of consistent quality and provide guidance as to how quality in EvD work is measured.	Complete: An EvD Quality Framework Checklist was produced that can be applied to all evaluations. It has been put together through reference to the DAC Quality Standards for Development Evaluation produced by the OECD.

Annex 7: Management comments

Key points

- Enhancing its focus on results, the EBRD has recently introduced result frameworks (RFs) at various levels of its activities in addition to the transition impact monitoring system (TIMS) that sets objectives for and tracks the transition results of investment projects. As of July 2013 it introduced results framework for the TC operations. Results frameworks are also being designed and will be rolled out for country strategies from the second half of 2014. Management is streamlining the various RFs, focusing on RFs for strategic initiatives and country strategies as the main building blocks of the Bank's results framework architecture, in addition to RFs for projects (investments and TCs) that are the basic building blocks of this architecture. Management will regularly consider the effectiveness and efficiency of the new results frameworks and propose necessary adjustments. Management therefore finds that the following statement in the Annual Evaluation Report does not reflect reality: "Management has consistently resisted the introduction of a functioning results framework,…, and on the basis of argumentation that is largely unpersuasive.…"
- Management has provided extensive comments on the issue of evaluability - results framework presentation - for investments projects that was the focus of the 2013 EvD special study, "EvD Special Study – Performance Metrics - How well do EBRD projects specify expected results?" The study proposes a presentation of results in a hierarchy that involves a "cause and effect" cascade from project objectives through outputs to outcomes to ultimate impact to replace aspects of the current transition impact benchmarks methodology and terminology used in the project documents. Management believes that given the Bank's mandate to promote transition to sustainable market economy, the causal link between the physical objectives of the project, outputs, and the higher level objectives, impact, is much weaker compared to typical development bank projects, and may be applicable for some types of projects only. Management did propose to EvD to jointly work on case studies to help clarify the types of projects where there may be benefits from introducing the suggested EvD approach. However, EvD declined to take joint action in this area on grounds of lacking sufficient resources.
- Management emphasises that the methodological difference between the transition impact rating of EvD and Management's TIMS ratings at the time of EvD evaluations affect the ratings gap in evaluated investment operations. Management believes that, while EvD could in parallel use its own evaluation methodology in line with international best practices (such as OECD DAC), the transition impact performance of operations should primarily be assessed against the same approved methodology that is used by Management.
- Management stresses that while the TC evaluability methodology was agreed with EvD, it appears to have been misapplied in EvD's analysis of a sample of TC projects submitted to TC Com during the first six months of the implementation of the TCRF. This has led to some incorrect findings and conclusions which Management considers undermining the effectiveness of the new TC results framework. Management believes these findings and conclusions may well warrant review and reconsideration.

Aggregate Performance – Chapter 2

Management welcomes and values the aggregate performance discussion. The findings of a general positive trend with improvements across different components, such as success rate, transition impact rating, and the overall performance are encouraging. Similarly, Management welcomes the improving ratings for environmental and social performance and change. Given the timeframe, Management suggests that it is a result of the introduction of the Performance Requirements in 2008, which brought a more systematic approach to appraisal of environmental and social issues, as well as the successful introduction of the Sustainable Energy Initiative.

Management notes however, that besides normal sample limitations some interpretive observations or postulations could be better based on evidence. These include statements such as: "…the Bank has assigned insufficient staff to manage projects in challenging contexts", or "environmental and social impact is a significant factor only in a minority of projects". Management believes that such important conclusions would warrant focused and deeper analyses on these particular issues. Evidence based findings are important if the results of performance analysis are used by the Bank to learn and take action to improve performance.

Findings from 2013 evaluations - Chapter 3

Management values the summary discussion of the findings and the overview of Management comments for evaluations conducted in 2013, and the thematic discussion of key findings from Operation Performance Assessment (OPA) validations. Management has already provided extensive comments on the specific findings of individual evaluation studies during the year as also summarised in the report.

Management comments focus on the cross-cutting issue of a number of these evaluations, the ex-ante results specification and presentation. The issue of results specifications ex-ante is discussed again in the chapter entitled Evaluability of operations (p. 24) under the section Evaluability assessment of 2013 Board approved investment projects and Preliminary quality assessment of TC submissions under the new results framework (both p.24). Also, in the chapter on Achievements and challenges (p.34) in the section Improved results specification (p.34) states that although the new format of the Board document designed by a joint Board and Management group "contains a section entitled 'measuring/monitoring success,' in EvD's view this falls well short of a robust basis for accomplishing either,…".

Results presentation was the focus of one of the 2013 EvD special studies, "EvD Special Study – Performance Metrics - How well do EBRD projects specify expected results?" The study proposes a hierarchy of results that involves a "cause and effect" cascade from project objectives through outputs to outcomes to ultimate impact. Management has provided extensive comments to this study (available on online, link) and would like to iterate some of the key comments.

