OCTOBER 2019

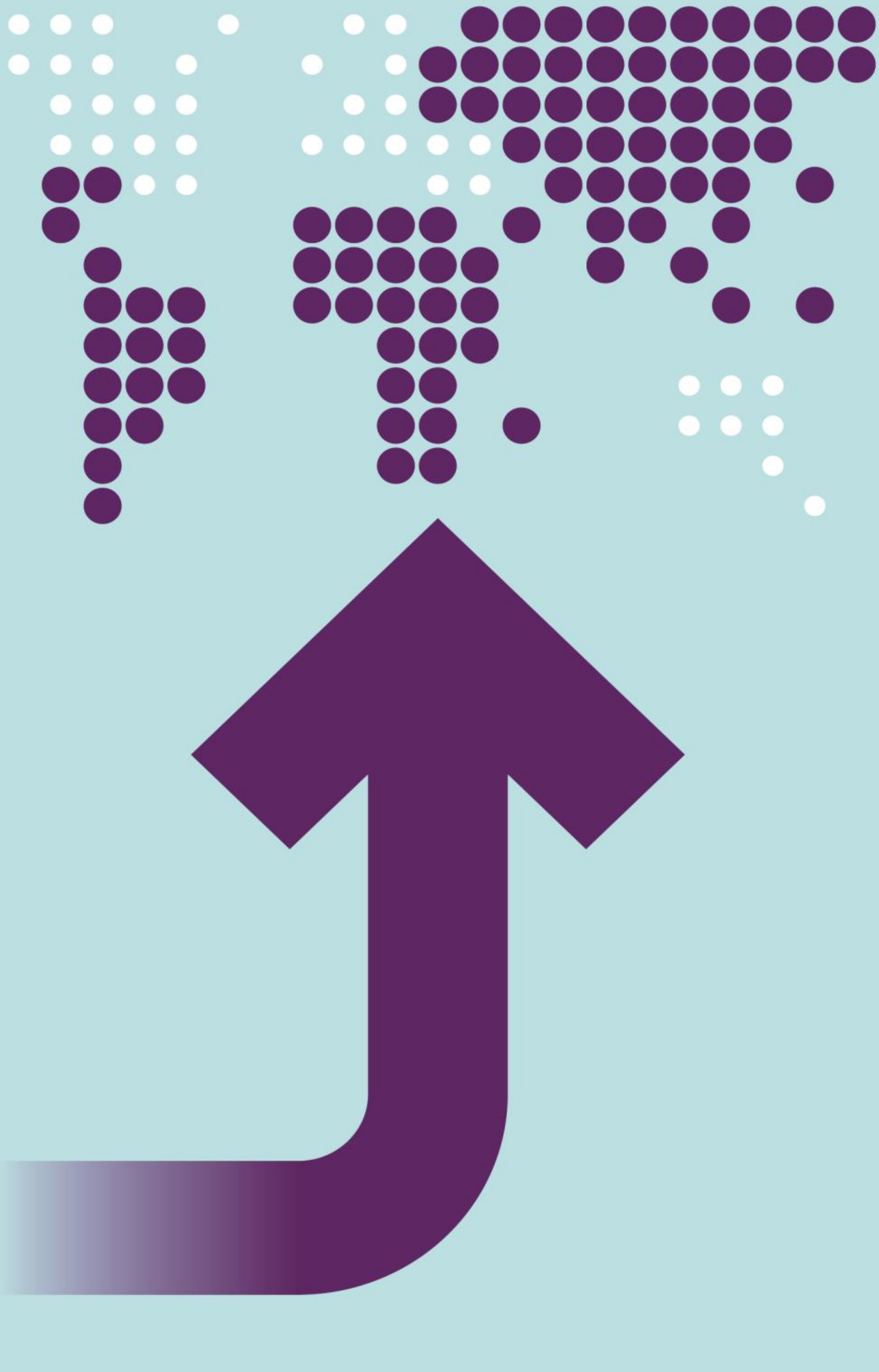
# AFRICAN DEVELOPMENT BANK

## Evaluation of the African Development Bank's self-evaluation system and processes

Independent Development Evaluation Department (IDEV)

Draft Evaluation Report

Centennial Group International



CENTENNIAL GROUP INTERNATIONAL
policy and strategic advisors

# Contents

## Boxes



## Figures



## Tables



## Annexes

## Acronyms and Abbreviations

| | |
|---|---|
| AfDB | African Development Bank Group |
| ADER | Annual Development Effectiveness Report |
| ADF | African Development Fund |
| ADOA | Additionality and Development Outcomes Assessment |
| AsDB | Asian Development Bank |
| ASR | Annual Supervision Report |
| PBO | Programme-based Operations |
| BTOR | Back-to-Office Report |
| CEDR | Comprehensive Evaluation of Development Results |
| CPO | Country Program Officer |
| CPPR | Country Portfolio Performance Report |
| CSP | Country Strategy Paper |
| CSPE | Country Strategy and Program Evaluation |
| DAM | Delegation of Authority Matrix |
| DBDM | Development and Business Delivery Model |
| DO | Development Outcome/Objective |
| ECG | Evaluation Cooperation Group |
| E&S | Environmental and Social |
| ESAP | Environmental and Social Assessment Procedures |
| ESS | Environmental and Social Safeguards |
| BDEV | Independent Development Evaluation |
| IED | Independent Evaluation Department |
| IFC | International Finance Cooperation |
| IFAD | International Fund for Agriculture Development |
| IPR | Implementation Progress Report |
| KPIs | Key Performance Indicators |
| LOC | Line of Credit |
| M&E | Monitoring and Evaluation |
| MDB | Multilateral Development Bank |
| MTR | Mid-term Review |
| NSO | Non-sovereign Operation |
| PAR | Project Appraisal Report |
| PCR | Project Completion Report |
| PCREN | Project Completion Report Evaluation Note |
| PIU | Project Implementation Unit |
| PSR | Project Status Report |
| PPP | Public Private Partnership |
| PRA | Project Result Assessment |
| QA | Quality Assurance |
| QaE | Quality at Entry |
| QoS | Quality of Supervision and Exit |
| RISP | Regional Integration Strategy Papers |
| RISPE | Regional Integration Strategy and Program Evaluation |
| RMC | Regional Member Country |
| RMF | Results Measurement Framework |
| RRS | Results Reporting System |
| SAP | System, Applications and Products |
| SESP | Self-evaluation System and Processes |
| SNOQ | Delivery Performance Management & Results Department |
| SOU | Special Operations Unit |
| TM | Task Manager |
| ToC | Theory of Change |
| TYS | Ten-Year Strategy |
| UA | Unit of Account |
| WB | World Bank Group |
| XSR | Expanded Supervision Report |
| XSREN | Expanded Supervision Report Evaluation Note |

## Executive Summary

**Introduction**

The evaluation of the AfDB's Self-evaluation System and Processes (SESP) comes at a critical time as the Bank has been going through profound changes and revisiting its strategic directions and operational processes. Such effort comes from the realization that ultimately it is the quality of operations that determines the capacity of the Bank to achieve development results. As a consequence, the Bank has initiated over the last couple of years an in-depth diagnostic of the Quality Assurance (QA) processes. This was complemented by a series of BDEV evaluations articulated along a theory of change that aims at examining the relevance, efficiency and effectiveness of the Bank's QA processes across the project cycle as a way to enhance the contribution to performance management, accountability and learning.

This evaluation is conceived in that logic and builds on previous BDEV evaluative work dealing with: i) quality at entry (QaE) and ii) quality of supervision and exit (QoS), as well as a "chapeau"[1] paper coalescing the different findings including the compliance with the Bank's Environmental and Social (E&S) Safeguards. Based on existing evidence, Management prepared a QA Implementation Action Plan, covering 5 areas of reform[2] and setting the basis for the improvement of the SESP, building on this evaluation and management's own diagnosis. This will complete the assessment of the QA framework and will potentially set the stage for institutionalizing best practices in a consistent way and positioning AfDB at par with comparator agencies.

The SESP are commonly defined as the assessment made of a project, country/regional program, and policy/thematic review by the entity engaged in the activity. As illustrated in the Theory of Change (Fig 1, Section II), the main SESP tools being used are the progress reports, the mid-term reviews, and the completion reports. The SESP are meant to pursue three main outcomes: i) performance, ii) accountability and iii) learning, and its functioning is assessed along three main dimensions of relevance, effectiveness and efficiency. The performance of the main SESP tools is summarized in Table 1 below. While not being SESP instruments as such, corporate reporting through dashboards and the Results Measurement Framework (RMF) are partly fed by the SESP.

While the SESP are carried out by staff/Management, BDEV complements the process through the validation of certain products such as the completion reports or through independent evaluations of projects, Country Strategy Papers (CSPs) and Regional Integration Strategy Papers (RISPs).

While project performance relies in good part on the performance of country Governments, partners and local stakeholders as well as external factors, the SESP are entirely under the Bank's control as well as in its capacity to adopt and implement recommendations. Ultimately the evaluation tries to answer the following question: "Do the SESP support performance management, accountability, and learning at the Bank?"; and two underlying sub-questions: i) how well are the SESP performing? and ii) to what extent are the SESP impacting the quality of development results?

The evaluation covers the period 2013-2018, and the implementation of the SESP over the time span that goes from approval to closing (exit). It covers both public sector and non-sovereign operations (NSOs). While specific references are made with respect to NSOs in each chapter, the report is mostly focused on

[1] Evaluation of Quality Assurance across the Project Cycle (2012-2017). October 2018.

[2] The 5 areas of reform are: i) reinforcing the knowledge of operations staff; ii) strengthening project preparation; iii) refining the assurance framework for quality at entry; iv) sharpened focus on delivery and results; and v) improving planning, budgeting and information.

public sector operations. In addition to the project or operation level analysis, case studies for a sample of CSPs and RISPs have also been carried out (Annex 7).

## Main Findings

1. **The Bank's SESP have many positive features. They lay out strong standards and procedures underpinning their functioning as well as a cogent articulation with the independent evaluation function carried out by BDEV.** Annex 5 presents a detailed description of the SESP of comparator institutions showing a good level of alignment which was to be expected from members of the Evaluation Coordination Group (ECG.) Many of the tools and processes in place are consistent with and as good as those of comparator institutions. The articulation between self-evaluation, validation and independent evaluations is similar, and the system produces corporate results data that are used to report to the Board. However, this evaluation finds that progress towards a culture of development effectiveness has been mixed and the potential of the SESP to make an impact on the three fronts of performance management, accountability and learning, is unmet.

2. **The main weaknesses of the SESP are in the application of the established procedures, standards and norms**. While the issues identified as part of this evaluation mirror similar constraints faced by comparator organizations, two factors seem to be affecting the proper functioning of the SESP to a higher degree in the AfDB: i) low compliance with established procedures; ii) little resources for M&E during supervision; and iii) deficient candour and a positive bias in assessing performance. This has affected the credibility of the SESP and contributes to a perception that the system is adding little value.

3. **AfDB has an independent evaluation policy for BDEV approved in 2016 and amended in 2019. However, the Bank does not have an integrated evaluation policy that covers both the independent function and self-evaluation of the Bank itself.** BDEV promotes the use of evaluation findings on specific topics in line with the demand coming from the Board and Management which are also the primary users, but BDEV is not a user of evaluation findings as such. Evaluation adds value only when its findings are used. BDEV evaluation policy necessarily focuses on the supply side of evaluation and not the use side. In the absence of such a common framework, alignment of strategic approaches and harmonization of methodologies and processes will remain at risk. Further, BDEV is not the only generator of evaluation findings. All completed operations and country/regional strategy papers (CSPs/RISPs) are self-evaluated by staff. Some comparator institutions have an integrated evaluation policy in place (IFAD and WBG) or are considering its introduction (AsDB) (Annex III for more details).

4. **Effectiveness of the SESP is constrained by its ratings methodology and structure and the way it is applied to Progress Completion Reports (PCRs)**. Several elements can be highlighted:

   - The Bank uses a 4-point rating system[3], but several discussions took place in the recent past, on the merit of adopting a 6-point system like other comparator organizations (AsDB, IDB, IFAD, the WBG). Admittedly the 6-point system allows for more flexibility and realism in assessing performance of the large pool of projects falling somewhat in between Satisfactory and Unsatisfactory, as the distance between the two is often perceived to be too wide and too stark, according to staff interviews. However, it was also noted that the 6-point scale does not permit to have a clear judgment about project performance, and that such scale would tend to classify performance of most interventions under "Moderately" or "Mostly" Satisfactory anyway, which ultimately will be aggregated together with the Satisfactory and above categories as being "above the bar". Other recommendations made by Evaluation Reference Group (ERG) members or by the 2016 PCR

[3] The scale is: 4-Highly Unsatisfactory (HS); 3-Satisfactory (S); 2-Unsatisfactory (U) and 1-Highly Unsatisfactory (HU).

Synthesis Report are in favour of an odd-number system (3 or 5), as it will allow to better account for the middle space of the distribution (assuming a symmetrical profile). However, an odd number rating scale would be against the agreed Good Practice Standards of the ECG.

- The main risk associated to any rating scale is to assess whether it could lead to more positive ratings without supporting evidence and subsequently to an increased "disconnect" with BDEV ratings. This happens to be the case in the Bank. Any rating system will have pitfalls and whether a 3, 4, 5 or a 6-point rating is adopted, it may not lead to significant improvements unless other measures are put in place for improving the reliability of the system and the implementation of more rigorous procedures for the generation of the required evidence.
- The use of simple averages of sub-ratings and dimensions (relevance, effectiveness, efficiency, sustainability), combined with the "range" methodology, whereby "satisfactory" (or better) is defined as a rating above 2.5, leads to mis-representation of the performance of the portfolio. The system allows to pass the satisfactory bar even if the project did not meet the satisfactory specifications as expected by the level-3 definition. This is particularly the case when results frameworks are too qualitative and imprecise, and leave a large degree of discretion in assessing delivery of outcomes.
- As a consequence, there is a significant difference on project performance depending on how "satisfactory" is defined. At 2.5 and above, 97% of the 137 PCRs analysed[4] are considered to be satisfactory, while at 3 and above it is 80%. Similarly, the "disconnect"[5] between the overall rating assigned by staff in the PCRs, and the lower rating provided by BDEV validation in the respective PCREN is 15% (at 2.5 and above) and 39% (at 3 and above). This raises questions about the credibility of the PCR ratings.
- Using the PCR/XSR ratings as the "official" score instead of the PCREN/XSREN ratings provided by BDEV through the validation process conveys a more positive picture than the reality. This is especially the case for RMF indicators on portfolio performance and departs from the practice of comparator's institutions. The issue is mostly related to the difficulty of generating the required validated information in a timely fashion so as to feed in the submission of the corporate reports;
- Bank and Borrower performance ratings in the PCR are very different from BDEV ratings in the PCREN. It has proven difficult for staff to rate negatively the Bank performance which is easily associated to their own performance and that of their clients, with whom staff try to maintain a good relationship and naturally they don't want to be seen as finger-pointing. The big perceived step between a "satisfactory" and an "unsatisfactory" rating has likely also contributed to the gap.[6]
- The downgrading of ratings by BDEV is a source of tension and is often rebutted by staff arguing that validation is a desk-based exercise that does not account for field realities. Managing ratings will remain a controversial subject as long as it is viewed as a tool to pass judgment on staff performance. Consultations between staff and the BDEV team around the PCR validation process remain rare. In the opinion of almost all the staff interviewed, the rating methodology needs revision. Whether the system is too heavy on ratings or not remains to be seen, but at a minimum there should be an attempt to mitigate the perception that ratings are easy to "game", that BDEV validation is out of context and mostly antagonistic in nature, that the disconnect with BDEV can be largely ignored, and that the rating methodology is too rigid and bureaucratic.

---

[4] BDEV. Synthesis Report on the validation of the 2016 and 2017 PCRs (2019).

[5] The "net disconnect" is the difference between the number of cases in which BDEV provides a higher rating (upgrade) and the number in which it gives a lower rating (downgrade).

[6] IFC, for example, also uses a 4-point scale for indicators, namely excellent, satisfactory, partly unsatisfactory and unsatisfactory. The step between satisfactory and partly unsatisfactory is perceived smaller, even though the latter is clearly in the bottom part of the rating scale and is being reported as such.

As illustrated in the Theory of Change (Section III), the underlying logic of the SESP is that it can play an important role in improving performance management, accountability and learning. The rest of the Findings are organized along these three main outcomes.

### A) Performance Management

5. **The lack of candour in self-evaluation, particularly in Implementation Progress Reports (IPRs)/Annual Supervision Reports (ASRs) and PCRs/XSRs, can be explained among other things by the lack of a proper incentive structure.** The perception that project performance is equated to staff performance undermines the motivation to rate poorly-performing projects candidly. The effectiveness of the SESP is undermined by their being viewed as a compliance mechanism that is driven by box-ticking, protecting one's reputation, and relying on weak generation of evidence (M&E and results frameworks). Candour is also undermined by weak accountability mechanisms. Finally, there is little recognition that being identified as a "problem fixer" could motivate staff towards greater proactivity in raising issues and corrective action.

6. **A number of issues constrain the contribution of the SESP to improving portfolio performance**

   - The Bank's culture, incentives, and institutional KPIs, are skewed in favour of lending approval, like in other comparator institutions, with limited emphasis on the quality dimension and development results. This issue has been recognized by Management and being addressed through the QA Action Plan;
   - There is insufficient attention to incentives that support a culture of quality and results. Opportunities for recognizing, celebrating, internalizing and learning from good quality M&E, results frameworks, proactivity in addressing issues or project restructuring, and successful implementation remain limited (this is also being addressed by the QA Action Plan);
   - There are weaknesses in M&E systems and how they are articulated with baselines and results frameworks. Many PCR ratings were downgraded by BDEV validation for lack of evidence that would support a particular assessment;
   - Excessive focus on accelerating project approval leads to critical design activities getting rushed or carried over to implementation (procurement plans, feasibility studies, validation of E&S assessment studies, setting up of the Project implementation Unit – PIU, baselines). This results in early implementation delays which require a stronger SESP to fix issues from the start;
   - There is a strong tendency to avoid addressing issues through formal project restructuring because the transaction costs are considered to be too high. This results in failure to introduce corrective measures and leads to retain appraisal targets that are no longer in line with the project reality;
   - Increased decentralization and the move towards continuous implementation support, together with the establishment of a new position of Implementation Support Manager for public sector projects is seen positively. However, new roles and the division of responsibilities between staff at HQ and at the country/regional level need clarification;
   - The IPR ratings of the current portfolio feed into the Country Portfolio Performance Report (CPPRs) which in turn also affect the assessment of the CSPs since they are presented in tandem. The Development Objectives (DO) and Implementation Progress (IP) ratings of the last IPR also migrate to the PCRs. Therefore, the compliance and candour issues affecting the IPRs permeate through other SESP outputs and all the way to the RMF. Also, issues of consistency between CSPs and CPPRs' assessment methodologies have emerged recently which require revision (now under way).
   - In the case of NSOs, the lack of clear measuring tools of the progress towards development objectives and the absence of a rigorous and institutionalised M&E system, lead to poor tracking and reporting of achieved results and reduce the likelihood of effective risk mitigation during implementation. This makes it difficult to assess, analyse and report adequately on portfolio

performance. These issues are now being contemplated under the NSOs QA Action Plan (September 2019) adopted by Management.

7. **Project teams and task managers are fundamentally motivated to help clients deliver results and by and large are committed to the supervision task**. However, most task managers and investment officers are overstretched and the additional support required to address issues and help in the proper implementation of the SESP has been wanting. The Bank seems to be short of staffing and skills required to implement the SESP efficiently and effectively. This is particularly the case for M&E and E&S safeguards during supervision, as also documented in the recent evaluation report on integrated safeguards system (ISS)[7].

### B) <u>Accountability</u>

8. **Low compliance stems from insufficient accountability mechanisms and deficient visibility**. In the absence of reliable information, the SESP lose their credibility. The current enforcement and incentive systems do not prevent staff from ignoring the rules that suit them least and getting away with not generating the required outputs. In the words of one of the interviewees: "if the Manager cares, staff will care". Moreover, the Bank has been lagging behind in the development of an IT-supported on-line portal that can provide access to portfolio and SESP data and raise the visibility and the efficiency of the system. Providing real-time information and compliance data to line managers will enhance accountability for supervision, and completion of IPRs/ASRs and mid-term reviews (MTRs) while easing the burden on task managers by simplifying and streamlining reporting requirements. It will also facilitate more effective portfolio reviews and planning exercises. The recent roll out of the Results Reporting System (RRS), as part of the QA Action Plan for both the public sector projects and the NSOs, is meant to address this issue. Once completed, it will facilitate automatizing the reporting of results, timely escalation of issues to senior management, accountability, reliability of information, efficiency of reporting at the corporate level and transparency at implementation.

9. **Reporting tools such as Dashboards and the RMF have proven to be powerful accountability mechanisms and should be enhanced to cover indicators of compliance**. The fact that the disconnect between Management ratings and BDEV ratings is not reported takes away an important element of accountability. Additionally, very little information is conveyed to the Board in the RMF with respect to the performance of NSOs. While there is a need to take into account issues of confidentiality, this could be addressed through aggregate reporting.

10. **Accountability processes have placed insufficient focus on quality of monitoring and closure.** The IPR/ASR and the PCR/XSR are not fully considered a decision-making or a learning tool but rather an administrative requirement with inadequate accountability or attribution. Since BDEV doesn't validate IPRs/ASRs, there are few opportunities for contestability of ratings as evidenced by the fact that IPRs are not much discussed nor systematically reviewed by management. Line managers are not systematically held accountable for quality checks at supervision and quality-related KPIs are missing. The role played by the Portfolio Management Unit (PINS) for NSOs is akin to a dual accountability approach and provides some degree of arms-length review and contestability. For public sector projects, the recent addition of the Implementation Support Manager position in each region could help engage on portfolio issues, liaise with the sector divisions, provide the regional and country perspective, including CPO involvement for better convergence with country portfolio management.

---

[7] BDEV Integrated Safeguard System (ISS) – Draft Technical Report on overall compliance of the African Development Bank Group operations with the Integrated Safeguards System across the project cycle - Draft Report -June 2019

11. **A consistent and harmonized framework that allows to assess performance throughout the project life from origination to independent evaluation, is lacking**. This would allow to link the SESP and its tools with the front-end portion of the QA process (quality at entry) and ensure that the same approach and indicators are being used on progress reports as at the end in PCRs and XSRs. This will also make clearer where the lack of candour comes in and allow for "no surprises" by providing predictability on the assessment metrics being used. This also means keeping results framework up-to-date in case of changes to the project environment and, on the NSOs side, including monitoring indicators in line with those adopted by the ADOA framework.

12. **Templates and formats for a number of SESP tools are not sufficiently differentiated and adapted to specific circumstances**. There is room to make the templates more efficient and user-friendly and avoid redundancies especially between IPRs/ASRs and BTORs, as also noted by the QoS evaluation. IPRs are not adapted to special investment vehicles such as Programme-based Operations (PBOs) and TA projects nor to fragility situations. The PCR and PCREN templates are overly repetitious, too long and some sections duplicative, as also evidenced by the Final Synthesis Report of the PCRs' validation (2019). Templates are not designed for optimum management attention and do not focus on priority issues or priority actions needed.

13. **The capacity of the SESP products to report and address specific issues is weak in the areas of safeguards, gender, climate and fragility**. In particular, as reported by the recent ISS evaluation, the Bank's supervision reports and BTOs do not capture the key E&S information to allow compliance checks. E&S information found in the available supervision reports is scanty except when a potential high corporate risk materializes. Other studies conducted by SNSC corroborate the point that once a project is approved, the Bank's internal reporting system offers very little information about the E&S follow-up. The gender dimension of M&E systems is particularly weak with average ratings in the PCRENs below a 2-rating. Project teams don't have enough support in these areas and coaching on SESP requirements to mitigate the effect of the high staff turn-over that the Bank has experienced in the last couple of years. The newly launched Operations Academy should help address this issue.

14. **Completion reports of CSPs/CPPRs and for RISPs are not validated by BDEV before being submitted to the Board**, unlike comparator institutions or similarly to PCRs and XSRs which are validated and submitted to the Board in the form of a synthesis paper. BDEV has carried out validation of one CSPs and one RISP on a pilot basis and is assessing whether there is scope for expanding the process. Management is in the process of starting a review of CPPR methodology and guidelines that should address this point including the articulation between CPPRs and CSPs[8].

### C) Learning

15. **Despite the fact that performance management and accountability aspects of the SESP are weak, they have overshadowed or even undermined learning objectives.** This finding resonates with the situation in comparator institutions. There is a fine line between the search for accountability and learning. Strategy papers (country or thematic) are more conducive to learning, possibly because no ratings are involved. Disagreement over ratings between staff and BDEV further undermines the incentive to promote learning. If the PCR/XSR is seen as a tool to judge the task manager, it will undermine candour in ratings and learning opportunities. The rating itself could become an obstacle to learning because it potentially makes the discussion unnecessarily contentious and personalized.

[8] Information Note to CODE on the Bank's Group's Assessment Methodology of CSP Performance in CSP Completion Reports. December 2018.

16. **There is no single place where SESP products and information can be accessed**. The SESP has not benefited from the availability of a solid repository of knowledge that is mined and shared regularly by staff. This is in the making with the advent of the RRS which will include data from all SESP products. Incentives, combined with new forms of learning and templates may be needed, providing solutions and lessons irrespective of how it may be self- or independently assessed in a single project context.

17. **The quality of PCRs, as measured by BDEV validation process, is variable, but average ratings are low (around 2.8), which hinders learning opportunities**. The PCR quality rating is affected by the same issue as for project performance. Results are good when measured against the 2.5 rating threshold (77% satisfactory in aggregate over 2016 and 2017), as per the current "range" methodology, but modest (52%) if measured against a 3-rating threshold. By and large, the majority of PCRs are prepared by consultants. Most of the task managers interviewed think that they should not do their own PCR for reasons of conflict of interest. However, while using consultants under the supervision of task managers may provide some level of arms-length review, candour remains an issue and consultants do not follow the same standards which raises issues of comparability and quality. PCRs are typically given lower priority by staff and there is little vetting on the choice of consultants.

18. **There is no systematic feedback from the SESP of NSOs that provide success or failure stories based on assessment of achievements in reaching development outcomes and profitability**. Lessons learned are not institutionalized and documented to ensure a strong capitalization of lessons learned. However, BDEV high level evaluations are contributing in capturing lessons and generating knowledge that enable new strategic orientations for NSO and private sector development.

19. **The SESP on its own cannot cater for the array of learning needs of the Bank.** Learning through the SESP is not sufficiently complemented by other sources of knowledge such as impact evaluations, thematic reviews, etc. There is a need for more creativity in terms of dissemination and sharing lessons with other countries, on the same sector or similar specific issues. Periodic workshops or events, as recommended in the PCR-EN guidelines, to be organized with relevant staff to enhance feedback on findings drawn from PCR/XSR-EN, rarely happens. If SESP documents are not sufficiently evidence-based and events are not seen as a safe space where people are willing to learn from success and from failure, incentives to learn will remain insufficient.

**<u>Recommendations</u>** (related Findings in brackets)

1. **Elaborate an overall Bank integrated evaluation policy**. Currently only BDEV has an independent evaluation policy but the Bank as such does not have an integrated evaluation policy that allows to cover both the independent function and the self-evaluation of the Bank itself. An integrated policy would provide a comprehensive governance framework, processes and procedures covering both the Bank and BDEV. This would be an opportunity to define how evaluation contributes to performance, accountability and learning in terms of desired outcomes. It would facilitate the realignment and harmonization of methodologies and ratings definitions, and link agreed outcomes to the Bank corporate results framework. It would also allow to describe the role of Management, the Board and BDEV, including BDEV degree of engagement with others in AfDB. Annex 3 provides a description of what such integrated policy would entail. (F.3)

2. **Reform the ratings methodology** of PCRs by: i) abandoning the current "range approach" and assigning the overall rating through a "judgment" assessment vs. the current practice of simple averages of sub-ratings and dimensions; ii) ) the official project rating should be the PCREN/XSREN rating provided by BDEV through the validation process, including for reporting to the Board in the RMF, and appropriate timelines should be agreed between BDEV and Management to ensure that validated ratings can feed into key reports; iii) self-rating Bank and Borrower's performance in the PCR should be discontinued

and left to BDEV only; iv) the on-going revision of the Operational Manual should ensure that strong and more rigorous procedures and guidance is provided regarding the required evidence for ratings. (F.2; F.4)

Points to be further discussed. Related to the area of reforming the ratings methodology, this evaluation thinks that there is merit to further discuss the following points as additional options for consideration or as part of a learning event.

- *Modify the current 4-point scale*. The analysis of whether other rating scales could be more appropriate and conducive to improving accuracy and candour was inconclusive. Pro and cons were identified with each model, including when looking at comparators, interviews, and ERG comments. The recommendations submitted above are seen to have more direct pay-offs for improving the system, irrespective of the rating scale. Notwithstanding, and if the above measures were adopted, the modification of the 4-point scale may still have merits to address the issues presented in the analysis and in the findings, as in the opinion of almost all the staff interviewed, the rating scale and methodology need revision.
- *Single-source PCR rating*. It could be argued that rating the PCR twice (by staff and by BDEV) is unnecessary and inefficient and only leads to tension around possible differences. The PCR rating could be provided by BDEV only through the PCREN as part of the validation process, as an independent source offering a single official rating for corporate reporting. Staff would complete the PCRs with the narrative only. Such approach has the main advantages of i) simplifying the process, making it less contentious and eliminating the disconnect issue, ii) taking care of the recommendation to stop self-rating Bank and Borrower performance, and iii) improve the PCR learning opportunities. It would require however to ensure i) timely completion of the PCRENs so that they can be ready for the preparation of the ADER report, and ii) a more collaborative process during PCREN preparation between staff and BDEV, whereby alignment between narrative and ratings can be discussed and adequate level of staff accountability be preserved. Such measure would represent a departure from the current practice in other comparator organizations, even if an excessive focus on ratings was already highlighted in the evaluation of the WBG self-evaluation system[9]. Also, an ECG practice note of 2018 (Annex 4) observes that self-evaluation systems may elect not to apply ratings in some instances and should be sufficiently flexible to accommodate management and evaluator judgement.

3. **Put in place an incentive structure conducive to enhance the capacity of the SESP to achieve results.** This recommendation comes as a complement to a number of measures decided by Management as part of the recently approved Quality Assurance Action Plan (2019-2021) to sharpen the focus on delivery and results. Management recognition of the importance not only of good design but also of proactive implementation, will be supported by a programme that rewards excellence and innovation in project design, project implementation and project restructuring - including project and portfolio turnaround successes. It also proposes information campaigns on quality assurance and new tools and templates, including a revision of the Operational Manual to reinforce reporting and compliance.

   In that context it would be important to ensure that task managers and task teams will be supported in the pursuit of achieving results and good quality SESP by providing additional capacity, especially at the regional/country level (for instance as an enhancement of the role of the Implementation Manager) in dealing with problematic/complex situations, implementing M&E systems and better articulated results frameworks, dealing with E&S safeguards, thematic issues (gender, climate) and specific situations (fragility). Simplifying project restructuring procedures is also required to improve the incentive for corrective action. Finally, SESP practices should be part of the staff performance conversations, with a

[9] IEG. Behind the mirror: a report on the Self-evaluation systems of the WBG (2015).

view towards greater recognition of proactive behaviour in dealing with complex, problematic issues and achievement of results. (F.5; F.6; F.7; F.10; F.13)

4. **Improve compliance and candour through enhanced accountability, the usefulness of performance monitoring tools and of corporate reporting**. Measures include: i) incorporating compliance indicators in the delivery dashboards for Management attention such as IPRs/ASRs validation by managers, timeliness of filing reports, implementation of the MTRs; ii) establishing new results-based KPIs and top-level targets to be cascaded down to line managers and task managers for staff performance management and accountability; iii) rolling out the Results Reporting System (RRS) for both public and private sector operations developed by Management for enhanced transparency, efficiency and accountability; iv) validating CSPs/CPPR (subject to a revision by Management of the CPPR methodology) and RISP Completion Reports for reporting to the Board; and v) identifying a relevant indicator on NSOs performance to be aggregated as part of the RMF level 3 indicators. (F.2; F.8, F.9; F.10; F.14)

5. **Redesign/Adapt the template of some of the SESP reporting tools and improve their quality.** The IPR and ASR formats should be updated and simplified to reduce redundancy and overlaps with BTORs, improve efficiency in reporting and making management discussion and vetting a requirement before filing and for compliance monitoring. IPRs should be differentiated for diverse typologies of operations such as PBOs and TA projects, and adjusted for country circumstances (e.g., fragility). ASRs should develop a more precise and actionable assessment of risks to development outcomes, a methodology that rates the progress towards the achievement of development objectives and a clear measurement of progress towards target outcomes and outputs. A number of these measures have now been adopted by Management under the QA Action Plan. PCRs/XSRs templates should be reduced in size and focused on items that require management attention with a view to make learning lessons more relevant and usable. The selection of previously-used consultants for PCR/XSR preparation should be vetted against former PCR quality scores. The on-going revision of the Operational Manual should ensure that the new guidance encompasses the adopted changes. (F.12; F.13; F.17)

6. **Continue to emphasize the learning outcome of the SESP but as a complement to other types of inputs that may be better suited to meet the Bank's demand for knowledge and best practices.** Rather than trying to fit the same instrument (such as the PCR/XSR) to achieve multiple objectives (accountability and learning), it may be more appropriate to develop a distinct learning approach that has the SESP as one of its inputs but relies on more specific and adapted venues, instruments and products, going beyond individual projects. Learning should remain an important outcome of the SESP but it should rely on a repository of information and data, and more dedicated and diversified sources of knowledge such as thematic reviews, impact evaluations, multi-country analysis, etc. Regular learning and "share" events, possibly jointly with BDEV, should be organized by the relevant Departments. Such events should be perceived as a safe space environment with its own rationale, different from what an accountability outcome would require. (F.15 – F.19)

## Table 1: SESP Reporting tools and performance

| SESP reporting tools | | Outcomes | Main issues | How well does it support the related outcome? | Recommendations |
|---|---|---|---|---|---|
| Public projects | IPR | Performance<br><br><br><br>Accountability | Low Compliance, filing delays and candour issues. Redundancy with BTORs and no differentiation for special situations (PBOs, fragility).<br><br>Managers are not sufficiently accountable for IPRs quality and reliability: little validation and discussions. | Weak tool for raising issues, corrective action and performance management.<br>IPRs weaknesses permeates through PCRs, CSPs, CPPRs and RISPs.<br>Insufficient accountability tool | The IPR format should be updated, simplified to reduce redundancy with BTOR, and differentiated for diverse typologies of operations and country circumstances (e.g., fragility).<br>Make Mgt. vetting of IPRs a requirement for compliance monitoring through the dashboard. |
| | MTR | Performance | Compliance is low and MTRs are only occasionally used for project restructuring, which is perceived as having high transaction costs. | The tool is considered adequate but its implementation weak. | Include MTR compliance as a dashboard indicator.<br>Requires higher degree of Mgt. accountability. |
| | PCR | Accountability<br><br><br>Learning | Weak incentive structure for candour. The rating system is not conducive to effective and reliable assessment, leading to disconnect with IDEV validation.<br>Occasionally used as learning tool esp. for follow up projects in the same country. Excessive focus on ratings hampers learning opportunities. Trade-off between accountability and learning. | Equating project performance with staff performance leads to candour issues and reduces accountability<br>Learning potential is unmet as there are too few opportunities for more structured lessons-sharing events | Reform the rating system to allow to better capture the reality and improve the guidelines for generating more solid evidence<br>Develop a distinct approach towards learning with the SESP as one of its inputs but relying on more adapted venues and products, a repository of information and a safe space environment. |
| NSOs | PSR | Performance<br><br>Accountability | Low compliance and redundancy with ASRs & BTORs.<br><br>Little discussion or validation by Mgt. | Weak tool for performance management.<br><br>Low compliance reflects low accountability. | Rationalize its use, format and content relative to the other reporting tools (ASRs, BTORs).<br>Make reporting on compliance more visible. |
| | ASR | Performance<br><br>Accountability | Low compliance and redundant with BTOR. Lack of consistency in using common assessment criteria and rating systems between the various documents. DO ratings are not systematically validated. | Weak assessment of risks to development outcomes.<br><br>Low compliance reflects low accountability. | Develop a more precise and actionable assessment of risks to development outcomes.<br>Improve accountability through better and harmonized results reporting. |
| | XSR | Accountability<br><br>Learning | Lenient ratings in the XSR with considerable validation backlog. Compliance issue. Projects assessed against indicators not included in the PAR or the ADOA.<br>Few learning opportunities. | Low compliance and candour affect accountability and efficiency of reporting.<br>Financial performance dominates over development outcomes. | Better alignment of criteria and harmonization of rating systems among the various reporting tools.<br><br>Make lessons more relevant and useful. |
| CSPs/CPPRs | CR | Accountability<br><br><br><br>Learning | The quality of CSPs/CPPRs is impacted by the quality of IPRs. CSPs/CPPRs completion reports are not validated by BDEV. The weight of CPPRs in assessing the quality of CSPs is questionable.<br>CSPs/CPPRs are reviewed by the Board and are more likely to be subject to discussion and learning. | Lines of accountability between the CSPs and the CPPRs are blurred.<br>BDEV carries out independent evaluations which facilitates accountability and learning. The current focus on the narrative, rather than on ratings, facilitates learning. | CPPR methodology and articulation with CSPs need revision.<br>Consider validations of more CSPs/CPPRs by BDEV. |
| RISPs | CR | Accountability<br><br>Learning | The quality of RISPs is impacted by the quality of IPRs. RISP Completion Reports are not validated by BDEV. RISPs are reviewed by the Board and are more likely to | BDEV carries out independent evaluations of RISPs which facilitates accountability and learning even if methodologies are not | Consider more validations of RISP completion reports by BDEV |

| | | | | | |
|---|---|---|---|---|---|
| | | | be subject to discussion and learning. | harmonized with CRs | |

## I. Background and Context

1. Following adoption of the Ten-Year Strategy (TYS), 2013-2022, the African Development Bank (AfDB or the Bank) has gone through the following major organizational restructuring, policy changes, and operational and institutional adjustments:

- Adoption in 2015 of the High5 priorities[10] within the context of the TYS, leading to the development of appropriate strategies for each of the High5s;
- Adoption of the New Development and Business Delivery Model (DBDM) in support of the High5s; and an enhanced drive towards decentralization, including the creation of regional hubs;
- Creation of structures such as Delivery Accountability and Process Efficiency Committee (DAPEC), and Technical Quality Assurance Committee (TQAC) for improving the operational processes.

2. The on-going institutional changes, combined with the ADF-14 commitments to improve the quality of project design and supervision and the performance of operations, provide an opportunity to take stock of Quality Assurance processes and examine how these processes can be optimized to promote efficiency, manage risks, and increase the likelihood of achieving sustainable development results.

3. BDEV has carried out an Evaluation of the Bank's Quality Assurance processes, complementing the work undertaken by the Transformation Management Team (TMT) and DAPEC with the ultimate objective to increase the extent to which the quality assurance chain contributes to learning and development impact. Three standalone evaluations have now been completed:

- quality at entry (QaE);
- quality at supervision and exit (QoS);
- the self-evaluation system and processes (SESP) (this evaluation).

4. Based on the conceptual synergies across these evaluations, BDEV delivered in October 2018 a "chapeau" report that examines the relevance, efficiency, effectiveness, and institutionalization of the Bank's Quality Assurance processes across the project cycle, which also included the assessment of the compliance with the Bank's E&S Safeguards requirements[11].

5. In addition, BDEV's recent evaluations also include:

- Independent audit of Bank results monitoring and reporting (RMR), which focuses on adequacy and compliance of Bank policies, procedures, organizational arrangements, monitoring and reporting frameworks, and data management (on-going);
- Independent Evaluation of the Integrated Safeguards System (2019);
- Quality retrospective report: an assessment of the Bank's quality assurance tools (2018);
- Comprehensive Evaluation of the Development Results of the AfDB Group (2016);
- PCR validation of the 88 projects that closed in 2017 and the 49 projects that closed in 2016;
- Independent evaluation of the Bank's additionality and development outcomes assessment (ADOA) framework (2014).

6. Based on existing evidence from evaluations and Management own diagnostic study (2018), Management prepared a QA Implementation Action Plan (June 2019), which sets the basis for the reform of the SESP, building on this evaluation.

---

[10] The High5s are: i) light up and power Africa, ii) feed Africa, iii) industrialise Africa, iv) integrate Africa and v) improve the quality of life for the people of Africa

[11] Evaluation of Quality Assurance across the project cycle of the AfDB (2012-2017) (2018)

## II. Purpose, Objectives and Scope

### A. Purpose

7. The definition and overview of the Self-evaluation System and Processes (SESP) are summarized in Box 1. Evaluation of the SESP builds upon the sequence of evaluations already carried out, leveraging evaluative evidence regarding supervision and quality at exit to examine how self-evaluation processes are implemented and can lead to better achievement of results.

8. The Bank implements both independent evaluations, which mandate rests with BDEV, and self-evaluation systems and processes, which resides with Bank's management. The two are mutually dependent. These systems and processes help the Bank to account for its investment effort, assess the quality of its portfolio, improve its performance, meet reporting requirements, learn from operational experiences, and make progress towards better demonstration of results.

**Box 1: Definition and overview**

- For the purpose of this report, self-evaluation is defined as the assessment made of a project, country/regional program, and policy, sector or thematic reviews by the entity engaged in the activity.
- As described in the Theory of Change, the system is meant to support three main outcomes:
  - Performance
  - Accountability
  - Learning
- The main dimensions being assessed are:
  - Relevance
  - Effectiveness
  - Efficiency
- The main tools being used are:
  Progress reports, mid-term reviews, completion reports.
- Self-evaluation is accompanied by validation of certain products by BDEV such as the PCRs/XSRs or independent evaluation of CSPs and RISPs.

9. The SESP are defined in different Bank documents including:
- The Operations Manual (OM), which was initially adopted in 1993, revised in 1999, and more recently in 2014. The next revision of the OM is on-going;
- Delegation of Authority Matrix (DAM) and relevant Presidential Directives (PDs);
- Additionality and development outcomes assessment (ADOA) framework for the Bank's non-sovereign operations; and
- The 4-level Results Measurement Framework (RMF).

10. The purpose of this evaluation is to support Bank management and operational staff through its findings and recommendations in:
- Improving self-evaluation and performance of operations, and country/regional strategies;
- Improving the relevance and quality of the Bank's Operations Manual to improve the SESP function;
- Enhancing operational effectiveness and the tools for accountability and learning;
- Supporting the implementation of the new DBDM, and process engineering;
- Informing the Board and other stakeholders on the functioning of the SESP and ways to improve it.

### B. Objectives

11. The main objectives of the SESP evaluation are:
- Assess how well the Bank's SESP perform, focusing on their relevance, effectiveness and efficiency, in serving three main outcomes—improving performance, enhancing accountability, and promoting learning;
- Identify and assess the enablers and barriers that affect the implementation and results of the SESP;
- Distil lessons and good practices, and formulate recommendations to enable the Bank to enhance the quality and performance (design, scope, implementation and results) of its SESP.

12. Key issues addressed by the evaluation include: i) the enabling environment for self-evaluation, ii) data collection tools and systems used to implement self-evaluation; and iii) the use of self-evaluation information for decision-making. The evaluation is forward looking and offers Management a number of recommendations that can enhance the performance of the tools, methods, indicators, processes and incentives that are most likely to establish trust in the SESP and the credibility of their results.

## C. Scope

13. This evaluation examines how the SESP and the ratings system are being implemented and applied to the following main outputs during the period going from approval to completion (exit).

- Public sector projects: Implementation Progress Report (IPR), Mid-Term Review (MTR), Project Completion Report (PCR);
- Non-sovereign operations (NSOs): Project Status Report (PSR), Annual Supervision Report (ASR), Expanded Supervision Report (XSR);
- Country Programs and Strategies: Country Portfolios Performance Reviews (CPPRs), Country Strategy and Program Completion Report (CSP-CR), and Regional Integrated Strategy and Program Completion Reports (RISP-CR).

14. An important distinction can be made between the mandatory self-evaluation products listed above and voluntary evaluation studies such as impact evaluations and occasional programmatic evaluations or retrospective studies commissioned by individual business units. PCRs and XSRs are independently validated by BDEV which also carried out, on a pilot basis, the validation of one CSP (South Africa) and one RISP (West Africa). BDEV also carries out on a regular basis independent evaluation of projects, CSPs (CSPE) and RISPs (RISPE) on the basis of potential need for evaluative information (e.g. revision of Policy or Strategy when one expires). Validation and independent evaluation by BDEV are important complements for the effectiveness of the SESP but are not covered as such by this evaluation.

15. Key indicators of the SESP information are aggregated into apex corporate reports and scorecards for corporate accountability. These include:

- The Results Monitoring Framework (RMF) and the associated Annual Development Effectiveness Review (ADER);
- Portfolio monitoring reports; and
- Reporting to the Board on progress in implementing strategies.

16. The evaluation covers the period 2013 – 2018, which represents a considerable part of the implementation of the TYS, the adoption of the High 5s strategies, as well as the DBDM and process reengineering reforms. This period encompasses the issuing of the updated Operations Manual in 2014; findings from the evaluation would inform the upcoming 2019 revision. This evaluation does not cover personnel, Board operations, control functions and Treasury operations.

17. Bank staff and managers as well as members of the Bank's Board are the primary intended audience of this evaluation. The evaluation, which is also expected to be of interest to comparator organizations, includes a comparative analysis of practices at the Asian Development Bank, the International Fund for Agriculture Development, and the World Bank Group.

18. Limitations to the evaluation relate to the availability of SESP products and documents, including monitoring reports and ratings, as filed by the task managers in the Bank's system. This has been especially the case for NSO data. Data on the SESP cost was scanty and could only be inquired through individual interviews as no centralized information was available, thus constraining the cost-effectiveness analysis. Also, the low response rate to the staff survey (6%) has limited its representativeness and the results were only considered on an indicative basis to be triangulated with other sources.

# III. Methodology

## A. Theory of Change (ToC)

19. The Bank's evaluation policy, the OECD-DAC criteria and the Evaluation Cooperation Group's Big Book on Evaluation Good Practice Standards guided the evaluation, which is based on a theory of change presented schematically in Figure 1. The theory of change underpinning the self-evaluation architecture is based on the fundamental logic that a well-functioning SESP can improve:

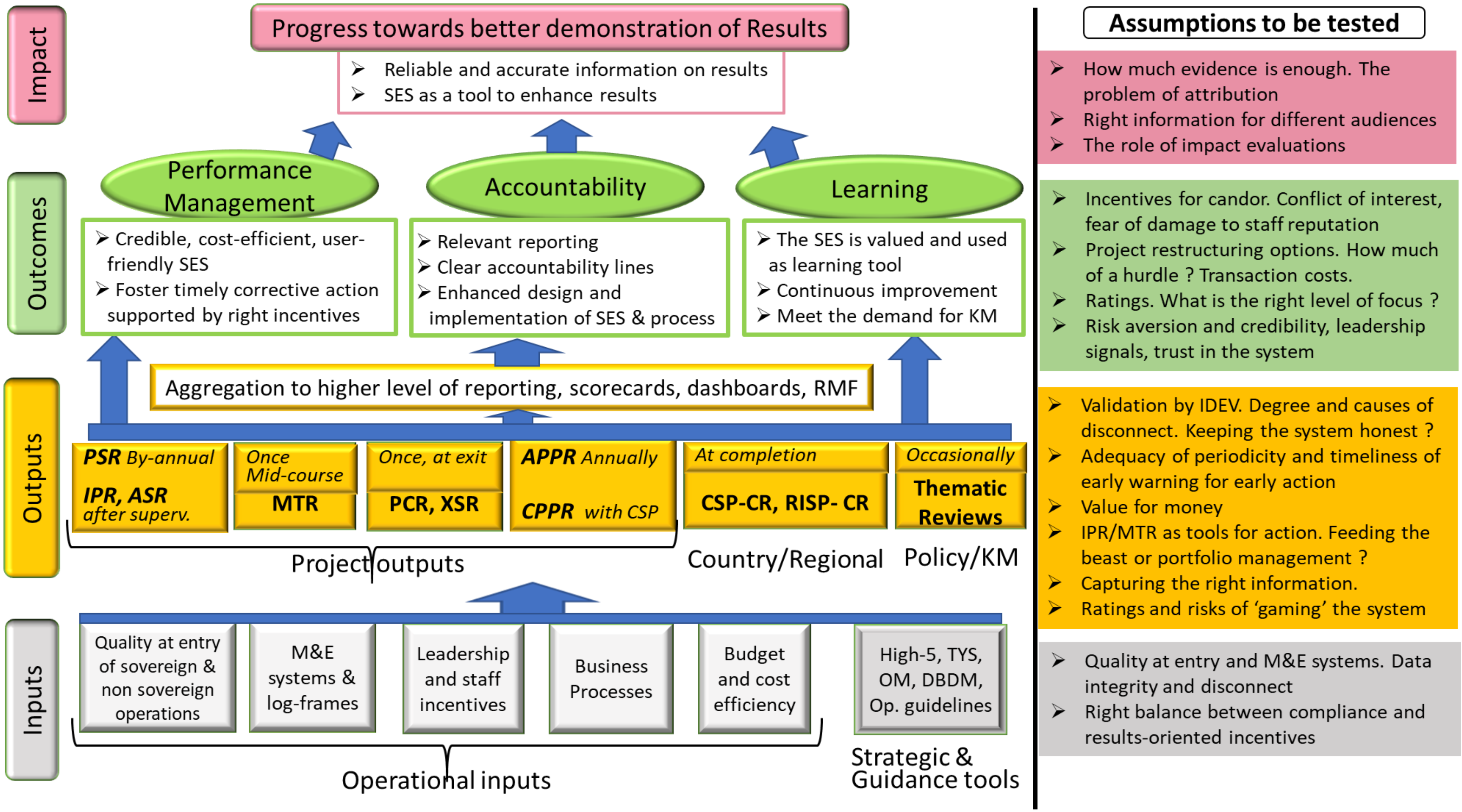
- performance management and how the availability of reliable information and evidence can help management take timely decisions;
- accountability and how the provision of key information at different levels (project, program, corporate) signals that the AfDB holds itself accountable for achieving results; and
- learning and how the SESP can be a tool for sharing lessons and continuous adaptation.

20. The Evaluation examined the causal pathways going from the inputs into the SESP (the portfolio at entry, the M&E systems, the business processes, the leadership signals and incentive structure, and the various guidance documents), and how they influence the achievement of outputs, outcomes and impact. It also examined the links between inputs and outcomes that are ensured through the production of a number of reports (outputs) during project supervision and at closing. These reports feed into broader reporting arrangements at corporate level. Other links between the self-evaluation systems and other systems were also assessed to determine how they influence the overall response culture and the incentive structure, such as project logframes; the articulation with BDEV's own independent evaluations and 'validation' exercises (e.g., PCREN, XSREN); the commitments made at the corporate level; the Operations Manual; and other requirements.

21. The interface between the various systems, gaps in coverage, overlaps, relevance, periodicity, and the overall supporting environment is analysed, also building on data from the QoS report. In examining the various causal pathways, a number of assumptions were tested to probe the robustness and credibility of the system, and to identify the weak links that could lead to recommendations for improvements. The key assumptions for the different levels of causality in the ToC cover the following:

- the effectiveness of the enabling environment and barriers to self-evaluation;
- the prevailing incentive structure and how it influences individual behaviours;
- the balance between compliance and the achievement of results;
- the effectiveness of what is produced compared to its added value;
- the adequate production, use and relevance of the project rating system; and
- transaction costs of project restructuring and aversion to risks undermining corrective action.

**Figure 1: Self-Evaluation System and Processes – Theory of Change**



## B. Evaluation Questions

22. The overarching question addressed by the evaluation is:
"Do the self-evaluation systems and processes (SESP) support Performance Management, Accountability, and Learning at the Bank"? The two underlying sub-questions are:

1. How well are the SESP performing?
2. To what extent are the SESP impacting on the achievement of quality development results?

Questions/sub-questions are organized in the Evaluation Matrix (Annex 2) as summarized in Table 2.

**Table 2: Summary Evaluation Matrix**

| | Performance Management | Accountability | Learning |
|---|---|---|---|
| Relevance | Are the SESP tools and processes relevant to improve performance of project/country programs? | Do the SESP generate relevant and credible information signalling that the Bank is holding itself accountable for achieving results? | Are the SESP being used as a reliable and relevant framework for learning and innovation? |
| Efficiency | Do the SESP provide a reliable, timely and efficient framework for portfolio management? | Do the SESP provide a reliable and efficient framework for reporting and accountability internally and externally? | Are the SESP being implemented as an efficient tool for learning? |
| Effectiveness & Impact Contribution | Is the SESP architecture being implemented as a tool to enhance performance and the achievement of results? | Is the degree of accountability exerted on the implementation of the SESP conducive to achieving results? | Have the SESP contributed to the identification and use of lesson learned? |
| Incentives & barriers | Are the incentives in place conducive to candid assessments and proactivity for corrective action? | Are the incentives in place conducive to exerting the right degree of accountability for the implementation of the SESP? | Is the incentive structure geared towards use of the SESP for continuous learning and innovation? |

## C. Evaluation Methods

23. This evaluation does not duplicate previous efforts, but builds on the existing base of evidence to focus more specifically on the performance of the SESP system itself. It builds on the relevant data and evidence already collected, while filling the gaps vis-à-vis new data requirements specific to the SESP. The evaluation has paired effectively and supplemented the QoS evaluation. Unlike the QoS evaluation which has looked at the various components of project supervision including from the Borrower's perspective, the SESP evaluation has focused on the internal processes, instruments and mechanisms, with an overall objective of assessing how the application of the SESP impacts on performance, accountability and learning. While project performance relies in good part on the performance of country Governments, partners and local stakeholders, the SESP is entirely under the Bank's control as well as in its capacity to adopt and implement recommendations.

24. A relatively small sample of closed projects was selected as part of the case studies (see below) in order to probe in more depth the actual application of the SESP during implementation of projects, CSPs and RISPs. Sampling was not intended to establish a base for inference on the rest of the portfolio or to be used as a prediction tool, but to complement already available data (mostly from the QoS and the QA Chapeau paper), and provide better insight on SESP implementation issues from the perspective of the respective task managers. The evaluation has also built its analysis on two sets of PCR validations carried out by BDEV through the PCR Evaluation Notes (PCRENs) for: i) 49 public sector projects closed in 2016, and ii) 88 public sector projects closed in 2017[12]. The evaluation has followed a mixed method approach and

[12] BDEV. Synthesis Report on the validation of the 2016 and 2017 Project Completion Reports (2019).

relied on diverse methodological approaches to answer particular evaluation questions. Data collection methods have targeted multiple sources to enable triangulation of information.

### 1. Meta-analysis

25. This comprises literature and desk reviews of evaluations of self-evaluation systems conducted by other MDBs, during which the team has assessed the experiences of organizations with similar SESP and examined how these institutions assess their SESP with respect to three pillars: Performance management, Accountability, and Learning (individual and corporate). It examined the common issues across the MDBs including factors affecting the three pillars.

### 2. Benchmarking

26. Benchmarking was conducted in parallel with the meta-analysis and compared various components of the self-evaluation system of comparator/sister organizations. It includes the key elements of each self-evaluation system, coverage, processes, and discussions with key stakeholders. It culls lessons of experience and good-practice. For the purpose of this analysis, the assessment examined practices at the Asian Development Bank, IFAD, and the World Bank Group.

### 3. Case Studies

27. Three case studies were developed around a sample of projects (both public and private), CSPs and RISPs. Project analysis is included directly in the main report with separate paragraphs dealing with NSOs and CSPs/RISPs. Annex 7 covers the analysis of the CSPs and RISPs case studies.

**Projects**

Public sector projects: a purposive sample of 12 closed projects was extracted from the pool of projects validated by BDEV in 2018. Selection was randomly based while ensuring that all sectors be represented and not more than one project in the same country.

NSOs: a sample of 5 operations was purposively selected jointly by BDEV and PINS from the batch of XSRs received by BDEV. Sampling covered the financial sector (line of credit, senior debt) and private equity in the power and energy, industry, infrastructure and multisector. As no XSRENs were available, the XSRs were reviewed externally by a non-member of the SESP evaluation team to assess the congruence of the assessment and the ratings.

Project documentation[13] was reviewed and the respective task managers interviewed. The projects were followed through their life cycle to evaluate the consistency and sequencing of the various reports and actions. The purpose was to review the effectiveness of the SESP with respect to: i) implementation of corrective actions, ii) application of existing reporting requirements, quality of the results framework, underlying assumptions, iii) clarity of results monitoring and M&E system, iv) relevance of project reports and candour in the use of ratings, and v) quality of the SESP tools and the use of IPR, PCR/XSR data (through the use of the 2017 BDEV Synthesis Report). Templates were used for better comparability of results. Detailed interview summaries for both public sector projects and NSOs are available as Working Papers.

**CSPs**

The CSPs for Côte d'Ivoire, Morocco and Tunisia were examined (inception to MTR and CR) in tandem with the respective CPPR. CSP selection criteria included the availability of complete SESP products through MTRs, Completion Reports, CPPRs, and CSP evaluations. The evaluation looked at the coherence and usefulness of the SESP with respect to: i) the clarity of strategic objectives and

[13] For public sector projects: PAR, IPR, MTR, CSP/CPPR, PCR, PCREN. For NSOs: PAR, ADOA, PSR, ASR, XSR.

underlying theory of change/results framework, policy dialogue, risk assessment process, consultation with RMC/ partners on strategic issues including cross-cutting issues, ii) clarity of results monitoring indicators and M&E system (data collection, data analysis and use), frequency and timeliness, iii) adequate production, relevance and candour of the CSP-MTR and CSP-CR and use for learning purposes, and iv) use of self-evaluation results in independent evaluation of the CSP by BDEV for the same countries. The task managers of the selected CSPs were interviewed as well as the staff associated with the CPPR as applicable. Interview templates were used for better structuring of the analysis and comparability.

**RISPs**

The East-Africa, Central-Africa, and the West-Africa RISPs were used to analyse the application and the effectiveness of the SESP for a regional integration strategy paper associated to a multi-country environment including a portfolio of regional projects. The main criteria of selection were the existence of a recent RISPE. The task managers of the selected RISPs were interviewed as well as other staff associated with them. Templates were used for better structuring of the analysis and comparability.

## 4. Interviews

28. Semi-structured interviews were carried out with 83 people including members of the offices of Executive Directors, Bank managers and staff as practitioners and resource persons knowledgeable about the SESP (Table 3). BDEV ensured that the people interviewed represented well-informed views of specific relevant divisions in the Bank including self-evaluation authors, portfolio management staff, operational quality, country/regional staff, the relevant Operational Complex and Executive Directors' offices as part of the analysis of external accountability. Interviews were structured around the key points of Table 4 and summarized in Annex 6.

**Table 3: Key Informant Interviews**

| | Abidjan | COs | TOTAL | Managers |
|---|---|---|---|---|
| **EDs' offices** | 9 | | **9** | |
| **SNVP** | 10 | | **10** | 4 |
| **RDVP** | 16 | 10 | **26** | 9 |
| **ECVP** | 4 | | **4** | 1 |
| **PEVP** | 1 | | **1** | 1 |
| **Audit** | 1 | | **1** | 1 |
| **Task Managers** | 4 | 8 | **12** | |
| **CPOs** | 1 | 4 | **5** | |
| **Private sector** | 11 | 1 | **12** | 3 |
| **Invest./portfolio officers** | 1 | 2 | **3** | |
| **TOTAL** | 58 | 25 | **83** | 19 |
| **BDEV** | 4 | | **4** | 1 |

**Table 4: Illustrative areas of enquiry for the interviews**

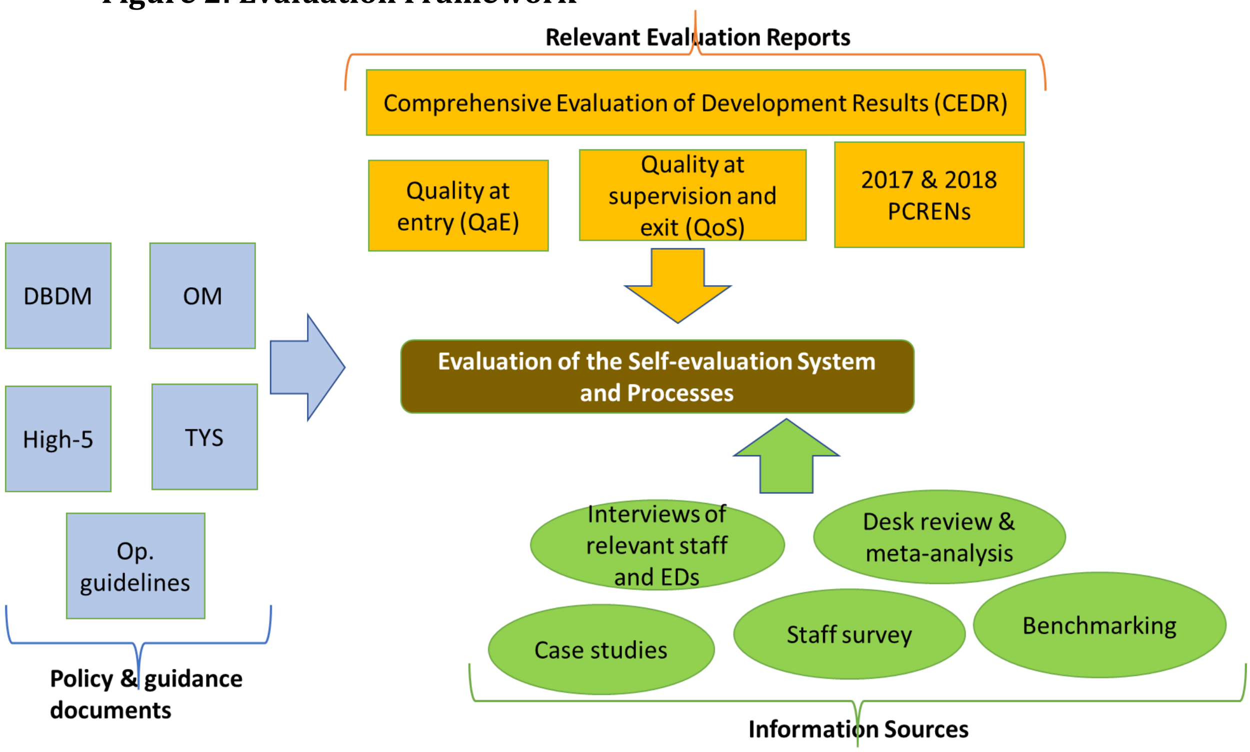
| Performance Management | Accountability | Learning |
|---|---|---|
| • Main lessons derived from the application of the guidelines with respect to production and use of the SESP documents and ratings.<br>• Existing incentives or disincentives around candour, objectivity and relevance.<br>• Use of the SESP as an alert system to raise issues and take corrective action for improved results.<br>• Main reasons for the source of disconnect between BDEV validation and the findings brought forth by the application of the SESP. | • The nature of leadership signals during the period of implementation of the task and the degree of accountability exerted by Management or the Board.<br>• Nature of the discussions on the application of the SESP and portfolio performance with Management.<br>• Issues around compliance-check vs proactive follow-up on problems identified during supervision.<br>• The validity of the RMF as a tool to distil the main outcomes and lessons of the SESP for external reporting. | • The extent to which the main outcomes and lessons learned (PCRs and PCRENs) were used for the preparation of follow-up projects, portfolio reviews, and thematic documents.<br>• The nature of the incentives required for a more effective learning through the use of the SESP mechanisms.<br>• The possibility that an excessive focus on ratings may undermine the learning potential. |

## 5. Staff Survey

29. A survey was circulated among staff who are directly or indirectly involved in the production or utilization of the information from the SESP. The survey has gathered information on the various components of the system, familiarity with issues related to compliance and quality, understanding and use of learning tools, incentives and various accountability mechanisms. Because of the low response rate (6%) with only 27 respondents, the survey was not sufficiently representative to allow rigorous extrapolation and results were triangulated with other sources (QoS staff survey, staff interviews, desk reviews). Annex 8 presents a summary of the results together with the survey template.