- Management believes that given the Bank's mandate of transition impact, the causal link between the physical objectives of the project, outputs, and the higher level objectives, impact, is much weaker compared to typical development bank projects aimed at the objective of poverty alleviation. The application of the causal results chain between transition objectives (impact), and project-level outputs or outcomes, may be applicable for some types of projects, (for example were demonstration effect of a new technology is a transition objective), but not for others (such as objectives related to regulatory or institutional reform).
- Management has asked the EvD to jointly work on case studies to help clarify the types of projects where there may be benefits from introducing this approach, (which would then need to be assessed against the costs), and looks forward to the joint future actions in this area. The final report of the TIMS Working Group finds in a sample of 57 projects signed between 2004 and 2007, that transition benchmarks already and as appropriate include a combination of output, outcome, and impact indicators, although not using the terms. In almost all analysed projects, benchmarks include "outcome" indicators, and in over 81 per cent of projects include both "output" and "outcome" indicators,
- Management also notes that the statement "...in many cases the benchmarks actually have limited value as indicators of performance..." appears at odds with the finding of the report that there is a strong correspondence (in 86 per cent of the cases for negligible risk) between TIMS transition impact rating and the EvD rating, as transition benchmarks are the main basis for TIMS assessment.

Special topic: evaluability assessments of TC and investment operations - Chapter 4

Management appreciates EvD's initiative to apply the evaluability checklist to a sample of TC projects submitted to TC Com in the first six months of applying the new TC Results Framework including the results matrix. Management agrees that improving the results matrix presentation is a challenge, and is encouraged by the positive overall findings of the evaluability study that suggests even in these early days of applying the results matrix staff are generally on the right track. The TC Team's support to TC users in 2014 will focus on improving the quality of TC submissions with a specific emphasis on their results matrices and how these are reported on.

Management questions the basis for the findings regarding the weak measures of the TC contribution to the transition impact for non-transactional TC, "TC outcomes contribute to transition impact,' where in almost a quarter of all cases, the results matrix scored either 'few' or 'none". Management notes that this section requires the assessor to make a value judgement on the strength of the relationship between the outcome statement and the identified transition impact criteria. As typically only one transition impact criteria (out of a possible seven) is identified for each TC project, a score of few is to be expected. More detailed information is required on those TCs scored as 'none' as this would suggest a complete disconnect between the TC outcome and its stated transition impact criteria.

Similarly, the report concludes that another weakness is the specification of transition impact indicators for both non-transactional and transactional TCs, with respectively over 70 per cent and a third of the sample specified either 'few' or 'none'. Management notes that the guidance provided to operation leaders is to identify the primary transition impact indicator or criteria that the TC project will contribute to. However, in the case of TCs supporting project preparation work such as feasibility studies it is not expected that a transition impact indicator will be documented as further work on the investment operation may not continue. Consequently, a rating of 'few' should be the norm. Although a rating of 'none' would be a concern, Management records show that with the above exception, TC project results matrices submitted to TC Com in 2013 identified at least one transition impact criteria.

Management finds that the EvD judgement that 35 per cent of the sample failed to score any points regarding the link between outputs and the outcome is of serious concern as this suggests a disconnect in the results hierarchy and consequently an inability to effectively evaluate the TC project. Management welcomes a review of those projects that have been assessed in this category and make the necessary adjustments with operation leaders.

Management also notes that the report's statement linking outputs to "...the larger objective" is not clear. Under the new TC Results Framework the term 'objective' has been replaced by 'outcome' in the results matrix. If the report is referring to the link between outputs and the outcome then Management suggests the correct terminology is used. If the 'larger objective' means something else and this link is not being referred to, then the statement should be reviewed and clarified.

Review of self -evaluation and monitoring in EBRD - Chapter 5

Management welcomes the findings of a continuing narrowing gap between ratings produced by self-evaluation and those produced independently by EvD. Management also welcomes better guidelines to improve OPA quality around project efficiency and notes that implementation of procedural details has been underway. It looks forward to further cooperation with EvD in this matter. Yet management feels that it is critical to extend guidance and training for OPAs beyond Banking Department staff and to all those who are involved in projects to extract discernible lessons. To date, the feedback from those staff members who have experienced in the new self-evaluation approach suggests that lesson aspects have yet to be fully transpired. Management would like to highlight the methodological difference between transition impact rating of EvD and the TIMS ratings at the time of EvD evaluations, also noted in the report. First, TIMS only considers transition impact through transition objective identified and benchmarked ex ante. EvD evaluations look at all seven potential sources of transition impact. Second, TIMS transition impact reflects only TI potential and distinguishes between potential and the remaining TI risk rating. EvD's transition impact ratings reflect the realised transition impact at the time of evaluation. Thus a comparison between EvD rating and TIMS rating could be done only for projects with "negligible" or "low" risk ex-post according to the TIMS. Indeed, for such projects the report finds a strong correspondence in 86 per cent of the cases between the two ratings. These methodological changes also affect the ratings gap in evaluated investment operations, as also noted in the report, *"...the single biggest reason for EvD revising the overall performance rating downwards was a lowering of the transition impact rating, which accounted for half the downgrades"*.