30. **Evaluation Framework and Data Analysis**: The evaluation framework was structured around three main elements that represent the source of the information as illustrated in Figure 2: i) the Bank's policy and guidance documents, ii) the main relevant evaluation reports (by BDEV), and iii) the documents and data sources highlighted above as well as part of the benchmarking exercise (Annex5). These formed the basis for the generation of findings, and drawing of conclusions and recommendations.

**Figure 2: Evaluation Framework**




31. **Potential biases and conflict of interest** that could be perceived in the role of BDEV have been addressed by relying on external consultants only motivated by their professional commitment to assess systems, evolutions, quality of work, effectiveness of methodologies and outcomes. They carried out most of the analysis, individual interviews, facilitation of group meetings, and responding to questions and comments. The evaluation examined self-evaluation by operational staff and is not an evaluation of BDEV validation or independent evaluation. Findings were the results of factual analysis, triangulation with multiple sources, experience with comparator institutions, own judgement, and do not represent self-serving positions by individuals. BDEV is conscious that its role might shape incentives and wants to mitigate any concern of objectivity possibly deriving from its interface with self-evaluation.

# IV. The Underlying Factors of Self-Evaluation

## A. The self-evaluation rationale

33. Few skills are more important to improving one's work than being able to step back and candidly evaluate yourself. Self-evaluation allows the raising of issues before they become too obvious to ignore and too late to correct. By human nature, people tend to be self-complacent about what they do, especially if the predominant culture in the work environment does not exercise a minimum level of accountability based on checks and balances. If the chances are that complacency will predominate, the feedback received on your own work could be warped and the self-evaluation distorted towards over-confidence.

34. Ignoring issues means they will not be fixed. This could only work until external parties start looking at outcomes and impact in a more thorough way, questioning the credibility of the system and making it clear that issues stand in the way of achieving results and that methods need to be found to identify them and bring them to light. Self-evaluation is often confronted with deep-rooted practices that are embedded in the "culture" of the institution. This could potentially change but would need to go through a revision of the predominant incentive structure.

35. Similarly, distorting self-evaluation towards negativity, to be on the safe side, may also not allow to take full advantage of the ability to achieve results with the instruments at hand and undermines the chances of proactive action. Candour does not mean having to be negative, but rather maintaining a dose of critical mindset towards oneself and objectively accept findings the way they are. As having to say something negative about your work can be uncomfortable, management signals need to be clear, in recognizing that project performance should not be equated with staff performance to the extent that things can go wrong for many reasons. Staff should be recognized for their willingness to raise issues as the best path towards taking corrective action and achieving results.

36. The Evaluation Cooperation Group (ECG), issued in November 2018 a Practice Note: "Self-evaluation in ECG member institutions" (Annex 4) noting that if self-evaluation is to be useful it must be an integral element of and used by wider "results" systems. To this end self-evaluations should: i) meet clear standards for quality, relevance and timeliness; ii) be clear and well-integrated with the overall evaluation policy to ensure management ownership and responsibilities for systems and processes; iii) be visible and widely accessible; and iv) rating systems must be rigorous and consistent but their application flexible.

37. Because it is difficult to gain such a level of distance from staff own work, a number of institutions have considered self-evaluation not to be a robust and reliable evaluation tool, although useful to support management in taking informed decisions. Most of bilateral aid organization and UN agencies do not carry out self-evaluations by the project teams themselves, and often do not use ratings that can be aggregated across time, sectors, and regions. They prefer to use consultants on an arms-length basis to ensure independence, but not in a systematic way.

38. As mentioned in the ECG Note, where self-evaluation is used as a primary evaluation tool, the main advantages are cost-effectiveness but the rigor in the validation of findings needs to be strong. An alternative to self-evaluation could be to exercise external or independent controls and mitigate the risk of conflict of interest. This would raise the level of trust in the system but would have important cost implications and would remove from the analysis the originator, the person who knows most about the product being evaluated and whose direct input would be most valuable. Most of the MDBs have tended to keep a combination of internal self-evaluation and arms-length/independent validation as a way to keep the system honest (checks and balances).

39. Two key building blocks stand out in assessing the functioning of self-evaluation systems: i) compliance with mandated reporting requirements to ensure that everyone follows some basic common procedures and metrics so that the various products are consistent and comparable, and ii) candour, to ensure that the outputs are credible and can be trusted as a basis for improving the performance of the portfolio, a driver of accountability and a key motivation for learning.

## B. Compliance

***Public Sector Projects***

40. Observed compliance with project reporting requirements in AfDB is very variable, pointing to the fact that there is little consistency in the way procedures and processes are applied as well as the level of enforcement and accountability exerted. Lack of compliance is an indication of how seriously staff and management are taking the implementation of the SESP. In the absence of a minimum of reliable information, the SESP loses its credibility. This also has consequences on the quality of reporting back to the Board through the RMF and, in the Management Dashboard, on key performance indicators.

41. If compliance is weak and fails to generate the required reports and outputs, the SESP loses its foundation. Also, if compliance is treated in isolation from the overall objectives of the SESP, it becomes a "feeding the beast" exercise whose main objective is to ensure that the right boxes are ticked and no flags are raised by the system. For compliance to be effective, accountability mechanisms need to work in two ways. First, they should ensure the actual production of the required outputs according to the established requirements, and second, whatever is produced should enhance the Bank's capacity to improve performance, ensure reliable reporting arrangements, and promote learning.

42. IPRs compliance is particularly lacking. This is corroborated by the QoS findings (Box 2). It should be noted that the IPR and PCR system was introduced in tandem and became mandatory as of January 1st 2013 for all public sector projects approved since January 1st 2011. For projects approved before 2011 the use of IPRs was optional and ratings could continue to be entered in SAP as before.

43. As reported by the SNOQ's Quality Retrospective Report (2018) there is a discrepancy between actual ratings and ratings based on the appropriate application of the methodology. The IPR coverage is reported to be at 64%, but only 32% of IPRs and 37% of the PCRs follow the correct rating methodology, pointing to low confidence in project ratings at exit. The following shortcomings were encountered: i) 41% of the IPRs reviewed have a missing rating; ii) in many cases the justification for the rating is missing, and 15% of the reports provide justifications that are contradictory to the ratings themselves; iii) 30% of the projects had missing baselines which can lead to biased results reporting; iv) just 64% of the IPRs were uploaded on time and 59% were cleared by management on time; and v) line managers do not verify the quality of the IPRs, as only 3% of the IPRs had comments and were signed off by the manager and only 2% were sent back for revision.

44. Compliance issues were further examined by this evaluation through the additional sampling of 12 public sector projects and 5 NSOs, including interviews with the respective Task Managers and Investment Officers. Of the 12 public sector projects reviewed, 4 projects were approved after 2011, had no IPRs, and were not rated in SAP either. Even for some projects approved after 2012, IPRs and associated ratings were not available. This has prevented a comprehensive analysis and understanding of the extent to which the rating of projects has been a relevant tool to raise issues, and to drive follow up and corrective actions.

45. IPR production requirements are described in some detail in the Operations Manual of 2014. Weaknesses in the timely submission of IPRs seem to be partly related to some redundancy and overlapping content between IPRs, Aide Memoires and BTORs (this point was already noted by the QoS

report). According to staff interviews, the main addition that IPRs bring in terms of content is the ratings, which are not featured in the other documents. Typically, staff consider the Aide Memoire to be the key document between them and the Borrower, and the BTOR as the key document between them and Management. If IPRs are of little use and are not filed, it also means that the capacity of the ratings to be a key tool for performance management is severely undermined as they cannot be acted upon.

46. MTRs can serve the purpose of mid-course correction in a more dedicated and targeted way than IPRs. Compliance with MTR requirements is described in the Operations Manual and most of the time is also included in the project appraisal reports and in the legal documents, which are stronger enforcement tools. MTRs are the main trigger point for entering into a discussion with the Borrower and Management on restructuring or corrective action. In principle MTRs are mandatory but in practice Task Managers decide on whether it is needed or not. Of the 52 projects in the QoS evaluation sample that were required to carry out an MTR, only 19 (37%) were available. Of the 12 projects reviewed as part of this evaluation's sample, only 2 carried out an MTR. Meeting compliance requirements depends, among other things, on the level of attention and priority demonstrated by Management. If accountability is weak, staff are more likely to ignore the rules that suit them least.

47. In the case of PCRs, Management own assessment indicates a compliance rate around 85% - 90% over the 2015-2017 period. Good compliance is higher in this case because BDEV carries out a validation of all PCRs, provides its own ratings in the PCR Evaluation Notes (PCREN), and sends a synthesis report to the Board for information. Timely PCR preparation is also a KPI in the Management dashboard. Compliance was also analysed in some detail by the recent QoS evaluation and the major findings that are relevant for the SESP are reported in Box 2.

**Box 2: QoS Evaluation findings on compliance (from 83 public sector projects)**

- 48% of the public sector projects reviewed had a launching mission.
- 26% of the projects had no supervision mission in 2017.
- 100% of problematic public sector projects were supervised at least once/year in 2017.
- IPRs were submitted with a delay averaging 25 days above the recommended maximum of 30 days.
- Completion rate for IPRs was not available; however, 358 IPRs were retrieved from 83 projects over 6 years. This compares with about 1000 IPRs that should be available if projects were supervised twice a year.
- According to quantitative analysis, 60% of the IPR sample was assessed and approved by the sector manager.
- 15% of the projects reviewed adhered to project classification methodology and 28% had credible IPR ratings.
- Of the 83 projects reviewed, 52 were eligible for an MTR of which 19 were made available.
- 44% of PCRs due in 2017 were submitted with delays.

***Non-Sovereign Operations***

48. In the case of NSOs, assessing compliance is more problematic as little information is available in annual portfolio reviews and reporting on the NSOs portfolio, which is not part of a corporate quality assurance and results system. Furthermore, since reporting on NSOs is not a KPI, little is reported to the Board as part of the RMF (level 3) or independently. The program of supervision is defined each year, depending on the type of NSO (Corporate loans, project finance loans, financial institutions, direct investments, private equity investments etc.). Supervision missions are conducted by obligors (debtors), meaning that 1 supervision mission can cover 1, 2 or more operations. The 2018 Supervision program included 151 planned missions out of a portfolio of roughly 300 operations.

49. The 2018 annual portfolio review indicates a compliance rate of BTORs/ASRs and PSRs at respectively 84% and 43% of the active portfolio, and XSR delivery (achieved vs planned) at 58%. To be noted that PINS reports an aggregate number for BTORs and ASRs as if they were equivalent SESP products.

However, the half-year 2018 report indicates that ASRs were available for 17% of the active portfolio and XSR delivery was at 20%. Although in principle BDEV should validate all XSRs made available, in reality there is a considerable backlog as the last XSREN was produced in 2014.

***Other factors***

50. Some additional key factors need to be addressed in understanding compliance issues. With many new staff coming on board in the last couple of years, the Bank has lost part of its institutional memory and new task managers are not equipped with all the tools and knowledge that come with experience. The staff survey carried out as part of the QoS evaluation reported that only 44% of the staff received sufficient support and training for supervision. Many staff simply do not know how the SESP and the various procedures and practices work. Management's decision to initiate an Operations Academy could go a long way in addressing this issue.

51. Workloads are unevenly distributed across task managers, which may explain the variability of attention paid to SESP. Some 30% of the task managers responding to the QoS survey declared handling more than 5 projects per year with some handling up to 10. This ratio has been increasing over the years and task managers' overload represents a serious risk to quality of supervision. Further, preparing new operations receives a higher priority from staff compared to supervision.

## C. Candour

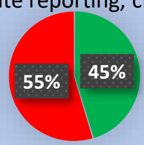
52. The difficulty to exert candour in self-evaluation is an issue that is often admitted by staff themselves and typically attributed to the lack of a proper incentive structure. Fortunately, incentive structures can be improved so as to send the correct signals that will influence staff behaviour in the right direction.

53. In self-evaluation, candour is applied to: i) the narrative, which includes staff capacity to raise and formulate the issues in a way that is conducive to an objective discussion; and ii) the ratings, which are quantitative and tend to lock the activity in a category at a particular point in time. For self-evaluation purposes, narratives are mostly used for strategy papers (RISPs, CSPs) and thematic reviews, while ratings are mostly used for projects (IPR/PSR and PCR/XSR).

54. Disagreements over ratings tend to be more difficult to resolve than over narratives, since by nature ratings offer little alternative choices and less room for compromise, as the discussion is often limited to being above or below the satisfactory bar. By downgrading staff ratings in the PCRENs/XSRENs, BDEV implicitly hints to a staff candour issue, which is often rebutted by staff arguing that PCRENs are desk-based exercises that do not account for the field

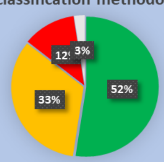
**Box 3: Candour: highlights from the QoS evaluation**

**Staff Survey:** incentives are conducive to accurate reporting, credible scoring

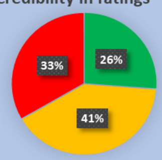



**Public Sector Projects – Desk Reviews**

Adherence to project classification methodology

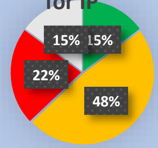



Credibility in ratings

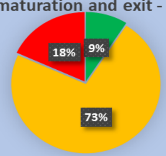



**Non-sovereign operations – Desk Reviews**

Credibility in scoring for IP




Quality of supervision at maturation and exit - XSR




QoS staff survey respondents were 128. Colour codes: Red – Fully and Partly Disagree. Green: Fully and Partly Agree.
Desk reviews were on 44 public sector projects and 27 NSOs. Colour codes measure the extent to which expectations are being met: green – fully, amber – partly, red – not met, white – not applicable)

reality. This evaluation's findings indicate a mixed picture in terms of the credibility of the scoring and the accuracy of reporting. This corroborates the main candour-related highlights from the QoS evaluation (Box 3) with 55% of the staff survey respondents who disagreed that the Bank's organizational setting and staff incentives are conducive to evidence-based reporting and credible scoring on project performance.

***Public Sector Projects***

55. A comparison between the satisfactory ratings (3 and above) assigned to the Development Objectives (DOs) and Efficiency indicators in the 2016 and 2017 PCRs and PCRENs, shows significant differences leading to questioning the candour of the PCR ratings (Table 5). However, the differences in ratings have been decreasing slightly between 2016 and 2017. As reported in the SNOQ Retrospective Report, many of the ratings of the PCRs reviewed did not provide adequate supportive evidence or justification, possibly because the PCRs are mostly based on qualitative assessments and are more subject to interpretation. The report notes that the challenge is that PCR ratings are neither impartial nor independent as long as the task manager or a consultant working under his/her supervision, drafts the completion report. This leads to inflated ratings in the self-evaluation report without supporting evidence.

**Table 5: Projects with satisfactory development objectives (Dos) and efficiency indicators (3 and above)**[14]

| | 2016 cohort (49 projects) | 2017 cohort (88 projects) |
|---|---|---|
| DOs | | |
| • **PCR** | 88% | 91% |
| • **PCREN** | 53% | 63% |
| Efficiency | | |
| • **PCR** | 65% | 57% |
| • **PCREN** | 41% | 43% |

56. Staff defensiveness on ratings is mostly due to the fact that they see BDEV validation as lacking sufficient context and in-depth information. Also, they see project ratings as a reflection of their own performance, as reflected in the interviews carried out. This is also a sign that the Operations Manual recommendations that ratings be applied through a team-based approach has not really been adopted and ratings remain to a large extent the exclusive responsibility of the task manager who takes the entire onus on him/herself and tends to personalize the assessment being made.

***Non-Sovereign Operations***

57. In the case of NSOs, candour seems to be more challenging than for public sector projects. Lack of supporting evidence and available data is a recurrent issue which also affects candour. The 5 NSOs reviewed independently as part of this evaluation raised candour issues for all the projects (Box 4). DO ratings seem to be the least credible.

**Box 4: Key issues observed in reviewing the sample of 5 NSOs**

- Outcome indicators in the PAR's logical framework and the ADOA ratings are at times imprecise and lack baselines, and there is frequently lack of symmetry between the objectives/indicators in these two documents.
- Exceedingly lenient ratings in the XSR, at times with projects being assessed against indicators that were not included in the PAR or the ADOA.
- Tendency to rate projects' outcome Satisfactory despite evident shortcomings in the achievement of outcomes.
- Imprecise definition of business success.
- Considerable deficiencies in the financial and economic analysis.
- Inadequate attention to E&S safeguards during supervision.

58. Information available in the PSR is scanty, duplicates the BTORs, and is of uneven quality. PSRs do not include much information on E&S issues or status of implementation of mitigation measures which makes it difficult to aggregate the information in the ASR. This resonates with the analysis done for 56 Financial intermediary operations conducted for the ISS Evaluation[15]. With respect to the XSRs, the timing can also play an important role in the accuracy of the DO ratings, especially when it is submitted too early before reaching maturity and it is based on unreliable client reporting on results. As

[14] Own analysis based on BDEV's Synthesis Report on the validation of the 2016 and 2017 PCRs.

[15] BDEV Integrated Safeguard System (ISS) – Draft Technical Report on overall compliance of the African Development Bank Group operations with the Integrated Safeguards System across the project cycle - Draft Report -June 2019

shown in Box 3, only 15% of the QoS sample of 27 NSOs met candour expectation and 70% met them partly or not at all.

## D. Managing Ratings

59. The advantage of ratings is that they are easy to report on, compare, average and aggregate. They can help to trigger action and assess attainment of targets through relatively clear-cut attribution. The disadvantage of ratings is that they can be seen as threatening, rigid, evaluative of staff performance, or not sufficiently nuanced to reflect real life situations.

***Public Sector Projects***

60. An important element of the Bank's ratings methodology is the use of simple averages of sub-ratings and simple averages of dimensions (relevance, efficiency, effectiveness, and sustainability) to determine the overall rating of the project in the PCR. Because this process generates decimal points, ranges are then used to attribute the final rating as an integer number within the current 4-point scale[16]. Accordingly, the current definition of Unsatisfactory (U) is a rating between 1.50 and 2.49 and the definition of Satisfactory (S) is between 2.5 and 3.49. It is therefore relevant to further examine whether these ranges can misrepresent the overall assessment by qualifying as satisfactory projects that are actually rated below 3, or more precisely between 2.5 and 2.9. The analysis shows that the number of projects falling in the latter category is 18% in the case of PCRs and 41% in the case of PCRENs. Hence, there is a substantial risk to distort the overall picture (Box 5 below).

**Table 7: Satisfactory cut-off rating and disconnect**

| | 2016 (#49) | | | 2017 (#88) | | |
|---|---|---|---|---|---|---|
| | PCR | PCREN | disconnect | PCR | PCREN | disconnect |
| **2.5 and above** | 98% | 78% | 20% | 97% | 85% | 12% |
| **3 and above** | 76% | 31% | 45% | 82% | 47% | 35% |

61. All PCRs are rated both by staff and subsequently by BDEV through the validation process[17]. Sometimes this generates divergence of views and a "ratings disconnect"[18]. The current rating methodology based on rating ranges can generate a different picture, depending on how "Satisfactory" is defined. As shown in Table 7, if it is defined as 2.5 and above, 97% of the projects are considered to be Satisfactory or better in the PCRs, which is questionably high by all comparable standards. In this case the disconnect with BDEV is 20% in the 2016 batch and 12% in the 2017 batch. However, if Satisfactory is defined as 3 and above, satisfactory projects are down to 76% and 82% in the PCRs for 2016 and 2017 respectively, while the disconnect is much larger, at 45% and 35%, respectively. Such big differences are due to the large number of projects located in the 2.5-2.9 range and raise questions about the credibility of the system[19].

62. Two of the comparators (WB and IFAD) do not use the averaging methodology or the ratings range system to assign the final ratings, but rather a judgment-based approach. For example, the WB uses a "tree view" with a cascading approach of 4 ratings (High, Substantial, Modest, Negligible) applied to each dimension, to arrive at the overall project rating on a 6-point scale (see Annex 5, Table 17, under WBG). The Inter-American Development Bank (IDB) has a similar system. One should be careful in comparing AfDB

[16] The scale is: 4-Highly Satisfactory (HS); 3-Satisfactory (S); 2-Unsatisfactory (U) and 1-Highly Unsatisfactory (HU).
[17] There is however an on-going discussion as to whether BDEV could move to a sampling approach; mostly for budgetary reasons.
[18] The "net disconnect" is the difference between the number of cases in which BDEV provides a higher rating (upgrade) and the number in which it gives a lower rating (downgrade).
[19] The WBG managed to substantially reduce its disconnect over the last 10-15 years from some 10%-15% to 2%. This is the result of full coherence of methods and ratings, strong public visibility and IEG having an impact year after year. In IFAD the disconnect is calculated differently as the difference of the overall ratings between the PCRs and the PCRENs and is reported to be -0.3 which is not comparable and tends to underestimate the issue.

performance to the one of comparators institutions that do not use the range methodology, as the cut-off rating for satisfactory performance is not the same.

63. The way ratings are structured can have an important effect on the incentive for candour. Contrary to other comparator agencies (AsDB, IFAD, the WBG) that have adopted a 6-point rating scale, the Bank implements a 4-point system. Admittedly, a 6-point rating system would allow to capture situations which are more middle-ground between Satisfactory and Unsatisfactory, such as "mostly" or "moderately" S or U. This would cater for a potentially large number of projects for whom un-anticipated issues have inevitably emerged in the course of implementation and that could not be completely fixed during supervision. The review of 137 PCRs and PCRENs from 2016 and 2017 reveals that the number of projects falling in an intermediate group with a rating between 2.1 and 2.9 were 20% in the PCRs and 57% in the PCRENs (Table 6). In these situations, it is often difficult to shoehorn project performance in the S or U category. Given the stark choice to be made between S or U, staff would typically lean towards the higher, more positive rating, while BDEV would go for the lower rating, creating the conditions for a bigger divergence. Ratings differences between the top half and the bottom half of the scale can also appear smaller depending on the choice of words, even on a four-point scale. Compare for example compared AfDB's satisfactory vs. unsatisfactory to the World Bank's substantial vs. moderate or IFC's satisfactory vs. partly unsatisfactory.[20]

**Table 6: Number of projects within rating ranges (out of 137 projects)**

| | Rating range 2.1 - 2.9 | Rating range 2.5 - 2.9 |
|---|---|---|
| **PCR** | 28 (20%) | 25 (18%) |
| **PCREN** | 78 (57%) | 56 (41%) |

[20] IFC's full rating scale for indicators is on a 4-point scale, Excellent, satisfactory, partly unsatisfactory and unsatisfactory. Overall outcomes are Highly successful, Successful, Mostly successful, Unsuccessful and Highly unsuccessful. IDB uses a similar system.

**Box 5: The unintended consequences of the ratings system**

The PCR guidelines provide the following main directives:

The PCR rates 11 criteria under four dimensions: i) the relevance of the operation from design/approval to completion, ii) the effectiveness in the actual achievement of results, (iii) the efficiency in the use of resources for achieving results, and (iv) the sustainability in the continuation of results after the project period. The performance of the Bank, the Borrower and others stakeholders are also assessed but do not form part of the overall PCR rating.

The score of each dimension is the average of the sub-criteria, and the overall score of the project is the average of the scores of the four dimensions. All criteria are equally weighted. The rating scale only allows for integer numbers from 1 to 4, however the averaging exercise inevitably leads to scores with decimal numbers. Therefore, to be able to bring ratings back in line with the original scale, the PCR preparation manual defines ranges: 1.00-1.49 (1-Highly Unsatisfactory); 1.50-2.49 (2-Unsatisfactory); 2.50-3.49 (3-Satisfactory); 3.50-4.00 (4-Highly Satisfactory).

The Satisfactory range (2.50-3.49) can be problematic to the extent that a project that reaches an average rating in the 2.50-2.99 range will eventually be categorized as Satisfactory even if there are by definition some areas that are less than satisfactory. As an example, if all dimensions of a project resulted in an average score of 2.5 (e.g., with 2 satisfactory 2 unsatisfactory ratings), the project would still be classified as satisfactory.

The four dimensions do not have the same importance in assessing the achievement of results and hence weighting them equally can affect the results. Most projects present high scores on relevance arguing that the project is in line with Borrower and Bank strategies (the average relevance rating of the 88 PCRs of projects that closed in 2017 is a high 3.7). This skews the overall average in a positive direction.

The Development Objective rating (DO), which arguably is the most important criterion in terms of delivering outputs and achieving outcomes in a combined way, accounts for only a ¼ of the overall score. The DO rating is extracted from the last IPR to feed into the PCR to assess progress made as follows:

4 – Highly Satisfactory: Both outcomes and outputs are rated highly satisfactory.
3 – Satisfactory: Both outcomes and outputs are rated at least satisfactory.
2 – Unsatisfactory: Either the outcomes or the outputs are rated unsatisfactory.
1– Highly Unsatisfactory: Either the outcomes or the outputs are rated highly unsatisfactory.

In reality, the majority of the PCRs reviewed present a mixed picture in terms of outputs and outcomes but typically the incentive structure at the Task Manager's level leans towards the satisfactory rating even if the project did not meet the expectations and the specifications of the satisfactory definition.

64. An analysis of the pro and cons of a 4-point vs a 6-point scale had been done by BDEV in the context of CSPs[21]. The main advantage of a 6-point system is to mitigate for the stark separation existing between S and U which is often perceived to be too wide, as conveyed by staff interviews. However, even a 6-point system will tend to classify performance of most interventions under "Moderately" or "Mostly" Satisfactory, which ultimately will be aggregated together with the Satisfactory and above categories as being "above the bar". This is evidenced by the case of IFAD and the World Bank. Comments received from the Evaluation Reference Group (ERG) as part of the review of this evaluation suggested to have a simple 3-point scale of "yes", "no" or "to some extent" arguing that the more the points the more complex it becomes and the more difficult it is to assess performance adequately. However, a 3-point scale would be against the Good Practice Standards established by the Evaluation Cooperation Group of Multilateral Development Banks, of which AfDB is a member.

65. In 2017 BDEV moved away from the 6-point scale and adopted a 4-point scale for the evaluation of CSPs (consistently with the public sector projects rating scale) on a pilot basis for two years. In doing so,

[21] BDEV. Strengthening Country Strategy and Program Evaluations (2017).

BDEV noted the risk that evaluators will give the benefit of the doubt and most ratings will fall on the satisfactory side of the scale, creating tensions between BDEV and Management and between Management and the Board as the distance between S and U may be too wide and cause a loss of granularity, and be detrimental to learning. Two years later, the risks proved to be valid. The same risk can also be applied to the PCR situation with the difference that in this case, PCR validations by BDEV resulted in increased levels of disconnect. The 2016 Synthesis report indicated that the current 1-4 scoring system does not allow for very nuanced evaluations.

66. Any rating system will have pitfalls, and whether a 3, 4, 5 or 6-point rating is adopted, it may not lead to significant improvements unless other measures are put in place for improving the reliability of the system and the implementation of more rigorous procedures for the generation of the required evidence. Notwithstanding, the rating scale would still be worth a deeper discussion as in the opinion of almost all the staff interviewed, the rating scale and methodology need revision.

67. The disconnect is not reported in the RMF, even though a target is now included on the "number of operations independently rated as satisfactory and above at completion". In general, a high rate of disconnect suggests problems in the M&E system and in the SESP, often reflecting unwarranted optimism about project performance. Such risk, however, does not seem to represent a major incentive for staff to rate projects with more candour, possibly because the validation ratings are not receiving much attention.

68. According to the PCR Guidance Note, the PCREN ratings are those to be used officially for the projects and they supersede the PCR ratings in case of differences. This is consistent with the practice of comparator institutions. In reality this is not the case, as PCR ratings are used as proxies for the official ratings instead. This has to do with the difficulty of having PCRENs ready in time[22], particularly when dealing with separate reporting processes such as the RMF and the ADER report, which are fed by the project ratings. This also means that typically, portfolio performance may present a relatively more positive picture than the reality and may be over-rated by at least the size of the disconnect. For instance, the RMF indicator on "operations that achieved planned development outcomes" uses the PCR rating.

69. Indicators pertaining to Bank and Borrower performance have a substantial disconnect, when considering the 2016 and 2017 cohorts of PCRs validated by BDEV (Table 8). Even if they are not included in the calculation of the overall project rating, they generally suffer of candour issues. This stems from the fact that asking staff to self-evaluate Bank performance is easily equated with staff performance, and the performance of their clients, with whom the Bank is striving to maintain a good relationship. Even in situations where the project was clearly unsatisfactory or worse, staff rated Bank performance as satisfactory. Of the 137 PCRs reviewed, the Bank's performance was rated unsatisfactory (below 2.5) only twice. This contrasts with BDEV's validation which rated Bank performance as unsatisfactory in 30 cases. The disconnect is also present in the case of Borrower performance (22 of the PCRs vs. 45 of the PCRENs). This shows that the assessment of these two indicators is not robust enough for a self-evaluation system to pursue.

**Table 8: Bank and Borrower performance (out of 137 projects)**

| | PCR | PCREN |
|---|---|---|
| **Unsatisfactory Bank performance** | 1% | 22% |
| **Unsatisfactory Borrower performance** | 16% | 33% |

70. Possible tension between Management and BDEV over discrepancies in ratings is unnecessary as reported during interviews. It could be argued that the PCR rating should be provided by BDEV only through the PCREN, as an independent source offering a single official rating. Therefore, staff would complete the PCR only with the narrative (similar to the CSP). This would simplify the process, make it less

[22] While the recommended time between the completion of PCRs/XSRs and the preparation of PCRENs is three months, the actual time lag was much longer. The 2016/2017 PCRs validations were combined and finalized in 2019.

contentious and eliminate the disconnect issue. It would also enhance the PCR learning capacity. Adopting such measure would require, however, timely completion of the PCRENs so that they are ready for preparation of the ADER report and a more collaborative process between staff and BDEV in completing the PCREN and matching narratives with ratings.

***Non-Sovereign Operations***

71. In the case of NSOs, the PSR, ASR and XSR evaluation dimensions are shown in Table 9. The SESP is not fully consistent in using common evaluation criteria and rating systems between the various documents, causing lack of clarity of the assessment process. The range of the various ratings for the different dimensions and subdimensions varies from 1-4, to 1-5 and 1-6, all with different definitions. The only harmonized ratings are those of the XSRs and the XSRENs.

**Table 9: NSOs Evaluation Dimensions**

| | | | |
|---|---|---|---|
| **PSR** | | | |
| Implementation Progress | Commercial Viability | Development Outcome & Additionality | |
| **ASR** | | | |
| Operational/financial performance | Implem. progress Devel. Outcome | E&S sustainability | Early warning |
| **XSR** | | | |
| Development Outcome<br>• Project business success<br>• Economic sustainability<br>• E&S effects<br>• Private sector development | Investment Outcome<br>• Equity<br>• Loan | AfDB Work Quality<br>• Screening, appraisal & structuring<br>• Supervision & administration | AfDB Additionality |

72. The Bank's XSR preparation guidelines of 2011 are outdated especially when it comes to defining a process for selection of NSOs. The link between the ASR and XSR documents, especially the consistency of the methodology of assessing and reporting on financial performance and impact on the Bank's Sector and Corporate Portfolio is weak. A comparison between Ex-Ante (per Board and ADOA) and Ex-Post (XSR) is made by the XSR team using DO and additionality ratings in numbers and narratively for key performance indicators. The lack of symmetry in the use of criteria and ratings makes the assessment of development outcome across the project cycle difficult. As a consequence, the NSO self-evaluation is not based on a consistent methodology for assessing performance, accountability and learning. Management is committed to address these issues under the recent NSOs Quality Assurance Action Plan (September 2019), notably by revising the guidelines and templates for ASRs and XSRs[23].

## E. Enforcement and Incentives to ensure compliance

73. Understanding what drives staff behaviour is essential for an effective SESP. In most cases, a combination of enforcement to ensure compliance and incentives will work in tandem to determine staff response. As illustrated in Figure 3, typically what works best for enforcement purposes is to make compliance and the production of the different requirements public and transparent through a central IT-supported repository of the information that can be accessed in real time. This is what has been missing so far in the Bank. It has been difficult for the evaluation team to access the information and it wasn't always clear whether the different outputs were simply not produced or whether they were not available through an easily accessible system. Management commitment to roll out the Results Reporting System (RRS) and the on-going use of the delivery Dashboard go in the right direction. By raising flags automatically in the system and making them visible to Senior Management in real time, the Dashboard works as an alert

[23] New dimensions are being included in the ASR template and assigned a rating scale for better reporting and assessment of implementation progress: i) project compliance with all covenants, ii) project systems and procedures (corporate governance, financial management, M&E functions), iii) project execution and financing; and iv) project business success (financial performance, profitability and overall contribution to an enabling environment).

system and an effective pressure point to look at issues, even if it is currently almost exclusively focused on disbursement, procurement and implementation delays, not no development effectiveness.

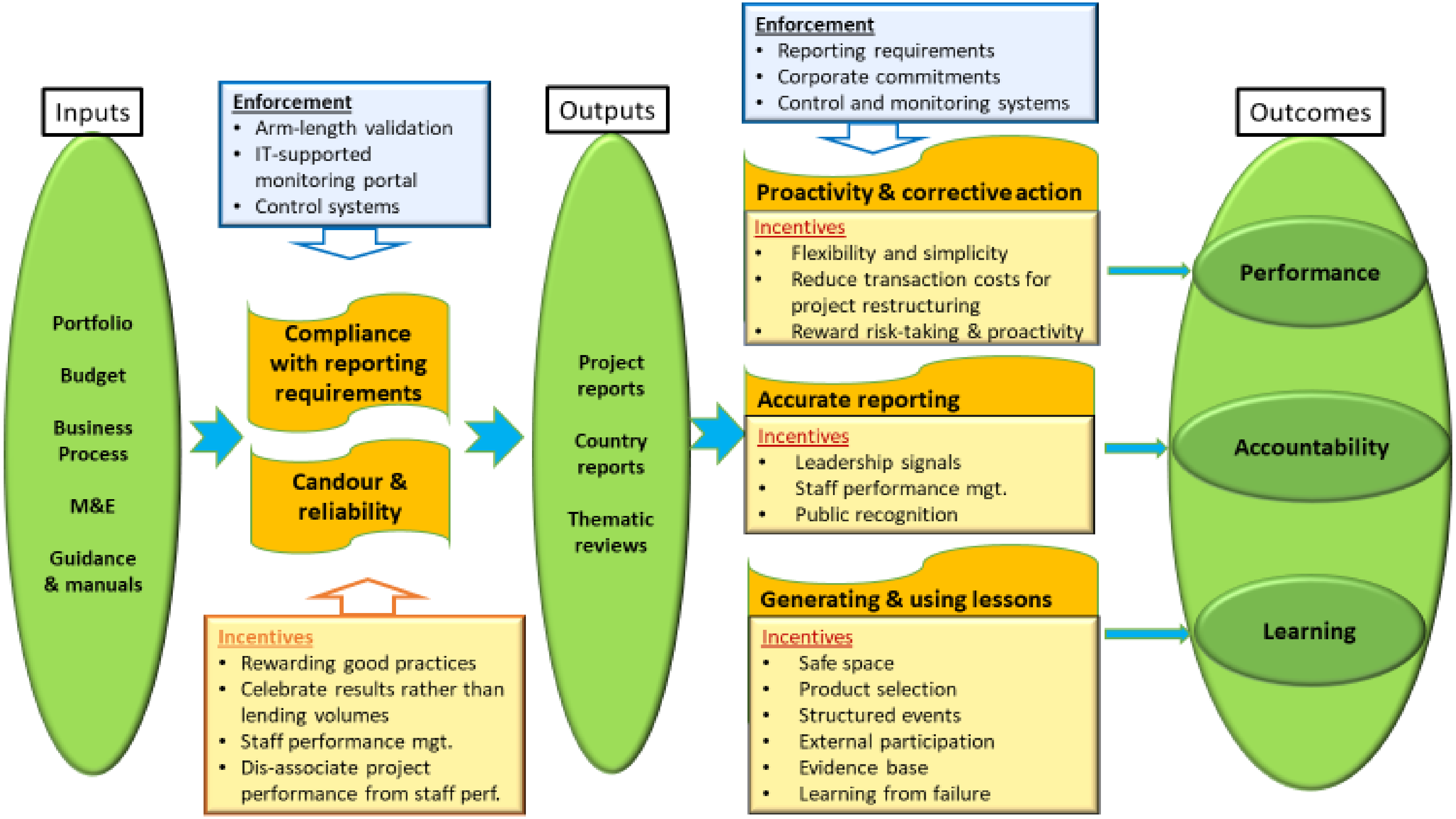
74. Independent or arms-length validation has a significant influence over behaviours and is an effective tool to improve candour and to keep the system credible. In addition to what is carried out by BDEV on Completion Reports, internal mechanisms are also possible. Private sector operations are already implementing a review and monitoring function with responsibility for project ratings through a Portfolio Management Unit detached from the originating Unit. Public sector projects have also recently instituted a new Implementation Manager Unit in each Region in RDVP, charged with ensuring effective project completion through results-oriented oversight and support of project management, portfolio performance in collaboration with the Country Managers, disbursement, legal, procurement and other service functions. It drives the actualization of the Project Implementation Plans for the achievement of project deliverables and Key Performance Indicators (KPIs).

75. Internal audits also play an important function in the mix of enforcement and incentive mechanisms and in helping to keep both the internal and external validation functions in check vis a vis compliance requirement. The delivery of KPIs is an important control measure for both compliance and performance, but currently there are no KPIs for quality and results that the SESP could cover.

76. The incentive structure presents some key weaknesses that would need to be addressed to influence key behaviours in the right direction: i) if project performance is equated with the performance of staff, the fear of damaging one's reputation in case of poor results will inhibit candour; ii) there is insufficient public recognition and rewarding of good practices related to SESP implementation, M&E systems, proactivity in taking corrective action, etc. – which could send a strong signal that Management cares about staff delivering results (in addition to loan approvals), and iii) as reported during staff interviews, there is no systematic discussion on the achievement of results during the staff performance evaluation.

77. The right mix between enforcement and incentives will need to be worked out to ensure that achievement of the SESP outcomes is not hampered by the same instruments being used. An excessive focus on ratings may lead staff to look at self-evaluation mostly from an accountability point of view undermining the capacity of the SESP to be a tool for learning. If ratings are biased towards over-optimistic views, it will hamper the SESP capacity to be a tool for corrective action.

**Figure 3: SESP Incentive Structure**



# V. Main SESP characteristics and outcomes

## A. Performance Management

78. One of the questions that the SESP evaluation tries to answer is whether the SESP architecture and its instruments are relevant and effective, in the way they are being implemented, to enhance portfolio performance and the achievement of results. The information gathered by the evaluation team, allows addressing the question by looking at the key elements that stand in the way of a well-functioning SESP.

### M&E systems

***Public Sector Projects***

79. M&E systems are the building block of any self-evaluation system, and they are also the weakest point as reported in both the 2016 and 2017 BDEV Synthesis Report. The analysis of the 49 PCREN of 2016 noted that 20 projects did not have enough information to be able to score the M&E at implementation and of the rest, 22 of the 29 projects that could be rated were considered to be unsatisfactory or worse. The average score was only 2.07 for design and 2.02 for implementation. The gender indicator scored only 1.97 and of the 33 projects for which a rating was available, 24 were considered to be unsatisfactory or worse. The 88 PCREN of 2017 noted a 41% satisfactory assessment of M&E overall with a rating of 2.7 at design and 2.4 at implementation. The gender criterion was not rated or was unsatisfactory in 49 cases. It appears that there is clearly a need to improve the M&E design especially in the gender dimension.

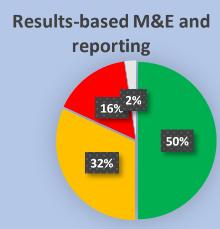
80. Weaknesses of M&E systems are often reported at design stage already, and they get carried over to implementation but they are typically more difficult to fix during supervision, as evidenced by the fact that ratings don’t improve from design to implementation. The lack of explicit links to outputs and outcomes, weak logical frameworks, delays in submitting progress reports by the borrower as well as difficulties in recruiting and retaining M&E experts, were noted.

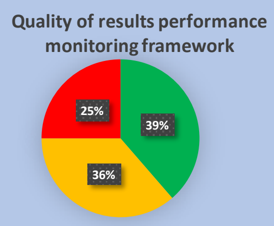
81. Findings from the 2016 and 2017 cohorts of PCRENs highlighted weaknesses in terms of information gaps with respect to evidence justifying ratings, which resonates with indications of absence of quality control mechanisms. Many ratings were downgraded by BDEV’s validation not necessarily for failing to meet the target but for lack of evidence that would support a particular assessment. This feedback resonates with QoS desk review (Box 6) also showing mixed results with respect to overall performance of the M&E system and results monitoring frameworks.

82. M&E issues are not new in AfDB and have also been identified in comparator institutions. The Task Managers interviewed stated that significant efforts were being made to implement and use a sound M&E system but that more measures were required to encourage adequate

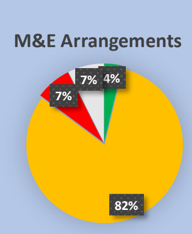
**Box 6: QoS M&E - Desk Review**

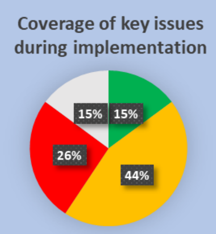
**Public sector (out of 45 projects)**




**Private sector (out of 27 projects)**




*Colour codes measure the extent to which expectations are being met: green – fully, amber – partly, red – not met, grey – NA)* use by borrowers. The limited monitoring capacity on the part of borrowers to make adequate use of M&E systems, constrains the use of the full information by the Task Managers when preparing IPRs. Data quality at the source remains an issue and affects the credibility of the whole system. Measuring and monitoring performance has often been an elusive target especially when process and outcome indicators are difficult to quantify, making the link with outcomes more arbitrary. If results reporting is to receive the required priority, the Bank and the Borrower must pay proper attention to the integration of: i) baselines, ii) logical frameworks, and iii) M&E systems.[24]

***Non-Sovereign Operations***

83. In the case of NSOs, M&E arrangements and coverage of key issues during supervision has been difficult to analyse for lack of information and limited availability of PSRs. M&E of financial sector projects is significantly worse than other projects (as reported in the QoS). It was difficult to assess the covenant compliance in the NSO consolidated portfolio report with respect to Annual E&S Monitoring Reports (ARM) which are required to be prepared by clients and cover, among others, E&S criteria, health and safety monitoring indicators, social programs, GHG emissions and energy efficiency. ARM findings are not systematically followed up and sufficiently reported upon in PSRs, ASRs and XSRs. As also noted in the ISS evaluation, E&S supervision of private sector projects is lagging in spite of recent ad-hoc efforts to improve the templates for some NSOs. However, the level of coverage of E&S information in XSRs was better than in the supervision reports analysed. Clients reports on development outcomes are not assessed or validated by the NSO portfolio managers making it difficult to ascertain their reliability.

84. The NSO consolidated Portfolio management and monitoring report, prepared jointly by the various departments and PINS, is mainly focused on financial risk assessment, value addition to Bank's income, and return on investments, but does not include an overall assessment of the development outcomes and achievement of private sector development objectives. The recent diagnostic study of the existing results system, undertaken by BDEV, revealed important gaps in terms of monitoring of achievements in private sector interventions, including potential inconsistency between development objectives set at origination and those that are tracked during implementation.

85. In the case of Lines of Credit, the problem stems from the poor definition of how they will be used and their likely impact on DOs, since the key success measurement is the timely reimbursement of the sub-loan. Half of the financial sector projects examined by the QoS (11 projects) had unsatisfactory supervision reports and the other half were partly satisfactory. To address this challenge, the Bank has put in place a new initiative, led by SNOQ and PINS departments, to harmonize the result reporting requirements and improve operational effectiveness as well as ensure pro-active management of development results[25].

## Raising issues in a timely fashion

***Public Sector Projects***

86. Task Managers are required to provide information on project progress following a field supervision mission or desk review within 30 days, through an IPR, which is subsequently approved by the line manager. The timeliness in reporting after supervision missions is critical for the prompt raising of issues

[24] The WBG has also highlighted weaknesses in M&E and low capacity both in-house and in client country data systems. The Results Measurement and Evidence Stream (RMES) was put in place to strengthen M&E skills and most "results staff" were absorbed into Global Practices in the centre. Capacity building is also offered by the Regional Centres for Learning on Evaluation and Results (CLEAR) initiative. This could lead to demand for more specialized skills to invest in good M&E. A similar CLEAR initiative is also being implemented by IFAD.

[25] New guidelines are under preparation with new templates and valuation process of Bank assets. A corporate portfolio repository system making PINS the stop-shop for data on projects from inception to independent evaluations (XSRENs) is being introduced.

and the taking of remedial action. One of the main functions of project supervision is the follow-up and resolution of issues identified in previous supervision missions through the SESP.

87. Deficiencies have been identified on all the points above: timeliness in submitting IPRs after a mission is poor, with an average time of 55 days and only 53% of supervision reports submitted on time against a target of 70%[26]. A review of the supervision missions carried out for 83 public sector projects from 2013 to 2017 by the QoS evaluation showed that the percentage of submitted and unassessed IPRs increased from 20% to 45%. Interviews carried out as part of the SESP evaluation indicated little reliance on the IPR documents across the board as they are typically approved automatically by the Manager with no discussion with the Task Manager and little validation. This does not necessarily mean that issues are not raised through other means such as the BTORs, but it does show that the IPR, as a SESP tool, is not serving its function.

88. The limited use of IPRs and their ratings for the purpose of addressing issues and taking corrective action may be a missed opportunity to the extent that Implementation Progress (IP) is a good predictor of overall project performance. The correlation rate between IP and the overall project rating for the 2016 and 2017 PCRENs cohorts is of 0.67 and 0.61 respectively. Addressing implementation issues as they emerge and in a timely fashion, is one of the main expected contributions of the self-evaluation system to portfolio performance and the achievement of results. If IPR ratings are not reliable, not comparable, difficult to access and not timely, the willingness to act on them is unlikely to be very high. Of the 12 projects reviewed as part of this evaluation, 5 have not been rated whether in SAP or in the IPRs (including 4 projects approved after 2011, when the IPR system was introduced).

89. Interview responses for the SESP evaluation indicated quality control gaps in the progress reporting process. It was reported that while supervision teams raise key implementation issues in Aide memoires and BTORs, in most cases these issues are not adequately addressed or in a timely manner. The tendency not to give projects a poor score seriously limits the opportunity for early identification of issues which is when they need to be addressed, before they become too entrenched and difficult to resolve. The IPR is not considered a decision-making tool but rather a compliance tool (administrative requirement) with insufficient accountability since they are not much discussed nor systematically validated by management.

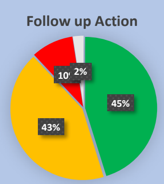
90. In addition, the quality of IPR ratings (or lack thereof) is even more important because the IP and DO ratings of the last IPRs are exported into the PCR. IPR ratings are also used to assess portfolio performance through the CPPR which is a key component of the CSPs. Hence the quality of the IPRs also permeates into CSPs. Moreover, IPRs are not adapted to programme-based operations (PBOs) which follow a different business model and for which a different format would be required (para 141).

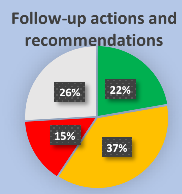
91. In that respect, the Delivery Dashboard does a better job than the IPRs as early warning and raising management attention. The portfolio dashboard inserts flags according to verifiable performance indicators and can categorize projects in potential or actual problem status even if the task managers think that all is well. IPR timeliness and validation are not captured by the dashboard but it does allow an assessment of the reliability of the ratings. There are currently no results-related KPIs that could complement the delivery dashboard.

92. MTRs are considered to be better vehicles for project adjustments. Performance upgrades did occur in five out of the 13 problematic projects that underwent an MTR in the QoS sample (45 projects). However, of the projects reviewed, only 40% showed evidence of follow-up on

**Box 7: QoS - Follow-up action.**
**Desk review**

**Public sector** **Private sector**




*Colour codes measure the extent to which expectations are being met: green – fully, amber – partly, red – not met, grey – not applicable)*

[26] 2017 Annual Quality Dashboard report.

previous mission's recommendations for public sector projects and 27% in the case of NSOs (Box 7).

***Non-Sovereign Operations***

93. In the case of NSOs, recommendations of supervision missions are often not sufficiently actionable and vaguely described. The capacity to manage and mitigate risks is low, especially for lines of credit, since the Bank has limited influence over financial institutions to improve compliance and provide relevant data with respect to the achievement of DOs that could be used to trigger corrective action as needed. There is a tendency to use ADOA results framework to forecast DOs years after the project has started, even when actual DO data should be available, and to leave PSRs rating unchanged unless there is a major change. The expectation that the project development rationale and the logical frameworks in the PAR and the ADOA be aligned, could not be evidenced.

94. An NSO dashboard is available with the distribution of the active portfolio geographically, by financial instrument, by sector, and with key performance indicators such as the non-performing loans (NPL), and the weighted average risk ratings (WARR). The dashboard assesses performance through the value addition to the Bank's income and return on investments, arrears and non-performing assets, the risk profile, and annual supervision status (based on PSRs undertaken during the year). It does serve as an alert system vis-a-vis the WARR but does not cover other aspects like for public sector projects.

95. The NSO project "Watch List" [27] is an additional tool to raise issues according to the status of the portfolio under: recovery/loss; rehabilitation; and active monitoring. The watch list is based on financial criteria with no assessment of risks to development outcomes. It is held under PINS coordination which decides on the ratings and on whether the NSO should be under close monitoring or to be managed by the Special Operations Unit (SOU) to find solutions. The SOU provides specialized knowledge and services to resolve issues of non-performing loans including through restructuring, turnaround or in some cases re-capitalization of distressed and underperforming companies.

## Transaction costs of project restructuring

96. The nature of the problems to be addressed plays an important role on the incentive for the Task Managers to raise and address issues, particularly if the project has to go through some form of restructuring involving high transaction costs. Task Managers interviewed note that raising issues may cause additional work and most of them are already over-stretched. Formal restructuring is rare, both because the transaction cost is perceived to be too high and the Borrower is not always forthcoming, fearing to add delays. Interviewees noted that countries change priorities all the time but typically projects are not restructured to reflect the changes, and the Bank cannot proceed without Borrower's agreement. Also, putting together a team with the right skills mix in view of changing priorities, has not been easy for many task managers.

97. Task Managers interviews conveyed a feeling that taking corrective action and project restructuring would be more proactively pursued if task managers felt that there was a support structure in place rather than being left entirely to themselves. Task Managers are typically very client-oriented and project-oriented and are committed to pursuing project objectives, but the Bank needs to show that in supporting the Borrower, there is an institution behind the Task Managers. The recent shift from project supervision to

[27] Projects are placed on the watch list for reasons including: non-payment of principal or interest when due; significant decline in collateral value; borrower reports a loss either in half or full year results and there is evidence the loss was not due to extraordinary circumstances; financial difficulties of the borrower; poor information disclosure; poor corporate governance issues relating to the competence of management; major management changes, especially of key decision makers without evidence of an acceptable succession plan; negative market trends, government directives; legal suits or threats of bankruptcy by other creditors; deterioration in the economic environment in general or in the industry the company operates in.

"implementation support", the appointment of the Implementation Support Managers in each region and the stronger focus on decentralization go in this direction.

98. The 12 public sector projects reviewed for this evaluation showed the need for some degree of project restructuring for the majority of them. The main reasons were initial procurement and disbursement delays often linked to lack of readiness for implementation, changes in Government priorities or local conditions, low capacity of the PIU, inadequate counterpart funds, complexity, and over-optimism. In most cases, targets were revised at the mid-term review (including major adjustments) but in no cases did the projects undergo a formal restructuring through the Board. The main reasons cited for that were the high transaction costs, avoiding delays, and Borrower's reluctance. It was also argued that changing targets could be done without changing objectives and therefore no formal restructuring was needed. This however creates a methodological issue for BDEV since it is required to use the formally-approved targets in conducting the validation of the PCRs, which in the absence of formal restructuring, remain the appraisal targets, and become a source of divergence.

99. This is also the case for CSPs/RISPs where there is reluctance to recourse to the Board at mid-term to modify coherence and relevance of the strategic pillars in an environment of changing priorities. As shown by the CSP case studies (Annex 7) and interviews, the issue is now amplified with the required alignment to the High5 priorities and possibly the need to retrofit existing strategies.

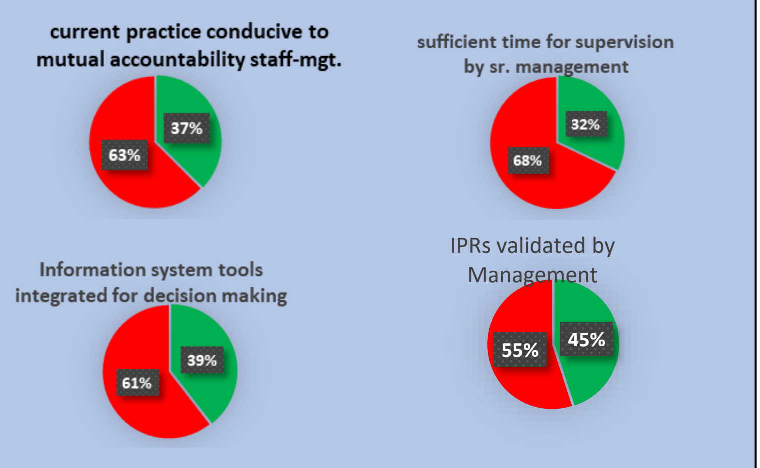
100. Restructuring procedures are not sufficiently flexible and tailored to different degrees of changes needed, which would also allow to realign with BDEV evaluation methodology applicable to revised targets in case Board approval is not needed. This mirrors similar concerns in comparator agencies[28]. If the Bank is adamant to achieve faster project design and approval[29], it should also recognize that notwithstanding PD 2/2015, critical design activities (e.g. procurement plans, feasibility studies, validation of Environmental & Social Assessment studies (as reported in the ISS evaluation), setting up of PIUs, baselines) are often rushed or carried over to implementation, which results in initial delays, greater inertia and more difficult problem resolution. Task managers should feel more empowered to proactively raise and follow up on issues that can improve implementation performance, including the possibility of project restructuring, if necessary, rather than supervising projects with a compliance-check approach.

## B. Promoting Accountability

101. One of the questions that the SESP evaluation tries to answer is whether the right degree of accountability is in place to allow the SESP architecture and its different instruments to function in a credible and effective way, signalling that the Bank is holding itself accountable for achieving results. Conversely, it is asking whether the SESP architecture is being implemented as a tool to enhance accountability towards: i) Management and ii) the Board.

102. Accountability is often the best way to achieve candid self-evaluation, meaning that someone in the hierarchy demands it. Procedures, guidance and

**Box 8: Accountability - QoS Survey Results (83 public projects)**




*Colour codes: Red – Fully and Partly Disagree. Green: Fully and Partly Agree*

[28] In the WBG it was recommended by the SES evalua legal agreements as well as through simplified Bank a frequently as needed, informed by relevant and time
[29] Presidential Directive N° 2/2015 applies to the impl seeks to achieve five objectives: (i) improve the opera (safeguards, procurement, project management) be c projects; (iii) improve institutional efficiency; (iv) redu accountability.

rules are well in place and sufficiently clear. The main challenges are related to their implementation and the level of accountability that would be required for the system to work effectively. Staff survey responses to some accountability-related questions from the QoS evaluations highlights a number of issues (Box 8). This requires looking at institutional incentive structures first. Management signals are very important in that respect, especially because accountability and incentives in the Bank have traditionally been skewed towards lending rather than the achievement of results. This is changing, however, and a number of measures have been put in place in the recent QA Implementation Action Plan (2019-2021) to establish a better balance.

103. Similar issues occur in all other comparator institutions, and it is understandable, since staff have greater control over project design and feel more accountable for it. Achieving results through project implementation and supervision is messy, it involves many stakeholders, especially Governments and there are many reasons for things to go wrong beyond the capacity of staff to fix, hence more reluctance to be held accountable for poor performing projects.

104. In looking at the two levels of accountability – internal and external – it is important to note that the two systems are running in parallel and there is little ownership or involvement of operations in establishing the right links between the SESP outputs and the way they get aggregated to feed into the RMF. The 17 Task Managers and Country Program Officers interviewed were little aware of what is required for the RMF and how the SESP can serve both purposes. Similarly, a number of RMF level 2 sectoral indicators [e.g. jobs created] are requested at the corporate level and trickled down at the operational level without much consultation on what is feasible to collect through the M&E systems.

### Internal accountability

***<u>Public Sector Projects</u>***

105. Self-evaluation by staff requires reliable evidence to function properly. Several factors influence the way accountability is being exerted in the implementation of the SESP:

- Effectiveness of M&E systems and how they link with result frameworks is at the core of credible reporting and accountability. As discussed previously, the results in this regard are modest, with issues raised both at design and implementation stage.
- There is little reviewing and probing by Management of the ratings and underlying issues when clearing IPRs. Staff interviews confirmed that project supervision is not receiving adequate attention or follow up from line Managers. Validation of IPRs by Management is often done in a mechanistic way (when done), as Managers have little time for IPRs in addition to the BTORs.
- Staff turnover is quite high as revealed by interviews and corroborated by the QoS evaluation, with insufficient handover. Having three to four task managers is quite common during the life of a project. This results in a lack of continuity, reduces incentives for proactive action, and dilutes accountability.
- Third party validation is a major part of the accountability process, serving to keep the system honest. This can be done externally (BDEV) or internally (through peer reviewing or by a different division). By and large the staff survey and the interviews confirm the usefulness of having independent ratings but most respondents disputed the relevance of the PCREN deskwork for not being sufficiently contextualized and discussed with staff. This can be the source of tension.

**Table 10: Quality of PCRs**

| | 2016 cohort | 2017 cohort |
|---|---|---|
| **# of PCRs rated 3** | 15 (30%) | 56 (64%) |

| | | |
|---|---|---|
| **and above** | | |
| **# of PCRs rated 2.5 and above** | 40 (82%) | 65 (74%) |
| **Average PCR rating** | 2.67 | 2.9 |
| **Delivered on time** | 77% | 58% |

106. PCR quality is assessed by BDEV in the PCRENs. The PCR quality rating is affected by the same issue as project performance. Results are good when measured against the 2.5 rating benchmark (77% in aggregate over 2016 and 2017), as per the current "range" methodology, but modest (52%) if measured against a 3 rating (Table 10). The PCRs quality situation can also be reflected through the average score which remains at 2.67 and 2.9 in 2016 and 2017 respectively, revealing an important scope for improvement. However, a significant upgrading from the 2016 to the 2017 batch should be noted for the 3 and above ratings (from 30% to 64%), while timeliness got worse (from 77% to 58%). Management own review of PCR timeliness in 2018 indicates a rebound to 85% as the result of a major push by the regional teams to improve timely completion. By and large, the majority of staff interviewed think that task managers should not do their own PCR. However, while using consultants under the supervision of task managers may provide some of arms-length review, candour does not improve much. Also, consultants do not follow the same standards which raises issues of comparability and quality. PCRs are typically given lower priority by staff and there is little vetting on the choice of consultants.

107. The recent establishment of the Implementation Support Manager position could help address accountability issue by engaging on portfolio issues with the sector divisions, providing the regional perspective and fostering better harmonizing across sectors. It would move in the direction of a team approach (with the task manager as the lead person), which is consistent with the objective to improve candour and objectivity and with the guidance provided by the Operational Manual. In that respect, CPOs would also have a stake and contribute to a more collegial approach, to the extent that individual projects are also part of the country portfolio that the country/regional teams need to manage, hence benefitting from greater convergence.

108. Interviews carried out as part of this evaluation suggest, however, that in the context of increased decentralization, a better definition of roles would be required and some concerns were raised as for the real capacity of this new function to be carried out consistently on the whole portfolio, given the scarce availability of human resources. Generally, sector managers appear to retain a direct line of accountability over IPRs and ratings, but the practice seems to vary from sector to sector and from region to region.

109. Even if some of the task managers interviewed have been prone to involve the Borrower in rating projects, this practice is not recommended as it will reduce flexibility and Managers will have even less of an incentive to question the ratings and exert accountability. Borrowers may, however, be part of the independent review process.

110. The Portfolio dashboard (and quarterly portfolio flash report) is an accountability mechanism that ensures better compliance and provides a framework for enhanced proactivity on corrective action. Portfolio performance as illustrated in the dashboard report is in the Performance Agreement of Vice-Presidents with the President. It could be cascaded down to Managers. The concept of making performance of some key project indicators widely available is typically very powerful especially if the system allows fixing issues before it is brought to the attention of Senior Management. Making key performance indicators (disbursement, procurement, implementation delays, etc..) directly accessible by all staff, is quite common in comparator institutions.

111. In the same vein, the rolling out of the RRS would permit to post, on a dedicated on-line portal, key project implementation data and the most critical SESP products (IPRs, MTRs, PCRs). It would potentially be a powerful instrument for improved accountability and reporting and an effective tool for Managers to streamline reporting requirements and easing the Task Managers' burden. Making information available

widely is also likely to improve compliance by flagging omissions or delays. This addition is much needed and one of the areas where the Bank has been comparing poorly with comparator institutions.

***Non-Sovereign Operations***

112. The PAR results framework lacks precision and clarity in identifying development indicators and baselines. NSOs also lack clear measuring tools of the progress towards development objectives during implementation and data to report in PSR, ASR and XSR. Mostly, discrepancies between development results' indicators assessed at origination (PAR and ADOA) and those tracked during implementation, lead to poor monitoring and reporting of achieved results and reduce the likelihood of effective risk mitigation during implementation.

113. An independent assessment of the quality of PSR, ASR and XSR undertaken for the purpose of this evaluation has revealed discrepancies in assessing development outcomes. In fact, there is a great tendency to rate project development outcomes as satisfactory despite evident serious shortcomings and data constraints. The E&S information is particularly scant but there have been ad-hoc recent efforts to improve the E&S supervision of lines of credit. These issues reduce the effectiveness of the SESP in assessing and managing performances particularly with respect to development outcomes and investment profitability. This is also due to insufficient knowledge of NSO evaluation guidelines and weak financial and economic analysis of projects.

114. Notwithstanding, there is more of an internal validation system in place for NSOs than in the case of public sector projects. Accountability for portfolio management lies with the Credit and Risk Committee (CRC) and third-party portfolio monitoring is played by PINS which offers the opportunity to pursue more of a dual accountability approach and oversight function (including rating projects). However, the low level of compliance compared to what is normally required seems to indicate low levels of accountability. Development outcome ratings are not systematically validated by the CRC nor is there a strong assessment of risks to development outcomes in the ASRs.

***CSPs and RISPs***

115. Country and Regional Programs (CSPs and RISPs) are all self-evaluated (by relevant country/regional offices) at mid-term and at completion (CSP-CR, RISP-CR). The CSP-CRs and RISP-CRs are prepared together with country/regional portfolio reviews (CPPRs) to assess achievement of the strategic objectives and possible review of the results framework in line with the performance of the country or regional portfolios. CSP-CR and RISP-CR use no ratings but a notional assessment of satisfactory or unsatisfactory, and are sent to the Board (CODE) for information. Even if CSPs and RISPs are not rated by staff, they are both fed indirectly by the IPR ratings of the individual projects in their portfolio through CPPRs (since CPPRs and CSPs are presented in tandem) whose reliability is uneven but typically quite optimistic. Therefore, the compliance and candour issues described above for projects, permeate into CSPs and RISPs.

### External accountability

116. There is strong demand for accountability from the Board. The main reporting tool being used by the Bank to inform the Board is the RMF and the ADER report. The SESP is particularly relevant for the RMF level 3 indicators which has to do with the Bank's operational effectiveness[30]. Most of these indicators are fed by the self-evaluation system in an aggregated way. Therefore, the degree of accountability that can be

[30] Level-3 indicators are: i) Increase the development impact of operations, ii) enhance the quality and speed of operations, iii) ensure strong portfolio performance, and iv) increase the quality of Bank's knowledge and advisory service. Each of these indicators relies on a number of sub-indicators.

exerted by the Board on operational effectiveness is only as good as the quality of the information provided through the SESP.

117. Interviews with ED's Offices raised a number of issues (see Annex 6) related to:

- The reliability of the RMF, the limited information received on the source of RMF data (and whether it is validated or not), and the role played by the SESP and how decisions get made.
- The low level of compliance with SESP requirements (IPRs, MTRs and PCRs) and their inconsistent implementation.
- The fact that Board's comments are recurrent, but nobody seems accountable for addressing them.
- There is willingness to re-discuss the 4-point scale ratings methodology, particularly the adequacy and definition of the Satisfactory and Unsatisfactory ratings and how they can be used consistently and effectively in projects, APPRs, CPPRs, CSPs and RISPs.
- The impression that the SESP is a box-ticking exercise. Staff turn to "satisfactory" very quickly, even when issues are present, and the narrative and the ratings do not match.
- The fact that CSPs are always satisfactory. The methodology for CSP self-evaluation should change. The Board does not focus much on PCRs but mostly on CSPs. They are however concerned about the disconnect mainly between CSPs and CPPRs.
- The mixed quality of project level logical frameworks as they are not used consistently to engage with counterparts and to measure success. The main problem remains at measuring properly the outcome level.
- Conflict of interest in having task managers preparing, supervising and then self-assessing their own work. This can only lead to candour issues.
- The need to eliminate the perception that project performance is associated with staff performance; build a reputation of "project fixers".

***Public Sector Projects***

118. A number of points can be made:

- In reporting the number of projects meeting their DOs, the RMF uses the PCR rating as a proxy, since the timeliness in the delivery of validated ratings has been challenging. Such approach can be debatable, not only because the Operations Manual notes that the official rating should be the one assigned by BDEV but also because the PCRENs Synthesis Report (with contains BDEV validation ratings) is separately transmitted to the Board, often with different and lower ratings, which may leave the Board to wonder which one to use and the origin of the disconnect. In all other comparator organizations, the project performance information conveyed to the Board is the validated one.
- A new target of >80% is included for the number of projects assessed positively by BDEV. However, results vary considerably depending on how satisfactory is defined. As discussed earlier in the report, using a rating range of 2.5 to 3.49 as the definition of Satisfactory conveys a rosier image than if the 3-rating was used as benchmark, as described in the current 4-point rating system.
- Some RMF targets are set at a level which is not realistic given current performance or given the issues raised above. For instance, the target of at least 93% for sustainability seems achievable when using the PCR rating of 2.5 and above (87% for 2016 and 2017 combined), but is overambitious when using the PCREN rating of 3 and above (43%) or even 2.5 and above (71%).
- The "net disconnect", which is the difference between the number of cases in which BDEV provides a higher rating (upgrade) and the number in which it gives a lower rating (downgrade), is a relevant indicator for the Board to consider. This indicator can be regarded as a proxy of the "candour gap".

***Non-Sovereign Operations***

119. The performance of NSOs is not subject to external reporting through the RMF level-3 indicators, which may be related to higher degrees of confidentiality attached to NSOs. Contrary to other sister institutions such as AsDB and IFC, the Bank's self-evaluation system does not inform the Board about NSO's performance at early maturity through expanded supervision reports evaluation notes (XSREN). Thus, the current practice does not offer the Board a timely and accurate independent opinion on achievement of NSO development outcomes and Bank's contribution and effectiveness. This also means that in general the SESP is less focused on results and less adapted to the DO assessment of the NSOs' business model. On a selective basis, however, projects are subjected to a comprehensive post evaluation, within two to three years of their completion. This exercise is conducted independently by BDEV. Upon completion of the evaluation, a Project Performance Evaluation Report (PPER) shall be produced and circulated to Management and the Board in compliance with the guidelines.

***CSPs/RISPs***

120. CSPs and RISPs self-evaluation reports are not validated by BDEV (see Annex 7 on the CSPs/RISPs case study), contrary to comparator organizations. However, BDEV has recently validated one CSPs and one RISP on a pilot basis and will assess scope for expansion[31]. BDEV, does carry out independent country program evaluations (CSPEs) on the basis of potential need for evaluative information (e.g. revision of Policy or Strategy when one expires) which encompass one or more CSPs or RISPs periods. These are distributed separately to the Board prior to the discussion of the next CSP/RISPs. For that, BDEV has shifted from a 6-point to a 4-point rating system in 2017 (in alignment with the Bank scale) on a pilot basis for two years. From 2014 to 2019, 22 CSPEs were prepared and discussed with CODE. The CSP/RISP-CRs assessment and the independent evaluation of CSPs and RISPs are not harmonized to ensure a full comparative analysis of what has worked or not based on common evaluation criteria and rating assessment of performance. Although an attempt was made to align country strategies to the High5 priorities, the assessment is not based on clear evaluation criteria such as relevance, efficiency, selectivity, effectiveness, sustainability or risks to development outcomes, using key performance indicators.

121. Issues of consistency between CSP and CPPR outcomes have emerged recently and the relevance of addressing at the same time portfolio management issues and the country assistance strategy has been questioned (see Case Study in Annex 7). CPPRs are mostly informed by the active portfolio (IPR ratings), which is affected by compliance and candour issues, while CSPs are informed by their own results framework and narrative. Currently, the CSP-CR is not included in the dashboard; thus, it is of little value to help inform the new CSP. In that regard, the whole CSP process should be looked at to strengthen the accountability system and learning from past experience and also to prepare for the next CSP. Also, with the recent introduction of the Diagnostic Notes for CSPs and RISPs, the articulation with the CPPRs and their performance could also be tackled. Management has embarked in a discussion with CODE on the required revision of the CSP/CPPR methodology which is now under way[32].

## C. Learning

122. The SESP spend a substantial amount of resources dealing with products whose main outcome is to promote learning. This is especially the case of completion reports for projects, CSPs, RISPs, thematic

[31] The pilot exercise for the South Africa CSP brought out some key issues, such as the absence of detailed assessment criteria or guidance on the ratings for the different dimensions for the CSP completion self-evaluation exercise. This leads to varied quality and content of the reports. The quality of the CSP results-based framework and its use as a monitoring tool needs to be improved, notably the linkages of proposed inputs, outputs and outcomes.

[32] Information Note to CODE on the Bank Group's Assessment Methodology of CSP Performance in CSP Completion Reports (12 December 2018)

reports, and sector strategies. In the case of projects, completion reports also get validated by BDEV. It is therefore legitimate to ask how much learning actually happens and whether such investment is producing the intended benefits.

123. Generally speaking, staff survey's respondents (see Annex 8) are broadly positive that SESP outputs are good vehicles for learning (59%). However, the individual interviews carried out on this topic show a great deal of scepticism across the board. PCRs/XSRs are reported to be potentially good tools for learning especially in the same country for follow up projects, CSPs and CPPRs, assuming they are produced in a timely fashion. They are admittedly less relevant for staff in other regions as they think that lessons are too generic, perfunctory and not actionable or too specific and not replicable. It was reported that learning from the SESP is happening sporadically as PCRs do not always allow to extract a credible story that can be used somewhere else. The fact that some issues are recurrent and can be found in almost every PCR (poor quality of M&E, insufficient attention to cost-benefit analysis, disbursement delays, procurement issues, low capacity of PIUs, restructuring issues), as documented in the project sample of this evaluation and the Synthesis Report for 2016 and 2017 PCRs, is a sign that not much learning is being translated into action.

124. There is no single place where SESP information and products can be accessed. The SESP has not led to a solid repository of knowledge that is mined and shared regularly by staff. Learning opportunities could be greatly enhanced if the information was accessible on-line through a common platform with new forms of learning and new templates. This is in the making with the advent of the RRS which will include data from all SESP products.

125. There is a need for more creativity in terms of dissemination and sharing lessons with other countries on the same sector or similar specific issues. According to the PCR-EN Manual, periodic half-day workshops are to be organized with relevant staff to enhance feedback on findings drawn from PCR/XSR-EN. This rarely happens. BDEV could take the initiative, jointly with Management, to organize learning events centred on their Evaluation Notes or corporate evaluation studies and add new dimensions. Learning could be organized thematically through structured events and with external participation, including Borrowers and beneficiaries. Most of all, learning events should be evidence-based and be seen as a safe space where people are willing to learn from success as well as from failure.

126. If the PCR/XSR is seen as a tool to judge the task manager, it will undermine candour in ratings and learning opportunities. The rating itself is actually seen as an obstacle to learning because it potentially makes the discussion unnecessarily contentious. If evaluation is about learning it should not be based on ratings. It should focus on generating and sharing knowledge with a forward-looking objective of adding value, providing solutions and improving a course of action irrespective of how it may be self- or independently assessed in a single project context.

127. There is therefore a fine line to be walked between the search for accountability and learning. Strategy papers (country or thematic) are more conducive to learning, possibly because no ratings are involved. Disagreement over ratings between staff and BDEV further undermines the incentive to promote learning. Task Managers interviews highlight that BDEV validation is often seen as too rigid and unfair, too focused on the initial results framework (that staff finds too cumbersome to update) and they disassociate themselves from the process and the information that goes with it.

128. The little communication between the PCR/XSR team and the BDEV team is not conducive to align views and better understanding of the issues. Debating and understanding the reasons for divergent views could actually be a source of learning in itself and improve the quality of the self-assessment function. Notwithstanding, the format of the PCR/XSR could also be improved to enhance learning, so that it is not seen as a perfunctory box-checking exercise at the end of the cycle. The majority of the staff survey's respondents agree that the lack of trust in the SESP undermines its learning potential.

129. Two of the most difficult aspects of evaluation and learning are the quality of the evidence being produced and the extent to which outcomes can be attributed to the Bank. The need for evidence-based lessons is normally addressed through the articulation between baselines, results frameworks and M&E systems. In the absence of solid evidence, projects are penalized at validation's time and learning becomes uncertain. Establishing attribution is harder and requires putting in place expensive statistically-proven methodologies at least for a sample of projects. Occasional impact evaluation studies could be used on an ad hoc basis to provide a solid set of evidence around which to organize learning events. Integrating impact evaluation methods into project monitoring systems could effectively provide quantitative and qualitative approaches to reporting not only on the achievement of results but also on implementation issues.

130. To facilitate learning it can help to assign responsibilities for coalescing knowledge from self-evaluation, existing thematic reviews and impact evaluations around themes, sectors or clusters of similar projects or issues. Also linking learning to the project cycle assumes that learning from PCRs can only happen once the project has closed, which means on average 7 to 9 years after the project started. PCRs are supposed to be delivered 6 months after closing and many of them (57%) incur delays and come too late even to inform follow-up operations, which are prepared before project closing. More flexibility and real-time learning would be required which also means better ways to extract knowledge gained during supervision.

131. The lending culture has not helped foster learning through the self-evaluation processes. However, the signals are changing and a number of measures put forward in the QA Implementation Plan are going in the right direction. Ultimately the Bank will need to decide whether the same instrument can pursue two outcomes at the same time: accountability and learning, or whether the trade-offs are such that dedicated and distinct mechanisms and products are required. The WBG came to the same conclusion in its recent SESP evaluation report[33] recommending that in looking for a better balance between performance, accountability and learning, the accountability function of the SESP should not be sacrificed and voluntary impact evaluations should be expanded to cover a wider spectrum of interventions in a given country, sector or region.

132. Reporting against objectives and outcomes through ratings for each individual project and carrying out a validation of PCRs are logical steps in an accountability framework but they are not conducive to promote learning. Learning goes beyond an individual project and its PCR, and should ensure that it can meet the demand for knowledge in the institution in combination with other products. This should be done in a "safe space" environment and therefore the use of ratings and the perception that staff performance is being judged would be counterproductive.

133. In the case of NSOs, the ratings and lessons learned are checked by sector managers in departmental meeting during which investment and portfolio officers normally attend to enhance the feedback loop. However, there is no consolidated lessons learned database accessible by NSO development and portfolio officers. Lessons learned should be institutionalized and documented to ensure a strong capitalization of lessons learned based on Operations, Risk Management, ECON and BDEV interaction.

134. Furthermore, there is need to increase learning from NSOs on E&S safeguards and climate change adaptation and also on SDGs and mainstreaming of safeguards by reaching out to clients and private stakeholders. In that respect, the ISS independent evaluation has recommended to strengthen the E&S work at project identification and ensure proper documentation to improve resource allocation for project preparation and supervision and reinforce the QA mechanisms for supervision reporting. A couple of factors limit the capacity of the NSOs' SESP to contribute to learning: i) the low level of compliance and

[33] IEG. Behind the mirror: a report on the Self-evaluation systems of the WBG (2015).

hence of reporting, ii) the required level of confidentiality and hence lower disclosure; iii) delays in XSR validation which may impact the perception of credibility of the SESP; and iv) the lower degree of public element embedded in NSOs makes the relationship with the client more based on financial performance and less on achievement of development results.

# VI. The SESP through the lens of the Evaluation Criteria

135. This evaluation used the OECD-DAC evaluation criteria of relevance, effectiveness, and efficiency. It focused on results beyond outputs and assessed consistency of the Bank's self-evaluation systems and processes with the Bank's independent evaluation function.

## A. Relevance

136. The Bank's SESP are to a large extent relevant and aligned with comparators' best practices. The existing guidance tools and procedures are clear, complete and aligned with the Bank's key policies, strategies (H5, TYS) and business model (DBDM). A theory of change is not explicitly available for the SESP but in line with the practice at comparator organizations the system has consistently pursed three main outcomes: promoting improved performance, fostering accountability, and enhancing learning. The operational manual and some of the SESP outputs would need however to be upgraded in view of a number of trends and priority areas (gender, climate and fragility) as well as special investment vehicles (PBOs, technical assistance). Results from the SESP evaluation show that the main weakness of the SESP is not in the established procedures, standards and norms but in their applications, particularly with respect to low level of compliance, and deficient candour.

137. A shortcoming emerging from the evaluation is that gender issues are not sufficiently covered in projects, CSPs or RISPs due to weak methodology in assessing results (lack of outcome indicators). Gender experts are not systematically involved during the project cycle, CSP/RISP mid-term reviews and CRs. Some countries have benefitted from thorough assessments of gender profile and helped in designing gender policy strategies and action plans. There is no clear guidance in place to report on gender results in the SESP through the M&E systems, as evidenced by low ratings in the 137 PCRENs reviewed. No dedicated budget is in place to cover specifically gender issues during preparation, implementation and completion. A categorization system of gender issues, based on gender profiles, will ease follow up and supervision and increase the credibility of the SESP for reporting on gender in projects, CSPs and RISPs.

138. Cross-cutting issues such as climate change (CC), and E&S safeguards have been raised during interviews as these topics could be better integrated in the SESP. This is confirmed by the ISS independent evaluation which found a low compliance with the Environmental and Social Action Plan (ESAP) requirements. The review of the available supervision reports found a partial and vague reporting focused on certain environmental/social mitigation measures, but not an overall picture of the total measures included in the E&S management documents of clients and borrowers. Also, the articulation of the climate change requirements with other E&S due diligence conducted by the Bank has been challenging and the corporate indicators used by the Bank, for instance in the ADER, are focused on work done before Board approval.

139. The status of E&S safeguards is missing in portfolio flagship reports. Risks to climate change adaptation are typically not reflected in assessing performance. The CSP/RISP SESP do not systematically report on countries commitments to National Determined Contributions (NDCs) and capacities in CC adaptation and resilience. Such reporting would help countries' accreditation for green climate fund. The M&E framework with indicators (2016-2020 update) provides an operational guidance for reporting on climate change issues at project or CSP/RISP levels. CSP/RISP SESP should also report on processes to deliver sector results on E&S safeguards with adequate risk assessment.

140. The level of attention devoted to safeguard issues in the SESP during implementation compares poorly with the attention devoted to these aspects during project design as evidenced by the fact that social, environmental and safeguard experts at the regional level are few and are not systematically involved in field missions or meetings. This finding is corroborated by the ISS evaluation which highlights

understaffing as the most significant constraining factor to the implementation of the ISS. The evaluation notes that it is often during project implementation that environmental risks are most pronounced but that the current staff and long-term consultants are not able to ensure that 50% of high risk projects are properly supervised. Climate change, social and environmental safeguards specialists do not systematically support teams during project and CSP/RISP implementation as well as to increase interaction with ECON complex and build capacity to address E&S safeguards at country or regional level.

141. An additional issue that the SESP face is that some of the Bank's formats and templates do not take into account the different business model of the PBOs. Many sections of the IPRs are not adequate as well as the reporting frequency as most PBOs only have one or two disbursement tranches and when the project goes to the Board, first tranche's conditions have already been met and the project is fully or partly disbursed shortly after. Therefore, design issues need to be tackled as a priority and different arrangements need to be made to be able to account for policy dialogue beyond the loan closing date (which is typically very short) to be able to assess how policy changes have impacted sector or macro performance.

142. Similarly, more specific and adapted formats, areas of enquiry, budgets and frequencies of reporting would need to be developed for fragile and conflict situations. Special consideration and requirements for additional analysis on fragility exist for project preparation, but this fades away during implementation and it is not considered a factor during completion, so that these projects are evaluated as if they were operating in a normal environment. Impact of fragility on outcomes should be more easily discernible through the SESP, as this is likely to improve accountability, to the extent that many issues related to fragility are beyond the capacity of the task manager to fix.

143. The SESP is closely integrated with BDEV independent validation function in line with comparators' institutions. However, while ensuring BDEV independence, the search for better convergence between staff and BDEV on ratings, harmonization of rating methodologies, definitions, and consistency of what is reported to the Board needs to be improved. The "disconnect" is currently not reported in the RMF and it is worth reviewing the ratings methodology as a way to improve candour and the credibility of the system.

144. The private sector department follows a risk-based approach to supervision of NSOs, which is performed by portfolio officers and a credit risk team that monitors the project credit risk performance. Overall, its approach is considered to be relevant and aligned with comparators. However, the originating team is insufficiently engaged in project supervision, as gathered from the staff interviewed in the private sector divisions. In general, the stronger focus on the bankability of the project may take incentives away from providing the right level of attention to the development impact of projects. The fact that little is reported publicly about the performance of private sector operations may be one of the causes of low level of compliance and lack of data.

## B. Effectiveness

145. The staff survey indicates that the SESP provide a framework for portfolio management which is reliable, timely and effective, and that it can be used as a way to proactively follow-up on problems during supervision. However, implementation is more problematic as the same staff survey reveals that the primary role of the SESP is often viewed as ensuring compliance through box-ticking, protecting one's reputation, and avoiding a disconnect with BDEV. Staff are split equally on the general question as to whether the SESP is seen as a compliance mechanism or a tool to achieve results (Annex 8). These results should be taken with caution due to the low response rate (6%) to the survey, but were triangulated and are broadly consistent with other sources (QoS, staff interviews).

146. Many Bank projects suffer from weak M&E systems and results frameworks and the lack of credible evidence generation, as described earlier in the report. As stated by a staff survey respondent "It is not very

useful for the Bank to invest in robust arms-length and/or independent validation while the inputs to such a system at project level continue to not have sufficient M&E, knowledge, and learning support". More robust M&E, objective results evidence, and data collection protocols should be integrated into project implementation systems. This needs to happen from design, include good baselines, solid monitoring system, and evaluation at closing. There are very few M&E specialists, and none in regions, which hinders the effective delivery of SESP outputs.

147. One area where progress is still needed is the capacity of the SESP to be an efficient tool for raising issues in a timely fashion and taking corrective action as demonstrated by the following findings: 39.8% of projects identified as problematic remained flagged for a long time (2015-2016) and 49.7% in the following year (2016-2017)[34]. The lack of swift corrective action may have several compounding causes: i) task managers hope that pending issues will go away by the next supervision mission if the borrower diligently follows the Aide memoire's recommendations, which may not happen; ii) low compliance in filing and timely submissions of IPRs compounded with inadequate management attention limits the effectiveness of proposed recommendations; iii) low levels of accountability by Management may exacerbate the situation and let issues drag along; and iv) reluctance to restructure projects to avoid additional burden and delays, resulting in retaining appraisal targets that are no longer in line with the project reality.

148. NSO clients are involved in reporting on the development outcomes based on a template provided by portfolio officers during implementation. To some extent, the nature of NSO clients makes them more concerned about financial returns than development objectives. This is particularly the case for Lines of Credit operations. Providing evidence that would support the project development rationale and causality between project activities and impact was found to be problematic in NSOs. This was evident from interviews, the in-depth analysis of the 5 NSOs and is corroborated by the QoS analysis. The NSO PAR results frameworks is weakly aligned with ex-ante metrics provided by ADOA and the reporting system lacks precision and clarity in identifying development indicators and baselines to be used for PSR, ASR and XSR preparation. This reduces the effectiveness of the SESP in assessing and managing performance, and it is therefore difficult for the SESP to look at NSOs through the same effectiveness lens used for public sector projects. However, the reviewed XSRs offer better coverage of development results and output/outcome achievements.

## C. Efficiency

149. As examined earlier in this report, lack of compliance seems to be one of the main hurdles that stands in the way of an efficient SESP. Staff survey respondents think that this issue can be mitigated through: i) a greater dose of accountability by Managers (for instance asking to include comments in the IPR form when clearing); ii) some form of simplification and making the reporting requirement more flexible and adapted (for instance expecting outcomes in the initial years of project implementation is unrealistic); iii) expanding the existing KPIs to achievement of results, iv) more coaching, mentoring and training.

150. Despite the fact that most of the interviewees and the staff survey consider SESP tools to add value (Annexes 6 and 8), there is scope for minimizing duplication and redundancy in the information conveyed by Aide memoires, BTORs and IPRs. In particular, the IPR is the one that suffers most from compliance issues, but it is the instrument that carries the ratings. An effort should be made to render its use as efficient and credible as possible. Its format should be updated to be simpler and friendlier to the various users according to the type of operation, context and priority areas. Most of staff survey respondents suggest some degree of merging between BTORs and IPRs, as they are both targeting an internal audience and they carry most of the overlap. Comparator agencies have confronted similar situations and while IPRs

[34] QoS evaluation results. Desk reviews (2018)

are generally considered the main reporting instruments, the use of BTORs is left to management discretion.

151. Cost-efficiency analysis of the SESP was constrained by lack of uniform data at the central level. Information gathered during individual interviews showed levels of budget allocation in the same range as comparator institutions albeit on the low side. However, contrary to comparators (IDB, WBG, IFC), the Bank does not use differentiated budgeting norms according to special context, complexity, and status of the active portfolio.

152. The recent decentralization and focus on continuous implementation support require further adjustments to the operational guidelines to ensure a common understanding of new roles and responsibilities. In particular, the role of the country office and of the new Implementation Support Manager, could help improve efficiency in supervision and portfolio monitoring. Increasingly projects are being managed from country offices but they are not yet fully staffed and are not entrusted with sufficient delegation of authority, as stated during staff interviews.

153. The Bank has struggled so far, to put in place an IT-supported on-line integrated platform that would facilitate access to all information pertaining to SESP in one single location. The new RRS which is being rolled-out will address this issue and be a powerful efficiency tool to improve accountability and performance.

154. The issues raised for public sector projects in terms of efficiency are also relevant for NSOs, possibly to a higher degree with observed redundancy between PSRs, ASRs and BTORs. The type of business model and relationship established with the client is admittedly less conducive to use the SESP efficiently for timely reporting on issues and taking corrective action, especially for financial sector operations, as also noted in the QoS evaluation. Comparator institutions also use sampling methods to produce XSRs (40% coverage in the case of IFC), sufficient to allow for statistical inference. Evaluating the XSRs enables IFC to account to its Board for achievement of its purpose and for learning. In particular, the results of the XSRs are analysed by the Independent Evaluation Group at the end of each XSR cycle and presented as part of the annual Results and Performance of the WBG, which is not the case in the AfDB.

# VII. Findings and Recommendations

**Main Findings**

1. **The Bank's SESP have many positive features. They lay out strong standards and procedures underpinning their functioning as well as a cogent articulation with the independent evaluation function carried out by BDEV.** Annex 5 presents a detailed description of the SESP of comparator institutions showing a good level of alignment which was to be expected from members of the Evaluation Coordination Group (ECG.) Many of the tools and processes in place are consistent with and as good as those of comparator institutions. The articulation between self-evaluation, validation and independent evaluations is similar, and the system produces corporate results data that are used to report to the Board. However, this evaluation finds that progress towards a culture of development effectiveness has been mixed and the potential of the SESP to make an impact on the three fronts of performance management, accountability and learning, is unmet.

2. **The main weaknesses of the SESP are in the application of the established procedures, standards and norms**. While the issues identified as part of this evaluation mirror similar constraints faced by comparator organizations, two factors seem to be affecting the proper functioning of the SESP to a higher degree in the AfDB: i) low compliance with established procedures and ii) deficient candour and a positive bias in assessing performance. This has affected the credibility of the SESP and contributes to a perception that the system is adding little value compared to its costs.

3. **AfDB has an independent evaluation policy for BDEV approved in 2016 and amended in 2019. However, the Bank does not have an integrated evaluation policy that covers both the independent function and self-evaluation of the Bank itself.** BDEV promotes the use of evaluation findings on specific topics in line with the demand coming from the Board and Management which are also the primary users, but BDEV is not a user of evaluation findings as such. Evaluation adds value only when its findings are used. BDEV evaluation policy necessarily focuses on the supply side of evaluation and not the use side. In the absence of such common framework, alignment of strategic approaches and harmonization of methodologies and processes will remain at risk. Further, BDEV is not the only generator of evaluation findings. All completed operations and country/regional strategy papers (CSPs/RISPs) are self-evaluated by staff. Some comparator institutions have an integrated evaluation policy in place (IFAD and WBG) or are considering its introduction (AsDB) (Annex III for more details).

4. **Effectiveness of the SESP is constrained by its ratings methodology and structure and the way it is applied to PCRs**. Several elements can be highlighted:

   - The Bank uses a 4-point rating system[35], but several discussions took place in the recent past, on the merit of adopting a 6-point system like other comparator organizations (AsDB, IDB, IFAD, the WBG). Admittedly the 6-point system allows for more flexibility and realism in assessing performance of the large pool of projects falling somewhat in between Satisfactory and Unsatisfactory, as the distance between the two is often perceived to be too wide and too stark, according to staff interviews. However, it was also noted that the 6-point scale does not permit to have a clear judgment about project performance, and that such scale would tend to classify performance of most interventions under "Moderately" or "Mostly" Satisfactory anyway, which ultimately will be aggregated together with the Satisfactory and above categories as being "above the bar". Other recommendations made by ERG members or by the 2016 PCR Synthesis Report are in favour of an odd-number system (3 or 5), as it will allow to better account for the middle space of the distribution (assuming a symmetrical profile).

[35] The scale is: 4-Highly Unsatisfactory (HS); 3-Satisfactory (S); 2-Unsatisfactory (U) and 1-Highly Unsatisfactory (HU).

- The main risk associated to any rating scale is to assess whether it could lead to more positive ratings without supporting evidence and subsequently to an increased "disconnect" with BDEV ratings. This happens to be the case in the Bank. Any rating system will have pitfalls and whether a 3, 4, 5 or a 6-point rating is adopted, it may not lead to significant improvements unless other measures are put in place for improving the reliability of the system and the implementation of more rigorous procedures for the generation of the required evidence.
- The use of simple averages of sub-ratings and dimensions (relevance, effectiveness, efficiency, sustainability), combined with the "range" methodology, whereby "satisfactory" (or better) is defined as a rating above 2.5, leads to mis-representation of the performance of the portfolio. The system allows to pass the satisfactory bar even if the project did not meet the satisfactory specifications as expected by the level-3 definition. This is particularly the case when results frameworks are too qualitative and imprecise, and leave a large degree of discretion in assessing delivery of outcomes.
- As a consequence, there is a significant difference on project performance depending on how "satisfactory" is defined. At 2.5 and above, 97% of the 137 PCRs analysed[36] are considered to be satisfactory, while at 3 and above it is 80%. Similarly, the "disconnect"[37] between the overall rating assigned by staff in the PCRs, and the lower rating provided by BDEV validation in the respective PCREN is 15% (at 2.5 and above) and 39% (at 3 and above). This raises questions about the credibility of the PCR ratings.
- Using the PCR/XSR ratings as the "official" score instead of the PCREN/XSREN ratings provided by BDEV through the validation process conveys a more positive picture than the reality. This is especially the case for RMF indicators on portfolio performance and departs from the practice of comparator's institutions. The issue is mostly related to the difficulty of generating the required validated information in a timely fashion so as to feed in the submission of the corporate reports;
- Rating Bank and Borrower performance in the PCR generates very large differences with BDEV ratings in the PCREN. It has proven difficult for staff to rate negatively the Bank performance which is easily associated to their own performance, and that of their clients, with whom staff try to maintain a good relationship and naturally they don't want to be seen as finger-pointing.
- The downgrading of ratings by BDEV is a source of tension and is often rebutted by staff arguing that validation is a desk-based exercise that does not account for field realities. Managing ratings will remain a controversial subject as long as it is viewed as a tool to pass judgment on staff performance. Consultations between staff and the BDEV team around the PCR validation process remain rare. In the opinion of almost all the staff interviewed, the rating methodology needs revision. Whether the system is too heavy on ratings or not remains to be seen, but at a minimum there should be an attempt to mitigate the perception that ratings are easy to "game", that BDEV validation is out of context and mostly antagonistic in nature, that the disconnect with BDEV can be largely ignored, and that the rating methodology is too rigid and bureaucratic.

As illustrated in the Theory of Change (Section III), the underlying logic of the SESP is that it can play an important role in improving performance management, accountability and learning. The rest of the Findings are organized along these three main outcomes.

### A) Performance Management

5. **The lack of candour in self-evaluation, particularly in IPRs/ASRs and PCRs/XSRs, can be explained among other things by the lack of a proper incentive structure.** The perception that project performance is equated to staff performance undermines the motivation to rate poorly-performing

[36] BDEV. Synthesis Report on the validation of the 2016 and 2017 PCRs (2019).
[37] The "net disconnect" is the difference between the number of cases in which BDEV provides a higher rating (upgrade) and the number in which it gives a lower rating (downgrade).

projects candidly. The effectiveness of the SESP is undermined by their being viewed as a compliance mechanism that is driven by box-ticking, protecting one's reputation, and relying on weak generation of evidence (M&E and results frameworks). Candour is also undermined by weak accountability mechanisms. Finally, there is little recognition that being identified as a "problem fixer" could motivate staff towards greater proactivity in raising issues and corrective action.

6. **A number of issues constrain the contribution of the SESP to improving portfolio performance**

- The Bank's culture, incentives, and institutional KPIs, are skewed in favour of lending approval, like in other comparator institutions, with limited emphasis on the quality dimension and development results. This issue has been recognized by Management and being addressed through the QA Action Plan;
- There is insufficient attention to incentives that support a culture of quality and results. Opportunities for recognizing, celebrating, internalizing and learning from good quality M&E, results frameworks, proactivity in addressing issues or project restructuring, and successful implementation remain limited (this is also being addressed by the QA Action Plan);
- There are weaknesses in M&E systems and how they are articulated with baselines and results frameworks. Many PCR ratings were downgraded by BDEV validation for lack of evidence that would support a particular assessment;
- Excessive focus on accelerating project approval leads to critical design activities getting rushed or carried over to implementation (procurement plans, feasibility studies, validation of E&S assessment studies, setting up of the Project implementation Unit - PIU, baselines). This results in early implementation delays which require a stronger SESP to fix issues from the start;
- There is a strong tendency to avoid addressing issues through formal project restructuring because the transaction costs are considered to be too high. This results in failure to introduce corrective measures and leads to retain appraisal targets that are no longer in line with the project reality;
- Increased decentralization and the move towards continuous implementation support, together with the establishment of a new position of Implementation Support Manager for public sector projects is seen positively. However, new roles and the division of responsibilities between staff at HQ and at the country/regional level need clarification;
- The IPR ratings of the current portfolio feed into the Country Portfolio Performance Report (CPPRs) which in turn also impact the assessment of the CSPs since they are presented in tandem. The Development Objectives (DO) and Implementation Progress (IP) ratings of the last IPR also migrate to the PCRs. Therefore, the compliance and candour issues affecting the IPRs permeate through other SESP outputs and all the way to the RMF. Also, issues of consistency between CSPs and CPPRs' assessment methodologies have emerged recently which require revision (now under way).
- In the case of NSOs, the lack of clear measuring tools of the progress towards development objectives and the absence of a rigorous and institutionalised M&E system, lead to poor tracking and reporting of achieved results and reduce the likelihood of effective risk mitigation during implementation. This makes it difficult to assess, analyse and report adequately on portfolio performance. These issues are now being contemplated under the NSOs QA Action Plan (September 2019) adopted by Management.

7. **Project teams and task managers are fundamentally motivated to help clients deliver results and by and large are committed to the supervision task**. However, most task managers and investment officers are overstretched and the additional support required to address issues and help in the proper implementation of the SESP has been wanting. The Bank seems to be short of staffing and skills required to implement the SESP efficiently and effectively. This is particularly the case for M&E and E&S

safeguards during supervision, as also documented in the recent evaluation report on integrated safeguards system (ISS)[38].

**B) Accountability**

8. **Low compliance stems from insufficient accountability mechanisms and deficient visibility**. In the absence of reliable information, the SESP lose their credibility. The current enforcement and incentive systems do not prevent staff from ignoring the rules that suit them least and getting away with not generating the required outputs. In the words of one of the interviewees: "if the Manager cares, staff will care". Moreover, the Bank has been lagging behind in the development of an IT-supported on-line portal that can provide access to portfolio and SESP data and raise the visibility and the efficiency of the system. Providing real-time information and compliance data to line managers will enhance accountability for supervision, and completion of IPRs/ASRs and MTRs while easing the burden on task managers by simplifying and streamlining reporting requirements. It will also facilitate more effective portfolio reviews and planning exercises. The recent roll out of the Results Reporting System (RRS), as part of the QA Action Plan for both the public sector projects and the NSOs, is meant to address this issue. Once completed, it will facilitate automatizing the reporting of results, timely escalation of issues to senior management, accountability, reliability of information, efficiency of reporting at the corporate level and transparency at implementation.

9. **Reporting tools such as Dashboards and the RMF have proven to be powerful accountability mechanisms and should be enhanced to cover indicators of compliance**. The fact that the disconnect between Management ratings and BDEV ratings is not reported takes away an important element of accountability. Additionally, very little information is conveyed to the Board in the RMF with respect to the performance of NSOs, notwithstanding the need to take into account issues of confidentiality.

10. **Accountability processes have placed insufficient focus on quality of monitoring and closure.** The IPR/ASR and the PCR/XSR are not fully considered a decision-making or a learning tool but rather an administrative requirement with inadequate accountability or attribution. Since BDEV doesn't validate IPRs/ASRs, there are few opportunities for contestability of ratings as evidenced by the fact that IPRs are not much discussed nor systematically reviewed by management. Line managers are not systematically held accountable for quality checks at supervision and quality-related KPIs are missing. The role played by the Portfolio Management Unit (PINS) for NSOs is akin to a dual accountability approach and provides some degree of arms-length review and contestability. For public sector projects, the recent addition of the Implementation Support Manager position in each region should help engage on portfolio issues, liaise with the sector divisions, provide the regional and country perspective, including CPO involvement for better convergence with country portfolio management.

11. **A consistent and harmonized framework that allows to assess performance throughout the project life from origination to independent evaluation, is lacking**. This would allow to link the SESP and its tools with the front-end portion of the QA process (quality at entry) and ensure that the same approach and indicators are being used on progress reports as at the end in PCRs and XSRs. This will also make clearer where the lack of candour comes in and allow for "no surprises" by providing predictability on the assessment metrics being used. This also means keeping results framework up-to-date in case of changes to the project environment and, on the NSOs side, including monitoring indicators in line with those adopted by the ADOA framework.

---

[38] BDEV Integrated Safeguard System (ISS) – Draft Technical Report on overall compliance of the African Development Bank Group operations with the Integrated Safeguards System across the project cycle - Draft Report - June 2019

12. **Templates and formats for a number of SESP tools are not sufficiently differentiated and adapted to specific circumstances**. There is room to make the templates more efficient and user-friendly and avoid redundancies especially between IPRs/ASRs and BTORs, as also noted by the QoS evaluation. IPRs are not adapted to special investment vehicles such as PBOs and TA projects nor to fragility situations. The PCR and PCREN templates are overly repetitious, too long and some sections duplicative, as also evidenced by the Final Synthesis Report of the PCRs' validation (2019). Templates are not designed for optimum management attention and do not focus on priority issues or priority actions needed.

13. **The capacity of the SESP products to report and address specific issues is weak in the areas of safeguards, gender, climate and fragility**. In particular, as reported by the recent ISS evaluation, the Bank's supervision reports and BTOs do not capture the key E&S information to allow compliance checks. E&S information found in the available supervision reports is scanty except when a potential high corporate risk materializes. Other studies conducted by SNSC corroborate the point that once a project is approved, the Bank's internal reporting system offers very little information about the E&S follow-up. The gender dimension of M&E systems is particularly weak with average ratings in the PCRENs below a 2-rating. Project teams don't have enough support in these areas and coaching on SESP requirements to mitigate the effect of the high staff turn-over that the Bank has experienced in the last couple of years. The newly launched Operations Academy should help address this issue.

14. **Completion reports of CSPs/CPPRs and for RISPs are not validated by BDEV before being submitted to the Board**, unlike comparator institutions or similarly to PCRs and XSRs which are validated and submitted to the Board in the form of a synthesis paper. BDEV has carried out validation of one CSPs and one RISP on a pilot basis and is assessing whether there is scope for expanding the process. Management is in the process of starting a review of CPPRs methodology and guidelines that should address this point including the articulation between CPPRs and CSPs[39].

### C) Learning

15. **Despite the fact that performance management and accountability aspects of the SESP are weak, they have overshadowed or even undermined learning objectives.** This finding resonates with the situation in comparator institutions. There is a fine line between the search for accountability and learning. Strategy papers (country or thematic) are more conducive to learning, possibly because no ratings are involved. Disagreement over ratings between staff and BDEV further undermines the incentive to promote learning. If the PCR/XSR is seen as a tool to judge the task manager, it will undermine candour in ratings and learning opportunities. The rating itself could become an obstacle to learning because it potentially makes the discussion unnecessarily contentious and personalized.

16. **There is no single place where SESP products and information can be accessed**. The SESP has not benefited from the availability of a solid repository of knowledge that is mined and shared regularly by staff. This is in the making with the advent of the RRS which will include data from all SESP products. Incentives, combined with new forms of learning and templates may be needed, providing solutions and lessons irrespective of how it may be self- or independently assessed in a single project context.

17. **The quality of PCRs, as measured by BDEV validation process, is variable, depending on how "satisfactory" is defined, but average ratings are low (around 2.8), which hinders learning opportunities**. The PCR quality rating is affected by the same issue as for project performance. Results are good when measured against the 2.5 rating threshold (77% satisfactory in aggregate over 2016 and 2017), as per the current "range" methodology, but modest (52%) if measured against a 3-rating

[39] Information Note to CODE on the Bank's Group's Assessment Methodology of CSP Performance in CSP Completion Reports. December 2018.

threshold. By and large, the majority of PCRs are prepared by consultants. Most of the task managers interviewed think that they should not do their own PCR for reasons of conflict of interest. However, while using consultants under the supervision of task managers may provide some level of arms-length review, candour remains an issue and consultants do not follow the same standards which raises issues of comparability and quality. PCRs are typically given lower priority by staff and there is little vetting on the choice of consultants.

18. **There is no systematic feedback from the SESP of NSOs that provide success or failure stories based on assessment of achievements in reaching development outcomes and profitability**. Lessons learned are not institutionalized and documented to ensure a strong capitalization of lessons learned. However, BDEV high level evaluations are contributing in capturing lessons and generating knowledge that enable new strategic orientations for NSO and private sector development.

19. **The SESP on its own cannot cater for the array of learning needs of the Bank.** Learning through the SESP is not sufficiently complemented by other sources of knowledge such as impact evaluations, thematic reviews etc. There is a need for more creativity in terms of dissemination and sharing lessons with other countries, on the same sector or similar specific issues. Periodic workshops or events, as recommended in the PCR-EN guidelines, to be organized with relevant staff to enhance feedback on findings drawn from PCR/XSR-EN, rarely happens. If SESP documents are not sufficiently evidence-based and events are not seen as a safe space where people are willing to learn from success and from failure, incentives to learn will remain insufficient.

**Recommendations** (related Findings in brackets)

1. **Elaborate an overall Bank integrated evaluation policy**. Currently only BDEV has an independent evaluation policy but the Bank as such does not have an integrated evaluation policy that allows to cover both the independent function and the self-evaluation of the Bank itself. An integrated policy would provide a comprehensive governance framework, processes and procedures covering both the Bank and BDEV. This would be an opportunity to define how evaluation contributes to performance, accountability and learning in terms of desired outcomes. It would facilitate the realignment and harmonization of methodologies and ratings definitions, and link agreed outcomes to the Bank corporate results framework. It would also allow to describe the role of Management, the Board and BDEV, including BDEV degree of engagement with others in AfDB. Annex 3 provides a description of what such integrated policy would entail. (F.3)

2. **Reform the ratings methodology** of PCRs by: i) abandoning the current "range approach" and assigning the overall rating through a "judgment" assessment vs. the current practice of simple averages of sub-ratings and dimensions; ii) ) the official project rating should be the PCREN/XSREN rating provided by BDEV through the validation process, including for reporting to the Board in the RMF, and appropriate timelines should be agreed between BDEV and Management to ensure that validated ratings can feed into key reports; iii) self-rating Bank and Borrower's performance in the PCR should be discontinued and left to BDEV only; iv) the on-going revision of the Operational Manual should ensure that strong and more rigorous procedures and guidance is provided in terms of the generation of the required evidence for rating attribution. (F.2; F.4)

   Points to be further discussed. Related to the area of reforming the ratings methodology, this evaluation thinks that there is merit to further discuss the following points as additional options for consideration or as part of a learning event.

   - *Modify the current 4-point scale*. The analysis of whether other rating scales could be more appropriate and conducive to improving accuracy and candour was inconclusive. Pro and cons were identified with each model, including when looking at comparators, interviews, and ERG comments. The recommendations submitted above are seen to have more direct pay-offs for improving the system, irrespective of the rating scale. Notwithstanding, and if the above measures were adopted, the modification of the 4-point scale may still have merits to address the issues presented in the analysis and in the findings, as in the opinion of almost all the staff interviewed, the rating scale and methodology need revision.

- *Single-source PCR rating*. It could be argued that rating the PCR twice (by staff and by BDEV) is unnecessary and inefficient and only leads to tension around possible differences. The PCR rating could be provided by BDEV only through the PCREN as part of the validation process, as an independent source offering a single official rating for corporate reporting. Staff would complete the PCRs with the narrative only. Such approach has the main advantages of i) simplifying the process, making it less contentious and eliminating the disconnect issue, ii) taking care of the recommendation to stop self-rating Bank and Borrower performance, and iii) improve the PCR learning opportunities. It would require however to ensure i) timely completion of the PCRENs so that they can be ready for the preparation of the ADER report, and ii) a more collaborative process during PCREN preparation between staff and BDEV, whereby alignment between narrative and ratings can be discussed and adequate level of staff accountability be preserved. Such measure would represent a departure from the current practice in other comparator organizations, even if an excessive focus on ratings was already highlighted in the evaluation of the WBG self-evaluation system[40]. Also, ECG practice note of 2018 (Annex 4) observes that self-evaluation systems may elect not to apply ratings in some instances and should be sufficiently flexible to accommodate management and evaluator judgement.

3. **Put in place an incentive structure conducive to enhance the capacity of the SESP to achieve results.** This recommendation comes as a complement to a number of measures decided by Management as part of the recently approved Quality Assurance Action Plan (2019-2021) to sharpen the focus on delivery and results. Management recognition of the importance not only of good design but also of proactive implementation, will be supported by a programme that rewards excellence and innovation in project design, project implementation and project restructuring - including project and portfolio turnaround successes. It also proposes information campaigns on quality assurance and new tools and templates, including a revision of the Operational Manual to reinforce reporting and compliance.

   In that context it would be important to ensure that task managers and task teams will be supported in the pursuit of achieving results and good quality SESP by providing additional capacity, especially at the regional/country level (for instance as an enhancement of the role of the Implementation Manager) in dealing with problematic/complex situations, implementing M&E systems and better articulated results frameworks, dealing with E&S safeguards, thematic issues (gender, climate), specific situations (fragility). Simplifying project restructuring procedures is also required to improve the incentive for corrective action. Finally, SESP practices should be part of the staff performance conversations, with a view towards greater recognition of proactive behaviour in dealing with complex, problematic issues and achievement of results. (F.5; F.6; F.7; F.10; F.13)

4. **Improve compliance and candour through enhanced accountability, the usefulness of performance monitoring tools and of corporate reporting**. Measures include: i) incorporating compliance indicators in the delivery dashboards for Management attention such as IPRs/ASRs validation by managers, timeliness of filing reports, implementation of the MTRs; ii) establishing new results-based KPIs and top-level targets to be cascaded down to line managers and task managers for staff performance management and accountability; iii) rolling out the Results Reporting System (RRS) for both public and private sector operations developed by Management for enhanced transparency, efficiency and

[40] IEG. Behind the mirror: a report on the Self-evaluation systems of the WBG (2015).

accountability; iv) validating CSPs/CPPR (subject to a revision by Management of the CPPR methodology) and RISP Completion Reports for reporting to the Board; and v) identifying a relevant indicator on NSOs performance to be aggregated as part of the RMF level 3 indicators. (F.2; F.8, F.9; F.10; F.14)

5. **Redesign/Adapt the template of some of the SESP reporting tools and improve their quality.** The IPR and ASR formats should be updated and simplified to reduce redundancy and overlaps with BTORs, improve efficiency in reporting and making management discussion and vetting a requirement before filing and for compliance monitoring. IPRs should be differentiated for diverse typologies of operations such as PBOs and TA projects, and adjusted for country circumstances (e.g., fragility). ASRs should develop a more precise and actionable assessment of risks to development outcomes, a methodology that rates the progress towards the achievement of development objectives and a clear measurement of progress towards target outcomes and outputs. A number of these measures have now been adopted by Management under the QA Action Plan. PCRs/XSRs templates should be reduced in size and focused on items that require management attention with a view to make learning lessons more relevant and usable. The selection of previously-used consultants for PCR/XSR preparation should be vetted against former PCR quality scores. The on-going revision of the Operational Manual should ensure that the new guidance encompasses the adopted changes. (F.12; F.13; F.17)

6. **Continue to emphasize the learning outcome of the SESP but as a complement to other types of inputs that may be better suited to meet the Bank's demand for knowledge and best practices.** Rather than trying to fit the same instrument (such as the PCR/XSR) to achieve multiple objectives (accountability and learning), it may be more appropriate to develop a distinct learning approach that has the SESP as one of its inputs but relies on more specific and adapted venues, instruments and products, going beyond individual projects. Learning should remain an important outcome of the SESP but it should rely on a repository of information and data, and more dedicated and diversified sources of knowledge such as thematic reviews, impact evaluations, multi-country analysis, etc. Regular learning and "share" events, possibly jointly with BDEV, should be organized by the relevant Departments. Such events should be perceived as a safe space environment with its own rationale, different from what an accountability outcome would require. (F.15 – F.19)

## Annex 1: Terms of Reference

AFRICAN DEVELOPMENT BANK

# Consultancy Services to Conduct an Evaluation of the Self-evaluation Systems & Processes of the African Development Bank

### I. Introduction

The Independent Development Evaluation Department (IDEV) of the African Development Bank Group (hereafter "the Bank") requires the services of a consultancy firm (hereafter, “consultant”) familiar with International Financial Institutions’ operations, and monitoring and evaluation systems and processes to evaluation its self-evaluation systems and processes. The evaluation aims to assess the relevance, coherence, effectiveness, efficiency and contribution of the Bank’s self-evaluation systems and processes. The assignment will be conducted under the general supervision of an IDEV Task Manager.

### II. Context

The Bank has both independent evaluation and self-evaluation systems and processes, which are mutually dependent. These systems and processes help the Bank to account for its investment effort, to learn from its experiences, and to monitor and improve its performance. The mandate of independent evaluation resides with IDEV, while that of self-evaluation rests with Bank management including operational complexes.

The Bank’s self-evaluation systems are defined in various Bank documents including the Operations Manual (OM), and policies. They are multi-purpose including monitoring, and judgement of development results, accountability for and learning from development implementation and results. They concern various aspects of the Bank including policies, strategies, programs, projects, processes and systems. They depend on multiple actors, guidelines, processes and tools including the results measurement framework (RMF) at corporate, program and project levels, and the additionality and development outcomes assessment (ADOA) framework. The RMF and ADOA framework are core to the Bank’s self-evaluation systems and processes.

In responding to the changing contexts within and outside the Bank, the Bank’s self-evaluation function has evolved over time. The Bank adopted its operations manual (OM) in 1993, and revised it in 1999, and then in 2014. The revision of the 2014 OM will start in 2019. Bank replaced in 2013 its Medium Term Strategy (MTS) with a Ten Year Strategy (TYS), 2013-2022. The Bank, since adopting its TYS in 2013, has gone through major organizational restructuring and adjustments in 2014 and 2016, and changes in policies, and operational and institutional processes including:

- The adoption in 2015 of the High5 priorities within the context of the TYS, leading the development of appropriate strategies for each of the High5s.
- The development and adoption of the New Development and Business Delivery Model (DBDM) in support of the High5s; comprising five major pillars.

- Creation of structures such as Delivery Accountability and Process Efficiency Committee (DAPEC), and Technical Quality Assurance Committee (TQAC) for improving the operational and institutional processes.

Certain aspects of the Bank's self-evaluation function have been the subject of evaluations/reviews. Some of these evaluations/reviews have been completed, and others are ongoing including:

IDEV's evaluations:

- Evaluation of the African Development Bank's quality assurance across the project cycle (2013-2017), September 2018. This evaluation covers project quality at entry, quality of supervision, quality at exit, and environmental and social safeguards.
- Evaluation of the Integrated Safeguards System (ongoing)
- Independent evaluation of the Bank's additionality and development outcomes assessment (ADOA) framework (2014)
- Synthesis Report – Second Independent Assessment of Quality at Entry in Public Sector Operations (2013). Independent Evaluation of Quality at Entry for ADF-11 Operations and Strategies African Development Bank Group (2010)
- Project Supervision at the African Development Bank (2001-2008) – An Independent Evaluation (2009)
- Country strategy and program evaluations, which include aspects of self-evaluation systems, and processes.

Other evaluations/reviews of the Bank include:

- The ongoing independent audit of Bank results monitoring and reporting (RMR), which focuses on adequacy and compliance of Bank policies, procedures, organizational arrangements, monitoring and reporting frameworks, and data management.
- An assessment of the Bank's quality assurance tools (2018).
- The diagnostic study of AfDB's practices to assure the quality at entry of public sector operations (2018).
- Multiple Mid-term reviews, and completion reports at strategy, program and project levels.

## III. Evaluation purpose, objectives, scope and questions

### a) Purpose and objectives

The purpose of the evaluation is to support the Bank's management and operational staff in:

- Improving self-evaluation, and performance management of operations, strategies, and policies;
- Improving the relevance and quality of the Bank's operations manual whose revision is to start in 2019;
- Promoting learning from experience, and operational effectiveness;
- Enhancing the implementation of the New Development and Business Delivery Model (DBDM), and process engineering;
- Accounting to the Board of Directors and other stakeholders for the results of the investments in the Bank self-evaluation systems and processes;

The evaluation will also build on and complement the:

- IDEV's Evaluation of the African Development Bank's quality assurance across the project cycle (2013-2017), and Evaluation of the Integrated Safeguards System.
- Independent audit of Bank results monitoring and reporting (RMR)

The evaluation's intended users are primarily staff and Board members of the Bank. The inception phase of the evaluation will clearly define the intended users and uses of the evaluation.

The objectives of the evaluation are to:

- Assess how well the Bank's self-evaluation systems and processes performed, focusing on their relevance, coherence, efficiency, effectiveness, and short-term impact;
- Assess the enablers and barriers that affected the design, implementation and results of the Bank's self-evaluation systems and processes;
- Distil lessons, good practices, and recommendations to enable the Bank to enhance the quality and performance (design, scope, implementation and results) of its self-evaluation systems and processes;

### b) Scope and questions

The evaluation will focus on the Bank's self-evaluation systems and processes during the period 2013-2018, which covers a substantial part of the implementation of the TYS, and recent institutional reforms including the DBDM and process engineering. This period is also that of the 2014 OM.

The evaluation will only cover self-evaluation systems and processes for sovereign and non-sovereign development. It will not deal with personnel self-evaluation systems and processes. The evaluation will mainly use the OECD-DAC criteria of relevance, effectiveness, and efficiency, and focusing on results beyond outputs. It will also assess coherence within the Bank's self-evaluation systems and processes, and with the Bank's independent evaluation function. The evaluation will be forward looking, and its key issues include:

(i) the self-evaluation systems' stakeholders, and their perspectives and inter-relationships;
(ii) the components, processes and mechanisms of the self-evaluation;
(iii) the enabling environment for, and barriers to self-evaluation;
(iv) appropriate data collection, analysis and storage tools and systems for self-evaluation;
(v) the use of self-evaluation outputs especially for decision-making, learning, accountability and for improving development quality and effectiveness.

The key evaluation questions, presented in the table below, are indicative. They will be refined and finalized during the inception phase of the evaluation.

| Criteria | Questions |
|---|---|
| How well are the Bank's self-evaluations systems and processes performed? | |
| To what extent are the self-evaluation systems and processes relevant and coherent? | |
| Relevance & coherence | To what extent are the parts (e.g. structures; instruments; processes; methods; mechanisms) of the self-evaluation systems complete, relevant and credible? |
| | To what extent is the theory of change for the self-evaluation systems explicit, complete, embedded, relevant, and credible? |
| | To what extent are the designs of the self-evaluation systems and processes adequate? |
| | To what extent are the self-evaluation systems and processes aligned with the bank key policies, strategies (H5s; TYS) and business model (DBDM)? |
| | To what extent are the self-evaluation systems internally coherent, and coherent with the Bank's independent evaluation function? |

| | |
|---|---|
| | How well the self-evaluation systems and processes mainstreamed cross-cutting issues; e.g. gender/inclusivity; safe guards; fragility, |
| To what extent are the self-evaluation systems and processes efficient, effective and impacting on development quality and organizational learning? | |
| Efficiency | How efficient are self-evaluation systems and processes in producing desired outputs; credible evidence/information? (are the self-evaluation systems' results cost effective?) |
| | How timely are self-evaluation systems and processes in producing desired outputs; evidence/information? |
| | To what extent has the Bank deployed adequate resources including human in self-evaluation and processes? How efficiently were these resources used to achieve the planned results? |
| | To what extent the Bank invested in good and/or innovative practices in self-evaluation processes, mechanisms, tools and methods? |
| Effectiveness | To what extent are the self-evaluation systems and processes producing relevant, reliable, timely and useful outputs? |
| | To what extent are the self-evaluation systems and processes' outputs used for (i) informing decision-making (e.g. programming/improvement; policy, strategy, & program/project design/implementation); (ii) accountability; (iii) learning? |
| Impact/contribution | What is the evidence of the contribution of the self-evaluation systems and processes to development quality (project; strategy; policy), outcomes, and learning? |
| | What contribution have the self-evaluation systems and processes made to corporate development effectiveness and organizational learning? |
| Enablers and barriers | What are the factors that have enabled or constrained the design, implementation and results of the self-evaluation systems and processes? |
| Lessons and recommendations | What relevant lessons, good practices, and recommendations that can be drawn for improving the quality and performance of the self-evaluation systems and processes? |

## IV. Methodology and processes

The IDEV evaluation policy and the Evaluation Cooperation Group's Big Book on Evaluation Good Practice Standards[41] will guide this evaluation. The evaluation approach will require a reconstitution of the supposed theory of change, underlying the Bank's self-evaluation systems and processes. The supposed theory of change will guide the refinement of the indicative evaluation questions, and the development of the evaluation methodological framework. The inception phase of the evaluation will clearly define and detail the most credible methodological framework for responding to the evaluation questions. The

[41] ECG Big Book on evaluation good practice standards, http://www.ecgnet.org/document/ecg-big-book-good-practice-standards. Both documents reflect the standard OECD-DAC development evaluation criteria and quality standards.

methodological approach should be of mixed designs and methods. The data sources, the basis for the evaluation streams of evidence, should include but not be limited to the following:

- Desk review of relevant documents/reports and databases including those of the Bank, other MDBs, and the literature.
- Substantive interviews and discussions with key stakeholders within and outside.
- Staff survey (staff; RMC official).
- In-depth case studies, based on appropriate sampling.
- Benchmarking with other MDBs and other appropriate agencies.

The evaluation process will include the following phases:

- Inception phase to produce the inception report, which will include the full evaluation methodology (including sampling, evaluation matrix, limitations, risks and mitigations, data collection and analysis tools/instruments, rating scale and standards), evaluation team composition and responsibilities for each of the individual evaluation team members. This will involve inter-alia desk reviews and discussions with key stakeholders, rapid assessment of available data, reconstruction of the supposed theory of change, stakeholder mapping and preparation of the inception report.
- Meta-analysis phase covering (i) evaluations on self-evaluation systems; (ii) Bank policies; strategies, programs/projects, and management action record system (MARS). This phase will overlap with the inception phase.
- Data collection and analyses for the generation of findings, and drawing of conclusions, recommendations and lessons learned: This phase will concern all the data sources, highlighted above. It will be the basis for the preparation of the background reports, and the evaluation synthesis report. Emerging findings will be shared with stakeholders for feedback.
- Synthesis, report writing and feedback leading to the draft evaluation synthesis report and its presentation to the evaluation Reference Group (defined under the quality assurance section below), and other stakeholders for feedback on the draft evaluation findings.
- Production and delivery of the final evaluation report in the appropriate format (in English) for dissemination and follow up.
- Communication and dissemination of evaluation results

**Risks and mitigation actions**: The evaluation risks and mitigation actions will be identified at the inception phase by the evaluation team.

Available **documents and databases** including the following: Bank strategies, policies, project databases, MARS, results measurement frameworks, annual reports, work programmes, progress reports, and self-evaluation guidelines and reports, budget reports, relevant IDEV evaluations/reviews.

## V. Deliverables and timeline

The consultant will deliver the following outputs (in English):

- Inception report (draft and final)
- Background reports (Benchmarking; meta-analysis; case study; survey…)
- Draft evaluation report and its presentation to the evaluation reference group, and for peer review; the evaluation report will include an executive summary, background and context, evaluation purpose, objectives and questions, key aspects of the methodological approach and limitations, findings, conclusions, lessons and recommendations, and annexes
- Final evaluation report including an executive summary of up to two pages and essential annexes
- Technical annexes including the methodology and its instruments and evidences.

- Electronic version of data collected and evidence set (analyzed data)

The evaluation will have an indicative duration **of 280 person days** over a period of five months, and its timeline is presented in the table below. The final evaluation report is expected to be completed and delivered in June 2019.

| Evaluation phase-Delivery-Timeline | | |
|---|---|---|
| **Phase/Output** | **Deadline** | **Responsibility** |
| Contracting phase | December 2018 | IDEV |
| Inception phase:<br>• Draft inception report<br>• Comments on draft report<br>• Final inception report incorporating comments<br>• Approval of inception report | Week 1 February 2019 | Consultant & IDEV<br>Consultant<br>IDEV<br>Consultant<br><br>IDEV |
| Data collection & analysis phase (including literature/document review, interviews, survey meta-analysis, benchmarking, and case studies) | Week 3 April 2019 | Consultant |
| Reporting phase:<br>• Draft & revised background reports<br>• Draft evaluation report<br>• Presentation of draft findings to RG for feedback<br>• Final report incorporating comments/suggestions<br>• Feedback on evaluation process | <br>Week 1 June 2019<br><br>Week 2 May 2019<br>Week 2 May 2019<br>Week 3 May<br>Week 1 June<br><br>Week 1 June | Consultant |
| Communication & dissemination phase: | From February- 2019 | IDEV |

## VI. Profile of the Evaluation Team (qualifications, experiences and competencies)

A firm (the consultant) will undertake the evaluation using a balanced team with demonstrated professional knowledge, skills and experience in:

- Evaluation/review/synthesis/benchmarking theories/practices
- Evaluation/monitoring and evaluation systems and processes in international development
- International development work and issues especially within the contexts of Africa and MDBs.
- How the MDBs work, and MDB activities including evaluation functions
- Evaluation report writing and presentation

- Fluency in English, and working knowledge of French; at least one evaluation team member being fluent in both English and  French will be an advantage
- Standard applications and analytical packages

## VII. Management and Quality Assurance Arrangements

An IDEV Task Manager will be responsible for: (i) providing overall guidance to the consultant, and approval of the evaluation process and outputs (inception report; background reports, draft and final evaluation reports); (ii) quality assurance process including the external peer review of the key evaluation products, and receiving comments from the Evaluation Reference Group (ERG); (iii) recruiting the consultant (iv) briefing the consultant; (v) establishing the ERG; (vi) receiving from the consultant all data, files (including raw data, coded data, interview notes, databases) that will be produced; (vii) communicating to the Bank's Management and Board of Directors, and disseminating the final evaluation results to the key stakeholders. IDEV will also recruit at least two competent and experienced international experts (content-area; evaluation) for the external peer review of the evaluation process and outputs; (viii) ensuring the payment of the consultant.

The **evaluation reference group (ERG**) will comprise selected Bank staff from the relevant complexes/Departments/Units. The ERG will review and comment on the evaluation process and outputs (inception report; evaluation reports), and provide a sounding platform for rapid feedback especially on the evaluation plan (including design and methods) and emerging evaluation findings.

## VIII. Evaluation Budget

The evaluation budget will comprise all expenses including fees, travel and taxes. The firm/consultant will provide a detailed budget with breakdown against activities and key milestones.

## Annex 2: Evaluation Matrix

| Relevance & Coherence | | | |
|---|---|---|---|
| | Evaluation Questions | Assessment Criteria and Indicators | Source |
| **Performance Management** | • **EQ1**. Is the SESP and its different instruments relevant and coherent on how it is aligned with the Bank's strategy and guidance and serves its purpose for performance management<br><br>[*To what extent there is evidence that the implementation of the SESP results in enhanced performance of project/country programs*] | ➢ Extent to which the set of policy and guidance documents describe a coherent process for the SESP, with added value by each instrument, clear roles and responsibilities, coherent and objective description of the rating process.<br><br>Main instruments to be analysed are:<br>i) Projects and project portfolio – IPRR/XSR (for non-sovereign operations), MTR, PCR; CPPR and APPR. Aide Memoires and BTORs will also be considered<br>ii) Country/regional strategies – CSP/RISP-MTR, CSP/RISP-CR<br>iii) Sector/thematic strategies – "Reviews"<br>➢ Extent to which the set of SESP instruments is aligned with the main policy and guidance documents:<br>i) Operational Manual, Presidential Directives and Delegation of Authority Matrix (DAM), DBDM, H5s, TYS, etc…<br>ii) Focus on priority areas: gender, fragility, safeguards, fiduciary, governance<br>➢ Extent to which the implementation of the SESP and its instruments comply with the main guidance documents<br>➢ Extent to which the provided ratings system reflects a consistent, timely and accurate approach to performance management for the various processes around each instrument<br>➢ Extent to which the SESP is used as a tool for corrective action and analysis of possible impediments | Desk-based review<br><br>Stakeholders survey, consultation, and focus group discussions (staff and managers; HQ and field)<br><br>Recent Evaluation reports, quality at entry, quality of supervision, CEDR, PCR and XSR evaluation notes, cluster evaluations<br><br>Case studies, process reviews<br><br>Benchmarking analysis |
| **Accountability** | • **EQ2**. Is the SESP and its different instruments a | ➢ To what extent the SESP relies on strong M&E systems as a critical input for | Desk-based review |

| | | | |
|---|---|---|---|
| | reliable and relevant framework geared towards verification of results, accountability and reporting at:<br>- staff/management level<br>- corporate level<br><br>*[does the SESP generate useful information for staff, management and the Board signalling that the AfDB holds itself accountable for achieving results]* | the credibility of the system<br>➢ To what extent are the different policy and guidance documents describing the SESP adequate to verify achievements of results and lines of accountability<br>➢ To what extent is the SESP internally coherent and consistent with the Bank's independent evaluation function<br>➢ To what extent the SESP provides a broad and relevant perspective on the results achieved and communicate overall performance in an easily understood way<br>➢ To what extent is staff and management being held accountable for the proper implementation of the SESP towards its intended results-based objectives.<br>➢ To what extent the SESP and its different levels of aggregation provides a relevant and accurate reporting of results through the RMF and other reporting tools (dashboard, etc..) | Stakeholders survey, consultation, and focus group discussions (staff, managers, and Executive Directors; HQ and field)<br><br>Recent Evaluation reports, quality at entry, quality of supervision, CEDR, PCR and XSR evaluation notes, cluster evaluations<br><br>Benchmarking analysis |
| **Learning** | • **EQ3**. Is the SESP being used as a reliable and relevant framework for learning and innovation<br><br>*[is the information produced being used to shape how learning takes place]* | ➢ To what extent are processes, roles and tools well-defined, comprehensive and integrated for learning and knowledge management<br>➢ To what extent are comprehensive scoring and IT systems in place to ensure learning from relevant project cycle and thematic activities<br>➢ Is the SESP providing the most relevant information to the different audiences<br>➢ To what extent is the SESP used as a tool for learning, and a repository of evidence of good practices, failures and lessons learned. | |
| **Efficiency** | | | |
| | Evaluation Questions | Assessment Criteria and Indicators | Source |

| | | | |
|---|---|---|---|
| **Performance Management** | • **EQ4**. To what extent the SESP and its instruments provide a reliable and effective framework for portfolio management<br><br>*[Is the SESP implemented in an effective manner in terms of time, financial and human resources to deliver its intended benefit]* | ➢ Are the costs of the Bank's SESP across the project cycle appropriate relative to the results achieved<br>➢ To what extent are the resource requirements comparable to the frameworks of other organizations<br>➢ To what extent the elimination of possible overlaps, redundancies or requirements could improve cost-efficiency<br>➢ To what extent the different instruments of the SESP are adding value compared to their costs<br>➢ To what extent is the budget or staff overload a factor in the proper implementation of the SESP as described<br>➢ To what extent is proactivity by staff for corrective action promoted or impeded by transaction costs, project restructuring procedures, etc… | Desk-based review<br><br>Stakeholders survey, consultation, and focus group discussions (staff, managers, HQ and field)<br><br>Recent Evaluation reports, quality at entry, quality of supervision, CEDR, PCR and XSR evaluation notes, cluster evaluations<br><br>Case studies<br><br>Benchmarking analysis |
| **Accountability** | • **EQ5**. Does the SESP provide a reliable and cost effective framework for reporting and accountability internally and externally<br><br>*[is the money spent for the SESP commensurate to the quality and the usefulness of the information provided]* | ➢ To what extent quality at entry is a factor in the implementation of a cost-effective SESP system for accountability and reporting<br>➢ To what extent are M&E systems a factor in the establishment of a credible accountability framework<br>➢ To what extent reporting requirements are adequate and cost-effective in providing the right information to the different audiences<br>➢ To what extent users find data input costly in terms of time compared to the benefits<br>➢ To what extent are the ratings and their aggregation being used as a cost-efficient tool for corporate performance reporting<br>➢ To what extent the way aggregation takes place for corporate reporting is adequate and cost-efficient | |

| | | | |
|---|---|---|---|
| **Learning** | • **EQ6**. Is the SESP being implemented as a cost-effective tool for learning<br><br>*[is the money spent for the SESP commensurate to the amount of learning and innovation that it generates]* | ➢ To what extent templates and instruments support efficient recording of lessons and add most value<br>➢ To what extent SESP and rating validation could be a factor in increasing the opportunities for learning<br>➢ To what extent the SESP is being used to record and gather data for learning purposes, events and knowledge management<br>➢ To what extent the SESP is trusted enough to be a source of learning and what could be the role of independent evaluations | |
| **Effectiveness & Impact Contribution** | | | |
| | Evaluation Questions | Assessment Criteria and indicators | Source |
| **Performance Management** | • **EQ7**. Is the SESP architecture being implemented as a tool to enhance portfolio performance and the achievement of results<br><br>*[What is the evidence that the SESP contributes to improve quality at exit]* | ➢ To what extent is the system designed to provide the right balance between ensuring compliance and pursuing results<br>➢ To what extent the SESP guidance is geared towards promoting proactivity in addressing issues and corrective actions<br>➢ To what extent the inputs in the SESP (quality at entry, M&E, processes, instruments, budget) are adequate to ensure proper delivery of the SESP<br>➢ To what extent the implementation of the rating system is sufficiently trusted to be a reliable tool for performance management<br>➢ To what extent has the SESP facilitated the implementation of mitigation and safeguard measures<br>➢ To what extent has the SESP contributed to effectively addressing cross-cutting issues such as gender, fiduciary, fragility | Desk-based review<br><br>Stakeholders survey, consultation, and focus group discussions (staff, managers, HQ and field)<br><br>Recent Evaluation reports, quality at entry, quality of supervision, CEDR, PCR and XSR evaluation notes, cluster evaluations |

| | | | |
|---|---|---|---|
| **Accountability** | • **EQ8**. Is the SESP architecture being implemented as a tool to enhance accountability consistently with provided guidance and Bank's priorities<br><br>*[Is the degree of accountability generated by the SESP conducive to improve results]* | ➢ Are roles and responsibilities sufficient clear in the definition of the different processes and requirements surrounding the preparation, conduct, review, sign-off, follow-up for the various steps of the different SESP instruments<br>➢ Is the design and guidance around corporate reporting processes effective and well integrated in the SESP system<br>➢ Has the Bank's SESP been delivered as expected for accountability purposes<br>➢ To what extent has enforcement of procedures been enacted upon when needed<br>➢ To what extent the implementation of the rating system is sufficiently trusted to be a reliable tool for accountability and reporting requirements<br>➢ To what extent is attribution a factor in determining the degree of the Bank's accountability | Case studies<br><br>Benchmarking analysis |
| **Learning** | • **EQ9**. Has the SESP contributed to the identification and use of lesson learned<br><br>*[is the SESP, the way it is designed and implemented, the right tool for learning]* | ➢ To what extent the focus on accountability can undermine the relevance and usefulness of the SESP for learning.<br>➢ To what extent independent or arm-length evaluation can play a role in enhancing learning opportunities<br>➢ To what extent the concerns over ratings and disconnects distract from learning<br>➢ To what extent the SESP can be used more strategically to meet knowledge gaps and for lesson learning | |
| **Incentives and barriers** | | | |
| | Evaluation Questions | Assessment Criteria and indicators | Source |
| **Performance Management** | • **EQ10**. Are the incentives in place conducive to candid assessments and proactivity for portfolio performance and | ➢ To what extent a compliance mindset and a focus on ratings can distort the systems and create biases or a candour gap.<br>➢ To what extent guidance provided can enhance incentives for results and corrective action rather than compliance and transaction costs | Desk-based review<br><br>Stakeholders survey, consultation, and focus |

| | | | |
|---|---|---|---|
| | corrective action<br><br>*[Do the incentives in place ensure that the SESP is implemented as designed for quality results]* | ➢ To what extent can team or third-party validation help the system to remain honest<br>➢ To what extent is evidence available to justify the ratings and what is the risk of "gaming the system"<br>➢ To what extent is poor project performance seen as having a link with staff performance potentially leading to risk aversion and fear of tainting staff reputation<br>➢ To what extent are the incentive in place reflecting the right balance between the focus on lending vs. supervision | group discussions (staff, managers, HQ and field)<br><br>Recent Evaluation reports, quality at entry, quality of supervision, CEDR, PCR and XSR evaluation notes, cluster evaluations<br><br>Case studies<br><br>Benchmarking analysis |
| **Accountability** | • **EQ11**. Are the incentives in place conducive to candid assessments for accountability<br><br>*[Can incentives influence behaviours so that accountability is towards the achievements of results rather than the compliance with rules]* | ➢ To what extent should the adequate implementation of the SESP be linked to staff performance evaluation<br>➢ To what extent greater focus on the inputs (quality at entry, budget, business processes, M&E systems) is likely to improve the incentive structure in place for effective implementation of the SESP<br>➢ To what extent is Management exerting leadership over the correct implementation of the system and lines of accountability<br>➢ To what extent staff and managers see the SESP as a relevant accountability tool that requires trust and close follow up for it to be useful | |
| **Learning** | • **EQ12**. Is the incentive structure geared towards the use of the SESP for continuous learning and innovation<br><br>*[Do the incentives in place ensure that the SESP* | ➢ To what extent new mechanisms for lessons sharing, "safe space" for debate and incentives for transparency would enhance the credibility and relevance of the SESP for learning purposes<br>➢ To what extent learning and feedback loops can lead to system improvement in enhancing flexibility, better procedures for project restructuring, recognition for excellence, differentiation according to specific situations (fragile context)<br>➢ To what extent a more direct involvement of third independent parties and | |

| | | | |
|---|---|---|---|
| | *serves as a learning tool]* | specific thematic/country events are likely to increase the credibility of the system and learning opportunities<br>➢ To what extent incentives can change behaviours in terms of documenting learning evidence of proactivity for corrective action, best practices of ratings follow up, awards for innovation, etc..<br>➢ To what extent better incentives can increase the perceived value of the knowledge created, mainstream risks and failures as part of the business, create opportunities for mining lessons and knowledge and organizational learning<br>➢ To what extent the focus on corporate results reporting for accountability and ratings contributes to weaken staff attention for other purposes such as learning | |

## Annex 3: Main features of an Integrated Evaluation Policy[42]

1. One issue that came up repeatedly during the review of other Development Banks was that they lack an overall evaluation policy, and that this absence negatively affects a number of areas. AfDB has an evaluation policy for IDEV rather than an evaluation policy for the AfDB. There are important differences. IDEV's primary role is to generate evaluation findings. As a secondary role, IDEV promotes the use of evaluation findings, but it is not a user of evaluation findings — this is the role of the Board and Management. Given this, an evaluation policy for IDEV necessarily focuses on the supply side of evaluation and not the demand and use sides.

2. Evaluation in multilateral development banks (MDBs) is a complex function covering separate but closely intertwined roles for the Board, management, and the Independent Evaluation Departments (IED). The parties involved devote substantial resources to the function. Evaluation adds value when there is clarity and consensus about the strategic purpose of evaluation and how that purpose can best be achieved. Evaluation adds value only when its findings are used, so an evaluation policy that focuses on the supply side will likely fail to cover the essential requirements for value addition. Further, IEDs are not the only generator of evaluation findings. All completed operations and country/regional strategy papers (CSPs/RISPs) are self-evaluated by staff.

3. An evaluation policy is the means by which the parties involved can create clarity and consensus about the evaluation's purpose and the ways and means to achieve it. When the evaluation policy does not fully cover the role of evaluation for the whole institution it could be revised as described below. The evaluation department is normally not part of the decision-making apparatus, thus, a policy for IEDs alone fails to establish the strategic importance and purpose of a very large part of evaluation in the MDB.

4. Some MDBs (the WBG and IFAD) have strengthened their evaluation policy to include all evaluation practices in all departments, and including the supply side as well as the demand side. AfDB could also strengthen its evaluation policy in a number of areas:

- establishing the strategic purpose(s) that evaluation should accomplish, particularly with a view to using evaluation to increase AfDB's relevance and add to AfDB's value now and in the future. Evaluation is the vehicle for accountability. To maximize the strategic purpose of evaluation, AfDB and IDEV could re-examine the meaning of accountability in light of current and future needs for accountability. This will help ensure that the tools deployed are fit for their purpose, i.e., to effectively and efficiently deliver meaningful assessments for the purpose of accountability, its contribution to learning, and any larger strategic goals.
- refining the identification of outcomes and their measurement in the RMF. Although IDEV is independent of management, it is part of AfDB, so it should demonstrably contribute to the achievement of the organization's expected results as contained in its corporate results framework. There should also be a clear articulation of how it will do so. For effective results-based management of the evaluation function, the evaluation policy would require Board, Management, and IDEV to agree on a set of outcome indicators against which the performance of the AfDB-wide evaluation system, and IDEV within this, would be tracked. Those outcome indicators should link to the corporate results framework. However, this should not preclude IED from raising issues about the corporate strategy and results framework where appropriate; this is implicit in its organizational independence.
- The policy should require that after the parties (the Board, Management, and IDEV) agree on the expected outcomes, they should also agree on what needs to be done to achieve these

[42] This section is inspired by the External Review of the AsDB and the World Bank Group and internal discussions.

outcomes and map these outputs and outcomes in a theory of change. The aim is to establish the basis for rational decision making about such matters as the mix of evaluation products, the amount of effort that should be directed toward knowledge management, evaluation capacity building, and meeting accountability requirements. IDEV's outputs will depend upon the outcomes it seeks.

5. A revised policy would:

- Cover evaluation throughout AfDB, not just in IDEV. Be readable and succinct and clearly describe the strategic purposes and outcomes that evaluation in AfDB should accomplish, and in particular i) define how evaluation contributes to accountability and learning in terms of desired outcomes; ii) clarify desired outcomes in developing evaluation capacity, within AfDB and externally; iii) realign and harmonize methodologies and ratings definitions between Management and IDEV; and iv) require that the agreed outcomes link to AfDB's corporate results framework;
- Describe the roles of management, including self-evaluation; ensuring that new projects, policies, and strategies are evaluable; demonstrably incorporating findings from evaluation in new projects, policies and strategies; and responding to and acting on evaluation reports.
- Describe the roles of the Board in overseeing IDEV, using evaluation findings in its own decision-making and giving clear direction to management;
- Describe the roles of IDEV including its lead role in establishing evaluation methodologies; preparation of multi-year evaluation work programs; preparing results frameworks and theories of change for evaluation in AfDB; and preparing medium to long-term staffing and staff development plans;
- Clarify the meaning of independence of IDEV, particularly in terms of the appropriate nature and degree of engagement by IDEV with others in AfDB; and
- Outline special provisions for IDEV regarding access to information, approval of IDEV documents, internal distribution and external disclosure of IDEV reports, appointment rules, budget, and staffing decisions.

6. To successfully develop an AfDB-wide evaluation policy, Management, the Board, and IDEV all need to be vitally engaged in drafting the policy through a highly engaged process designed to generate consensus on what all of the parties involved want out of evaluation. The process requires a commitment of time for a series of reflections and discussions on the strategic purposes that evaluation should serve—including refreshing and redefining the meaning of accountability and learning in more tangible and specific ways in light of current and future changes in the context—and defining those purposes in terms of a series of outcomes that can be measured and reported on.

7. Moreover, a participatory process to reach agreement on the essential features of evaluation in AfDB should be an effective way of refreshing the relationship between IDEV and management. As they develop the new policy, the dialogue and engagement between IDEV and Management offer an opportunity for them to develop a constructive and collegial relationship, while preserving the essential features of independence and objectivity.

## Annex 4: Evaluation Cooperation Group (ECG) Practice Note

In its: "Self-evaluation in ECG member institutions" (November 2018), ECG laid out a number of recommendations and features of self-evaluation systems that are likely to contribute effectively to wider institutional performance. They are summarized below.

**Starting points:**

- Self-evaluation by responsible management/operational teams is an important feature of the evaluation systems of most institutions represented in the ECG, and is a central element of management's specific responsibilities for well-functioning evaluation systems.
- Self-evaluations can simply be reviewed by the Independent Evaluation Department (IED) for completeness, compliance with agreed methodology, presence of evidence, etc, for advisory purposes without additional independent analytical work or validation.
- Validation is a higher level of IED assessment, and is generally a desk-based assessment looking in particular at the sufficiency of evidence and analysis, relevant determinants of performance - offering an independent analysis. Where self-ratings are used, validations will ordinarily provide comparable IED ratings.
- IEDs have a vital role in helping boards and management to ensure that self-evaluation is credible, value-adding and effective.
- Self-evaluation is essentially a management responsibility under Board oversight. It may be conducted on the full range of institutional operations, activities, structures, processes, policies and strategies, and employ a variety of different methods.
- If self-evaluation is to be useful it must be an integral element of and used by wider "results" systems. To this end, self-evaluations should meet clear standards for quality, relevance and timeliness; management-level ownership and responsibilities should be clear and well-integrated; work should be visible and widely accessible.
- A dedicated Management unit with assigned responsibility for self-evaluations at an institutional level would reinforce system effectiveness and integrity, can provide quality assurance and be a focal point for technical assistance and information dissemination.
- Self-evaluations may be done by directly responsible operational team(s), operational team members not directly involved (such as new staff), or consultants.
- Management should determine the operational coverage of self-evaluation in accordance with the evaluation policy or practice and with input from the IED.
- Self-evaluation may embody accountability or learning, or a combination.

**Recommended Features**:

- The major components of self-evaluation systems (scope, methods, timing, data content, outputs, and definitions) should follow from and reflect those used by IEDs.
- Self-evaluation methods should incorporate an articulated theory of change even if none had been explicitly presented for the matter being self-evaluated.
- Baseline circumstances targeted in the operation should be effectively captured in the self-evaluation, along with all relevant targets and benchmarks – initial and modified. The adequacy of collected/available data should be considered explicitly.
- Self-evaluation systems should be subject to periodic assessment.
- Management should ensure that self-evaluation systems have high internal circulation, visibility and accessibility across units for wider learning and awareness.
- Internal confidentiality should be minimized.

- Self-evaluation should seek to assess all factors contributing to performance. This encompasses all stated or implied objectives of the subject operation or issue as well as the possible effects of unintended or unanticipated developments.
- Rating systems must be rigorous and consistent but their application flexible. The use of formal performance ratings in self-evaluation is context-sensitive.
- Rating systems should reflect ECG Good Practice Standards and provide a consistent methodology based on a theory of change, a plausible and definable basis for ratings, and the capacity to capture innovation and exogenous factors.
- Self-evaluation systems may elect not to apply ratings in some instances and should be sufficiently flexible to accommodate management and evaluator judgement.
- Policies, strategies and initiatives not related to investments should be reviewed as needed, drawing from self-evaluation. It is desirable for an independent evaluation to be completed upstream so as to inform management and the Board of revisions, otherwise some form of review of management's self-evaluation should be completed.

## Annex 5: Comparators Review

**Table 12: Public Sector Lending**

| | ADB | World Bank | IFAD | AfDB |
|---|---|---|---|---|
| 1. Is there a corresponding validation instrument (VI)? | Yes | Yes | Yes | Yes |
| 2. Do SE and VI use same harmonized methodology? | Yes | Yes | Yes | Yes |
| 3. Does the VI also include ratings not in the SEI? | Yes: Quality of SEI | Yes: Quality of SEI | No | Yes: Quality of SEI,<br>M&E performance |
| 4. Does VI involve any field mission? | No | No | No | Yes<br>Roughly 10% |
| 5. Are there other field-based project evaluations or validations? | Yes: Project performance Evaluation Report (PPER) | Yes: Project Performance Assessment Report (PPAR) | Yes: Project Performance Evaluation (PER) | Cluster evaluations of projects (learning purpose) |
| 6. SEI for all projects or using sampling?? | 100% | 100% | 100% | 100% |
| 7. VIs for all SEIs? | Yes | Yes | Yes | Yes |
| 8. What is key rating? | Overall Performance | Project Outcome | Project Performance &<br>Overall Project Achievement | Overall Performance |
| 9. How calculated? | Based on the assessments of relevance, effectiveness, efficiency, and sustainability, weighed equally on a four-point scale. | Derived through guidance from Relevance of objectives, Efficacy in achieving each objective, Efficiency | <u>Project Performance</u>: Average of ratings for Relevance, Effectiveness, Efficiency, and Sustainability<br><br><u>Overall Project Achievement</u>:<br><br>Overarching assessment, drawing on nine dimensions. | Average of sub-ratings and averages of 4 dimensions: Relevance, Effectiveness, Efficiency, & Sustainability.<br><br>Same weights |
| 10. Rating scale for overall rating? | Six-point scale:<br>HS, S, MS, MU, U, HU | Six-point scale:<br>HS, S, MS, MU, U, HU | Six-point scale<br>HS, S, MS, MU, U, HU | Four-point scale<br>HS, S, U, HU |
| 11. Use of rating ranges | Yes | No | No | Yes |
| 11. Are SEI and VI submitted to the Board? | Yes | Yes | Yes | Yes<br>VI synthesis Note |
| 12. Are SEI and VI disclosed to the public? | Yes | Yes | Yes, if government agrees | Yes |

## Table 13: Private (non-sovereign) Lending/Investments

| | ADB | IFC | AfDB |
|---|---|---|---|
| 1. Is there a self-evaluation instrument (SEI)? | Yes | Yes | Yes |
| 2. Is there a corresponding validation instrument (VI)? | Yes | Yes | Yes |
| 3. Do SE and VI use same harmonized methodology? | Yes | Yes | Only XSR and XSREN use harmonized criteria |
| 4. Does the VI also include ratings not in the SEI? | Yes: ADB Work Quality | Yes: IFC Work quality, Screening, Appraisal and Structuring, and Supervision and Administration | Yes: Work quality |
| 5. Does VI involve any field mission? | No | No, except for selected environmental and social aspects | No |
| 6. Are there other field-based project evaluations or validations? | Yes – Project Performance Evaluation Report (PPER) | Yes – Project Performance Assessment Report (PPAR) | Provided for in the guidelines but seldom conducted |
| 7. SEI for all projects or using sampling?? | For all projects | Sampling | PSR/ASR: all projects (but low compliance)<br>XSR: Sampling |
| 8. If using sampling – approach and percentages? | Purposeful sampling | Random 40% of mature operations, by IEG (but 100% for E&S aspects) | Purposeful Sampling (in consultation with portfolio manager) |
| 9. VIs for all SEIs? | Yes | Yes | In principle yes for XSRs, but large backlog |
| 10. What is overall or key rating? | Overall Performance | Development Outcome | No key rating |
| 11. How calculated? | Based on rating matrix, with no fixed weights, of Development Results, ADB Investment Profitability, ADB Additionality, and ADB Work Quality | Synthesis and not averaging of 4 components | Qualitative judgment drawn from underlying indicators ratings: Development Outcome, Investment Outcome, Work Quality, AfDB Additionality |
| 12. Rating scale for overall rating? | Four-point scale from Highly Successful to Unsuccessful | Six-point scale from Highly Successful to Highly Unsuccessful | Rating in XSR/XSREN on a 6-point scale. Ratings of sub-dimensions and other SEIs vary from 1-4, 1-5 and 1-6 |
| 13. Are SEI and VI submitted to the Board? | SEI are but VI are not. PPERs are submitted | No[43] | XSREN synthesis report and PPERs<br>Not in the RMF |
| 14. Are SEI and VI disclosed to the public? | Yes, if rules to preserve confidentiality are met | No | No |

[43] IEG will now try "clustered PPARs" that would be distributed to the Board and eventually to the public after removing information on projects' identity and client-specific data.

## Table 14: Country/Regional Integration Programs

| | ADB | World Bank Group | IFAD | AfDB |
|---|---|---|---|---|
| 1. Is there a self-evaluation instrument (SEI)? | Yes | Yes | Yes | Yes |
| 2. Is there a corresponding validation instrument (VI)? | Yes | Yes | Yes | No |
| 3. Do SE and VI use same harmonized methodology including for ratings? | Yes | Yes | Yes | Not Applicable |
| 4. Does the VI also include ratings not in the SEI? | Yes: ADB Performance, Borrower and executing agency performance | Yes – Quality of SEI | No | Not Applicable |
| 5. Does VI involve any field mission? | Yes – one mission | No | No | Not Applicable |
| 6. Are there other field-based program evaluations or validations? | Yes – Country Assistance Program Evaluation (CAPE) | Yes – Country Program Evaluation | Yes – Country Strategy and Programme Evaluation | Yes – Country Strategy & Program Evaluation (CSPE) |
| 7. SEI for all programs or using sampling? | 100% (either a validation or a CAPE) | 100% | 100% | Not Applicable |
| 8. VIs for all SEIs? | Yes | Yes | Yes | Not Applicable |
| 9. What is overall or key rating? | Overall Rating | Development Outcome | Overall Country Strategy & Programme Performance | No rating used for SEI<br>Ratings used for CSPE |
| 10. How calculated or estimated? | Scored from five ratings: Relevance, Effectiveness, Efficiency, Sustainability, Development Impacts | Based on ratings for individual objectives (rated by 4-point scale) using set criteria | Not an arithmetic average of the ratings of Relevance and Effectiveness | Not Applicable |
| 11. Rating scale for overall rating? | 6-point scale from Highly Successful to Unsuccessful | 6-point scale from Highly Satisfactory to Highly Unsatisfactory | 6-point scale from Highly Satisfactory to Highly Unsatisfactory | No ratings for SEI<br>4-point scale for CSPEs |
| 12. Are SEI and VI submitted to the Board? | Yes | Yes | Yes | Yes (SEI) |
| 13. Are SEI and VI disclosed to the public? | Yes | Yes | Yes | Yes (SEI) |

## Comparators Benchmarking

### The World Bank Group (WBG) Public Sector Projects

**1. Results Framework**

**Timely and appropriate adaptive management** requires building a culture of continuous improvement and problem-solving, based on evidence. Evaluation plays a key role in generating the evidence about what works in different contexts, and in identifying lessons for World Bank Group (WBG) stakeholders.

Evaluation Roles and responsibilities:

i) **Boards and Committee on Development Effectiveness**. CODE supports the WBG's Boards in assessing the development effectiveness of the WBG, monitoring the quality and results of operations, and overseeing or liaising on the work of the entities that are part of the Group's accountability framework. On behalf of the Boards, CODE also oversees the work of IEG as well as the adequacy, efficiency, and robustness of the WBG's monitoring and evaluation systems.

ii) **WBG Management.** This is responsible for fostering a culture of accountability and learning, and for providing appropriate incentives for staff at all levels. Management is also responsible for ensuring the availability of adequate resources for self-evaluation, learning, and knowledge management, and for formal review and follow-up on recommendations, including those agreed from independent evaluations carried out by IEG.

iii) **World Bank Group operational units.** Operational managers and staff are responsible for managing and conducting mandatory and demand-driven self-evaluations. Findings and recommendations from supervision processes, which are used in managerial monitoring of portfolio performance, ultimately feed into project and country program completion reporting (that is, mandatory self-evaluations). IEG validation of self-evaluation reports is a primary source of feedback on, and accountability for, the relevance, efficiency, and effectiveness of the WBG operational work program, as well as on the completeness, quality, and candor of self-evaluations.

iv) **Independent Evaluation Group**. IEG conducts independent evaluations of specific projects and programs, country programs, and a range of sectoral, thematic, and corporate issues. In addition, it reviews and validates individual project and country program completion reports from the World Bank, IFC, and MIGA. IEG's annual report on the *Results and Performance of the World Bank Group* summarizes lessons and recommendations that emerge from syntheses of portfolio-wide reviews of individual projects and country program performance. IEG and WBG management use the Management Action Record, a formal process of discussing actions to be implemented as a result of key recommendations from each major evaluation; progress on these actions is followed for a set number of years and reported to CODE. To promote learning from evaluations, IEG and operational units and management promote learning from evaluations through different types of learning activities.

2. **Evaluation Policy**.

A WBG-wide Evaluation Policy was issued in April 2019 by the managements of IBRD/IDA, IFC and MIGA and by the DGE (Director General, Evaluation) in response to earlier comments from an external panel reviewing IEG that the WBG needed an overarching evaluation policy because it "lacks a framework that outlines the principles, criteria and accountabilities for evaluation across the organization, that provides clarity to all staff on the merits of robust, high quality and credible evaluation, and that clearly delineates the respective roles of all parties." This is the first time that common principles for evaluation have been articulated for the whole Group. These include core principles for evaluation, and principles for selecting, conducting, and using evaluations. These aim to strengthen accountability and learning for evidence-based decision making and program improvement to enhance development results.

## 3. Self-evaluation and independent Evaluation Systems

**The WBG distinguishes among two main evaluation modalitie**s:

**Independent evaluation**s

i) They are conducted by IEG, which reports directly to the Boards. The target audiences for independent evaluations are the Boards, WBG management and staff, clients, and development partners, as appropriate.
ii) Evaluations in the WBG system should adhere to three core principles: i) Utility (the relevance and timeliness of evaluation processes and findings); ii) Credibility as a prerequisite for utility, grounded in expertise, objectivity, transparency, and rigorous methodology; and iii) Independence as a prerequisite for credibility.
iii) The principles are designed to: (a) align the World Bank Group's evaluative efforts with global challenges and the World Bank Group's strategic focus; (b) clarify the roles and responsibilities of key actors and encourage synergy among them throughout the evaluation process; and (c) ensure that all WBG evaluations are robust, of high quality, and credible.
iv) Two primary purposes: Promoting accountability, with results evaluated for their contribution to achieving the World Bank Group's goals; and promoting learning, feedback, and knowledge sharing for the systematic use of evaluation findings
v) Monitoring and evaluation enhance and feed into each other but serve different purposes. The availability of good monitoring data is necessary for good evaluation. Evaluations can, in turn, provide lessons for improving the design and implementation of monitoring systems and processes.
vi) The WBG optimizes the value of evaluations by adopting an evaluation use perspective throughout the entire evaluation process. Resources for producing evaluations and disseminating their findings are inevitably constrained, and evaluation competes with other operational and managerial demands. In addition, the capacity to absorb evaluation findings and use them effectively in accountability and learning processes varies across different evaluation audiences.
vii) Building an evaluation culture is not only about strengthening the supply side (that is, the quality and focus of evaluations) but also about strengthening the demand side—for example, by building a common understanding of the role of evaluative evidence in learning and accountability.
viii) When planning an evaluation, it is important to consider its evaluability. An evaluability assessment can be conducted as part of an evaluation process or as an exercise that is entirely separate from an evaluation. The World Bank Group 12 Evaluation Principles are Strategic selectivity. Responsive planning. Adequate resources. Collaborative approach. Rigor. Quality assurance. Customized reporting. Broad dissemination. Adequate follow-up.

**Self-evaluations**

They are conducted by operational staff or specific units within the management structures, and are therefore not fully independent of WBG management. They are usually closely linked to decision making and organizational learning processes within each institution. Self-evaluations are also conducted for purposes of accountability to WBG management and/or development partners/investors. The target audiences for self-evaluations are primarily operational units, management, clients, and development partners. In addition, to different degrees self-evaluations also inform other actors such as (representatives of beneficiaries and the general public). There are two broad types of self-evaluation in the WBG:

i) **Mandatory self-evaluation**.[44] At the core of the evaluation system across the WBG are mandatory self-evaluations of specific lending operations, investments, guarantees, country programs, and advisory services. These evaluations are prepared by the responsible operational units and are embedded in the project and program cycles. They are neither structurally nor functionally independent, but the principle of behavioral independence applies - further strengthened by IEG's review and validation (sometimes on a sample basis). Mandatory self-evaluations complement the implementation and monitoring arrangements that are embedded in each institution's project and portfolio management processes. Self-evaluation adheres to methods and guidance that are jointly accepted by WBG management and IEG, with pre-determined concepts, formats, and scope that are closely linked to the premises applied at the time of intervention approval and during reporting cycles. Aggregate analyses of (validated) self-evaluation reports enable cross-sectoral and cross-regional comparisons of performance as well as reporting at the corporate levels and to the Boards.

**ii) Demand-driven self-evaluation**. A variety of evaluation activities are undertaken in response to specific donor, client, or internal demands, or as an element of operational or research work—for example, retrospective studies of various products and instruments, trust fund evaluations, and impact evaluations. Demand-driven self-evaluations are structurally embedded in managerial processes. However, they are often either conducted, managed, or commissioned from external consultants by functionally independent units within the institution. In addition, the principle of behavioral independence applies.

**Previous systematic self-evaluations.** Some years ago the Bank also had a practice of regular retrospectives of sector strategies and reviews of some policies – such documents were presented to the Board which also received IEG comments. And a unique self-evaluation exercise was initiated in 2006 in response to a demand from the IDA Deputies for a review by IEG of IDA's internal controls. This was addressed through a multi-year and multi-stage self-evaluation by Bank management of its controls, advised by and reviewed by IEG which presented the concluding reports to the Board.

**An IEG evaluation report (Behind the Mirror**)[45] of 2015 found that the compliance with requirements was mostly strong, and that the self-evaluation systems meshed well with the independent evaluation systems for which they provide information. Having all operational units write substantive end-of-project reports was a noteworthy accomplishment that not many other organizations afforded themselves. However, the evaluation also found that the self-evaluation systems primarily focused on results reporting and accountability needs and did not provide the information necessary to help the WBG develop learning to enhance performance. Information generated through the systems was not regularly mined for knowledge and learning except by IEG, and its use for project and portfolio performance management could be improved. ISR ratings and indicators derived from them were not always precise because of weak project monitoring and optimistic reporting. Thus in the view of that evaluation the ISR would be more effective for early warning if team leaders had incentives to proactively acknowledge issues and raise risk flags. Many mid-term reviews also occurred late, as did remedial action to address identified problems because Bank and client procedures complicated and delayed restructuring of Bank projects.

The above IEG comments may be well taken, but may perhaps also have been setting the ambitions too high for self-evaluation or any systems. As one example, the ICRs serve an important accountability function that – as Bank management pointed out - cannot easily be replaced by other tools. (The ICRs are also an important record-keeping tool – the only document that in one place puts together the key aspects

[44] World Bank mandatory self-evaluations are sent to its Board. IFC and MIGA mandatory self-evaluations are reported to their Boards on an aggregate level.

[45] Behind the mirror: a report on the Self-evaluation systems of the WBG (2015).

of a completed project.) Management pointed out that the data from self-evaluations feed into corporate results measurements and serve other purposes as well. So management would explore ways to maximize learning while maintaining the accountability function.

**4. <u>Harmonization and self-evaluation instruments</u>:**

**Self-evaluation and Evaluation instruments** have expanded over the past forty years, cover many operational activities, and now include:

i) Implementation Status and Results Report (ISR) – regular reports for Bank lending during implementation; and Implementation Completion and Results Report (ICRR) for Bank lending at closing. Validated by IEG in ICR Review (ICRR Review).
ii) Currently, World Bank carries out self-evaluations for all IBRD/IDA operations regardless of funding size and all recipient executed trust funds above $5 million (with a few exceptions). The evaluations assess the project against the original project objectives and any subsequent formal revisions and rate outcomes based on criteria for relevance, effectiveness, and efficiency.
iii) For ICRs and CLRs there is full harmonization of criteria and ratings with the respective IEG validation instruments (ICRR Reviews and CLRRs), and both self-evaluations and their validations are prepared in standard formats.
iv) Ratings are based on achievement of objectives normally through indicators. This can at times lead to an emphasis on "bean-counting" with the normally six-point rating scales concentrated at the middle two ratings,
v) Any good evaluation system ought to be able to address both "doing the right things" and "doing things right". In our view, both the WB instruments tend to focus on the latter, and to give modest to little attention to whether the Bank has been doing the right things. This may be particularly noticeable for the CLRs and the CLRRs due to their emphasis on assessing against the objectives and indicators given in the PLRs.
vi) CPFs, ICRRs, ICRR Reviews, PLRs, CLRs and CLRRs are all made publicly available, as are the redacted versions of the ISRs.
vii) An earlier estimate was that ICRRs cost on average $40,000-$50,000 each to produce.
viii) The Bank has put in place a reporting infrastructure for advisory and knowledge services but has yet to develop a reliable way to evaluate the effectiveness of this work.
ix) The Bank had earlier for a number of years the Quality Assurance Group (QAG) that provided a mechanism for monitoring the quality at entry and quality of supervision of Bank lending projects. After its abolition in 2010 the Bank no longer has a mechanism for such monitoring in real time.

**Impact Evaluations**. The Bank started impact evaluations some 15 years ago, including through the Development Impact Evaluation Initiative (DIME). These are supplements to other monitoring and self-evaluations and normally measure outcomes at discrete points in time. IEG does not validate impact evaluations. This instrument is less suited for corporate accountability purposes, cannot be used for aggregated reporting, and is more akin to research.

**<u>Ratings and evaluation principles</u>**

**ICRRs.** Implementation Completion and Results Reports (ICRRs) are - according to the guidelines to World Bank staff - intended to:

i) Provide a complete and systematic account of the performance and results of each project.
ii) Capture and disseminate experience from the design and implementation of a project in order to: (i) improve the selection of interventions to achieve the goals of the country partnership framework (CPF) (or, previously, the country assistance strategy [CAS]); (ii) improve the design and

implementation of interventions through lessons learned; and (iii) help ensure greater development impact and sustainability of projects.

iii) Provide accountability and transparency at the level of individual projects with respect to the activities of the Bank, the borrower, and involved stakeholders.

iv) Provide a vehicle for realistic self-evaluation of performance by the Bank and borrowers.

v) Contribute to databases for aggregation, analysis, and reporting, especially by the Independent Evaluation Group (IEG), on the effectiveness of lending projects in contributing to development strategies at the sector, country, and global levels.

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

**The ICRR Review**, conducted by IEG, is an independent, desk-based, critical review of the evidence, results, and ratings of the ICRR in relation to the project's design documents. IEG arrives at its own ratings for the project, based on the same evaluation criteria used by the Bank.[46] IEG reviews each ICRR that is submitted to IEG, as a validation and not as an independent evaluation of the project.[47] The World Bank and IEG share a common, objectives-based project evaluation methodology for World Bank projects to assess achievements against each project's stated objectives, as well as the relevance of the objectives and the efficiency of resource use in achieving the objectives.

**There are three main project ratings that IEG validates through the ICR Review**, and one rating that is assigned by IEG only.[48] The four project ratings are –:

i) **Outcome**: the extent to which the project's major relevant objectives were achieved, or are expected to be achieved, efficiently. Both the Bank and IEG rate Outcome. (more on this rating below).

ii) **Bank Performance.** The Bank's performance is defined as the extent to which services provided by the World Bank ensured quality at entry of the project and supported effective implementation through appropriate supervision (including ensuring adequate transition arrangements for regular operation of supported activities after loan/credit closing), toward the achievement of development outcomes. Bank performance is rated using the six-point scale by assessing two dimensions:
   - **Quality at entry**, which is shorthand for "Bank performance in ensuring quality at entry" refers to the extent to which the Bank identified, facilitated preparation of, and appraised the project so that it was most likely to achieve development outcomes.
   - **Quality of supervision** refers to the extent to which the Bank proactively identified and resolved threats to the achievement of relevant development outcomes and the Bank's fiduciary role.

---

[46] When insufficient information is provided by the Bank for IEG to arrive at a clear rating, IEG downgrades the relevant ratings as warranted.

[47] For a subset of operations—perhaps on the order of 20 percent—IEG conducts Project Performance Assessments (PPARs) in the field.

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

iii) **The M&E quality rating** is based on an assessment of three main elements: (i) M&E design, (ii) M&E implementation, and (iii) M&E utilization. Monitoring and evaluation are distinct, and the rating is informed by both the quality of monitoring and the quality of evaluation. The quality of M&E is rated on a 4-point scale—High, Substantial. Modest and Negligible.

iv) **Quality of ICRR**. Because the ICRR Review is almost entirely based on the information found in the ICRR, the reliability of IEG's ratings based on the desk review depends critically on the accuracy and quality of the evidence it provides. For this reason, IEG rates the quality of the ICRR, on the four-point scale: High, Substantial, Modest and Negligible. This indicator is only rated by IEG.

IEG also writes Project Performance Assessment Reports (PPARs) on a purposefully selected share of projects.

**The Outcome rating** is built up from three secondary ratings: Relevance, efficacy and effectiveness. These three are rated on a four-point scale: High, Substantial, Modest and Negligible:

i) **Relevance of objectives** is the extent to which an operation's objectives are consistent with *current Bank country strategies* (expressed in CPFs). For the ICR, "current" refers to the time of project closing.
ii) **Efficacy** is defined as the extent to which the project's objectives were achieved, or are expected to be achieved, taking into account their relative importance, and are attributable to the activities or actions supported by the project. For the purposes of this section, the objectives refer to each of the key <u>outcomes</u> indicated in the statement of PDOs from the legal agreement (credit/lending/grant agreement) in the case of investment projects section. The achievement of each objective is assessed based on the level of achievement and the concept of "plausible causality."
iii) **Efficiency** is a measure of how economically resources and inputs are converted to results. For a development project, the central question is whether the costs involved in achieving project objectives were reasonable in comparison with both the benefits and with recognized norms (value for money).

**The project outcome is defined as "the extent to which the project's major relevant objectives were achieved, or are expected to be achieved, efficiently."** Thus, the outcome rating is derived from the prior assessment of the relevance of objectives, efficacy in achieving each objective, and efficiency. To assure consistency across IEG ICR Reviewers, IEG has developed guidelines for deriving the project outcome rating from the sub-ratings on relevance, efficacy, and efficiency in the previous sections. As the Bank is an objectives-based institution, achievements against the project development objectives (PDOs) are paramount. Thus, the benchmark for evaluation is the project's own stated objectives – not any absolute standard or someone else's conception of what good performance is.

## Table 16: Rating of Outcome.

This uses a six-point scale:

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

| | |
|---|---|
| *Moderately unsatisfactory* | There were significant shortcomings in the project's achievement of its objectives, in its efficiency, or in its relevance. |
| *Unsatisfactory* | There were major shortcomings in the project's achievement of its objectives, in its efficiency, or in its relevance. |
| *Highly unsatisfactory* | There were severe shortcomings in the project's achievement of its objectives, in its efficiency, or in its relevance. |

The Outcome rating contained in Table 1, should, in every case, be cross-checked against 'best judgment,' stepping back and asking the question "Were the shortcomings in Outcome non-existent, minor, moderate, significant, severe, or major?

**Table 17. Deriving the Overall Outcome Rating for a Project, Tree View**

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

| | | | |
|---|---|---|---|
| | | Negligible | **Highly Unsatisfactory** |
| | Negligible | High or Substantial or Modest | **Unsatisfactory** |
| | | Negligible | **Highly Unsatisfactory** |
| Negligible | High | High or Substantial or Modest | **Unsatisfactory** |
| | | Negligible | **Highly Unsatisfactory** |
| | Substantial | High or Substantial or Modest | **Unsatisfactory** |
| | | Negligible | **Highly Unsatisfactory** |
| | Modest | High or Substantial or Modest | **Unsatisfactory** |
| | | Negligible | **Highly Unsatisfactory** |
| | Negligible | High or Substantial or Modest or Negligible | **Highly Unsatisfactory** |

*Source: IEG and OPCS*

## WBG – Reviews of Country Programs

### Country Partnership Strategy Completion and Learning Reviews (CLRs) and in the IEG validations (CLRRs)[49]

1. **General:** All WBG country programs are self-evaluated (by relevant country unit) and validated (by IEG) before the next Country Partnership Framework (CPF). Normally, the CPF addresses all WBG entities – IBRD/IDA, IFC and MIGA. Ratings are fully harmonized. Most CPF periods are now around five years. Normally, there is a review around the midterm – the Performance and Learning Review (PLR) – this gives an opportunity for the country unit to modify objectives and results framework, and the outcome ratings in the CLR and CLRR assess progress against these mid-term reformulations. The CLR is included as an attachment to the subsequent CPF, while the IEG validation (CLRR) is distributed separately to the Board prior to the CPF meeting; so both are available when the Board discusses the next country program. The CLRs follow a standard format with about 15 pages of text

2. **The CLR is a tool for both accountability and learning**. It draws lessons on what works and what does not work to inform the next CPF and future strategies. The assessment is also important for determining if the WBG's program was effective and achieved its objectives. As a part of each CLR, the WBG team assesses and rates the overall effectiveness of the WBG's program in achieving its stated objectives and the WBG's own performance. IEG reviews the final CLR and provides an independent validation of the country team's self-assessment. The shared methodology is used by both the WBG team in its self-assessments and the IEG in its validation.

**3. The CLR assesses the CPF along three dimensions:**

i) **Development Outcome** – the extent to which the CPF was successful in achieving its stated objectives.

ii) **WBG performance** – how well the WBG designed and implemented the program.

iii) **Alignment with the WBG Corporate Strategy** – how well the CPF program as implemented was focused on assisting the country to reduce poverty and boost shared prosperity in a sustainable manner (the WBG Corporate Goals). This dimension is not rated.

4. **The CPF Development Outcome rating evaluates how successful the CPF program was in helping the country achieve the CPF Objectives identified in the PLR results framework**. Only achievements made during the formal CPF period are considered, no matter when the activities that supported the achievements were initiated. The Development Outcome rating considers only achievement of CPF Objectives and not completion of WBG activities or outputs. If during the CPF period, the WBG made significant progress in implementing interventions that were not expected to achieve their objectives until the next CPF period, such progress is captured under the WBG Performance rating, not under the Development Outcome rating.

[49] Much of this discussion of ratings is taken from an IEG note: "Assessing Country Partnership Frameworks (CPF): A Shared WBG and IEG Approach".

5. **The WBG Performance assessment rating is based upon two factors**: 1) how well the CPF was designed, and 2) how well the WBG implemented the CPF program. It is possible for the WBG performance rating to diverge from the rating of the CPF Development Outcome, which can be affected by outside forces including poor implementation by the government of its own program.

6. **The IEG review is an independent validation of the WBG team's self-assessment and not an evaluation of the CPF program.** Based on the evidence presented in the CLR, IEG will provide an independent judgment on 1) whether there is sufficient evidence to support the self-assessment and 2) whether the self-ratings for CPF Development Outcomes and WBG Performance are consistent with the methodology. IEG provides a rating for Development Outcome and WBG Performance. IEG bases its ratings and discussions on the relevant data and analysis presented in the CLR and additional IEG evidence if the evidence presented in the CLR is insufficient to demonstrate that the CPF Objectives have been achieved. A lack of evidence to clearly demonstrate achievement of objectives and relevance of objectives to CDG is an important factor for downgrading in the IEG validation.

7. **Rating Methodology - CPF Development Outcome:** Each CPF Objective is rated according to the five-point scale found in the table below. In determining the achievement of each CPF Objective, the WBG self-evaluation and IEG validation examine the results chain from the WBG interventions through the CPF Objective. In addition to assessing the extent to which the targets for Objective Indicators have been met, the WBG self-evaluation and IEG validation consider how well these indicators measure the achievement of the CPF Objective. If there is insufficient evidence that an Objective was achieved or not, the Objective is reported as Not Verified.

**CPF Objective Rating Scale**

| | |
|---|---|
| **Achieved:** | The program fully achieved the Objective during the CPF period (e.g. all the quantitative targets were met). |
| **Mostly Achieved** | The program made good progress towards achieving the Objective during the CPF period (e.g. more than half of the quantitative targets were met and the program is on track to meet the remainder of the targets). |
| **Partially Achieved** | The program made only limited progress toward achieving the Objective during the CPF period (e.g. less than half of the quantitative targets were met). |
| **Not Achieved** | The program made little progress toward achieving the Objective during the CPF period (e.g. few if any of the quantitative targets were met). |
| **Not Verified** | There is insufficient evidence to assess the achievement of the Objective. |

8. **If CPF Objectives remain in the results framework but the WBG took no actions to help achieve them**, they will be rated as Not Achieved even if lack of action was justified for reasons beyond the WBG team's control.

9. **The individual Objective ratings are aggregated to arrive at a rating for the CPF Development Outcome**. The overall rating is based on the six-point scale and criteria in the tables below. When determining the overall rating, Not Verified Objectives are treated as Not Achieved.

- **Highly Satisfactory**: All objective ratings Achieved

- **Satisfactory**: Majority objective ratings Achieved
- **Moderately Satisfactory**: Majority objective ratings Achieved and Mostly Achieved
- **Moderately Unsatisfactory** Majority objective ratings Mostly and Partially Achieved
- **Unsatisfactory**: Majority objective ratings Partially Achieved and Not Achieved
- **Highly Unsatisfactory**: Majority objective ratings Not Achieved

10. If certain CPF Objectives are more important than others, the CPF or CLR identifies them as such and makes a strong case for the differential weighting. Otherwise, as a starting point, all objectives are considered equal for determining the Development Outcome rating.

**CPF Outcome Rating Scale**

| | |
|---|---|
| **Highly satisfactory** | The program Achieved all Objectives; or it Achieved the majority of the Objectives and had exceptional development outcome in one or more areas. No major shortcomings (e.g. safeguard violations or significant unintended negative consequences) were identified. |
| **Satisfactory** | The program Achieved the majority of the Objectives; or the program either Achieved or Mostly Achieved the majority of the Objectives and had exceptional development outcome/impact in one or more areas. No major shortcomings were identified. |
| **Moderately satisfactory** | The program either Achieved or Mostly Achieved the majority of the Objectives; or the program at least Partially Achieved the majority of the Objectives and had exceptional development outcome/impact in one or more areas. No major shortcomings were identified. |
| **Moderately unsatisfactory** | The program Achieved some of its Objectives and the majority was at least Partially Achieved; or it either Achieved or Mostly Achieved the majority of its Objectives but produced major shortcomings. |
| **Unsatisfactory** | The majority of the program's Objectives were either Not Achieved or only Partially Achieved; or the majority of the objectives were at least Partially Achieved but the program produced major shortcomings. |
| **Highly unsatisfactory** | All of the program's Objectives were Not Achieved; or the majority of the Objectives were Not Achieved and the program produced major shortcomings. |

11. **The rating of the WBG's performance is an overall judgment** on how well the WBG has performed along two key dimensions: (i) the design of the CPF; and (ii) implementation of the CPF program. The WBG self-evaluation and IEG validation should consider, among other things, the 11 key factors listed below and discuss WBG performance based on those that are relevant.

**Factors to Consider in Assessing WBG Performance Design:**[50]

i) Design of WBG interventions for achieving CPF Objectives, including selection of Focus Areas and instruments, adequacy and appropriateness of interventions, consistency between financing and AAA, IFC additionality, synergy across WBG, and consideration of other development partners' programs.

[50] These factors are considered as appropriate, without any individual ratings.

ii) Tradeoff between risk and development impact, particularly in a fragile state.
iii) Strength of results framework and intervention logic, including realism of the CPF Objectives and the relevance of objective indicators and CPF Objectives to the support of the CDGs.
iv) Identification of critical risks and mitigation measures.
v) Integration of lessons learned from the previous CPF or CEN.

**Factors to Consider in Assessing WBG Implementation:**
i) Appropriate collaboration and appropriate division of labor between the Bank, IFC and MIGA.
ii) Quality of supervision, including managing program risk, the risk and quality of the portfolio, timeliness of program implementation and adjustments.
iii) Relevance, quality and dissemination of knowledge services.
iv) Responsiveness to changing circumstances, priorities and demands of the country, including introducing mid-course correction when needed and updating the results matrix in the PLR.
v) WBG's efforts for improving alignment with country systems and coordination with other development partners.
vi) Attention to safeguard and fiduciary issues.

**12. Performance rating is based upon the four-point scale below.**

| | |
|---|---|
| **Superior** | The design and implementation of the program successfully contributed to the pursuit of CPF Objectives with a strong results framework, timely adaptation to changing circumstance and priorities, exceptionally successful interventions and or innovations. A sound program of ongoing activities in place for the next strategy period. |
| **Good** | The design and implementation of the program successfully contributed to the pursuit of the key CPF Objectives and timely adaptation to changing circumstance and priorities. A sound program of ongoing activities in place for the next strategy period. |
| **Fair** | While successful in contributing to achievements in some areas, the design and implementation of the program failed to contribute to the achievement of a significant number of CPF objectives. The WBG did not proactively engage to address implementation problems and adapt to changing circumstances. |
| **Poor** | The design and implementation of the program failed to adequately contribute to the pursuit of the CPF Objectives. The program of ongoing activities may be in need of improvements. |

**13. Aspects Addressed in IEG's CLRRs**

Under the harmonized approach, the CLRR uses the same ratings as for the CLR. Broadly, in addition to the rated aspects: Development Outcome (based on the ratings of the individual objectives) and WBG Performance, the CLRR also discusses the following aspects:

- Relevance of the WBG Strategy:
- Congruence with Country Context and Country Program
- Relevance of Design
- Selectivity
- Alignment
- WB: Lending and Investments
- WB: Portfolio Performance
- IFC: Portfolio and New Investment

- IFC: XPSRs and DOTS ratings
- MIGA: Outstanding and new Guarantees
- WB: Analytic and Advisory Activities and Services
- IFC: Advisory Activities
- Results Framework
- Partnerships and Development Partner Coordination
- Safeguards and Fiduciary Issues
- Ownership and Flexibility (but not assessing government performance)
- WBG Internal Cooperation
- Risk Identification and Mitigation
- Overall Assessment of Bank Performance and Rating under two headings: Design and Implementation
- Assessment of CLR Completion Report
- Findings and Lessons (not recommendations)

## IFC – self-evaluation instruments - The XPSR[51]

### Overview

1. IFC's investment operations are evaluated using an XPSR (Expanded Project Supervision Report). The XPSR system enables IFC to account to its Board and shareholders for achievement of its purpose and for learning. IEG synthesizes XPSR findings for presentations on IFC's portfolio performance in specific sectors or departments. In particular, the results of the XPSRs are analyzed by IEG at the end of each XPSR cycle and presented in the annual Results and Performance of the World Bank Group (RAP).

2. XPSRs are on a sample basis – cost was one issue. The size of the sample is currently at 40%, a sampling level sufficient to allow for statistical inference at 95% confidence level on 3-year rolling average basis. This is sufficient (just) to allow IEG to draw relative performance inferences about the outcome quality of the net approval population and of the main strategic groups (e.g. country risk/income group and sector). Coverage for environmental and social aspects is however at 100%. Validated by IEG.

3. The XPSR evaluation criteria cover project or program effects on stakeholders and include financial, economic, environmental and social, and private sector development dimensions, along with IFC's investment return, work quality, and additionality. The evaluation standards and guidelines reflect a combination of benchmarks, qualitative criteria, and performance standards. To protect commercially sensitive client information, IFC self-evaluations are restricted and are not shared with the Board.

4. DOTS (Development Outcome Tracking System) – a monitoring system to measure the development effectiveness of IFC's investment and advisory work has been used for aggregation purposes and are verified by IFC's auditors – a unique aspect of this system. IFC's credit market operations are not covered. DOTS has now been replaced by AIMM (Anticipated Impact Measurement and Monitoring).

5. IEG had a searchable data base of Lessons, but this was terminated. An earlier IEG report (the 2013 Biennial Report on Operations Evaluation (BROE)) found that the quality of evidence on the outcomes of IFC's advisory services was weak, but had improved over time. The DOTS system had at the time limit information on end beneficiaries of IFC investment.

6. A previous evaluation noted that IFC had sought to reform and reduce the scope of its results measurement and self-evaluation. Some stakeholders perceived a risk of erosion of the accountability function and that arbitrating between different positions proved difficult in the absence of a policy or other guiding principles. There had been only limited progress toward systems that better met learning and business needs yet would maintain a credible level of accountability and the tone at the top of the institution has not been supportive of self-evaluation.

7. The evaluation also considered that for IFC support and guidance on writing and learning lessons was missing. Lessons were recorded but rarely used and too often of low quality, many of them were too generic, not sufficiently based in evidence, failed to recommend what specifically should be done differently in the future, or failed to address critical internal organizational issues. The evaluation identified three broad causes: excessive focus on ratings, attention to volume that overshadowed attention to results, and low perceived value of the knowledge created.

[51] Most of this discussion is taken from IEG's Instructions for Preparing an XPSR. November 2016.

**<u>The XPSR process</u>**:

i) IEG is responsible for designing and maintaining the XPSR infrastructure. It also reviews the XPSRs after they are prepared to verify the findings.

ii) At the start of each program year, IEG selects a random sample of investment operations for evaluation with an XPSR. This sample is drawn from projects typically five years old (from approval), which have generated at least 18 months of operating revenues (covered by at least one set of audited financial statements). Different maturity criteria apply to funds and greenfield financial sector projects.

iii) Operational departments' directors and portfolio managers are responsible for assigning XPSR teams, and should provide guidance and quality control.

iv) IEG reviews each completed XPSR and prepares an Evaluation Note (EvNote). It does this to assess the evaluative ratings independently, and to ensure corporate-wide consistency in the application of rating guidelines. The EvNote provides the ratings as assessed by IEG (utilizing the same definitions and criteria as in the XPSR), and also a few ratings not required in the XPSRs themselves

v) Following a circulation of the draft EvNote to the XPSR team for their comments on factual inaccuracies that the team believes should be corrected, IEG revises the EvNote if needed and as appropriate and sends it to IFC management. The final EvNote including XPSR's team's comment is filed alongside the XPSR.

vi) For closed projects IEG will prepare a Project Evaluation Summary (PES) in lieu of an XPSR, applying the same evaluation methodology, rating framework and benchmarks.

8. An XPSR should complement the data in IFC's regular supervision and credit monitoring reports, and should have the following content:

i) A summary of the project's business objectives, financial structure, sponsors and financiers;

ii) An assessment of the project's strategic relevance, and extent of achievement of the operation's objectives (as defined and presented in the Board Paper);

iii) A comparison of appraisal projections for key financial and operating indicators with the actual results to date, along with an explanation for material variances;

iv) A comparison of project performance with relevant competitors and/or sector benchmarks for the specific industry/niche;

v) An analysis of the prospects for the operation going forward to assess the sustainability of results in the longer term;

vi) Rating of the project's emerging development results, IFC's investment performance, IFC's operational effectiveness and additionality, with evidence to support each performance rating; and

vii) Identification of emerging lessons from the experience to date. (There used to be available a fully searchable database of XPSR lessons, but this has now been closed and to be handled by the IFC's knowledge management platform.)

9. The main XPSR rating is for Development Outcome. This rating reflects the overall contribution of the project to the development of its host country, and thus implicitly addresses how well the project has contributed to fulfilling IFC's purpose. A project's development outcome encompasses

all effects on a country's economic and social development. The development outcome rating is a synthesis of indicator ratings of four sub-components; it is not a simple average; it is the only rating applying a six-points scale: Highly Successful, Successful, Mostly Successful, Mostly Unsuccessful, Unsuccessful, and Highly Unsuccessful.

10. The subcomponents of Development Outcomes are rated by a four-points scale: Excellent, Satisfactory, Partly Unsatisfactory, Unsatisfactory. These are:

i) **Project Business Performance,** considering the project's impact on its owners and financiers.

ii) **Economic Sustainability** (the project and/or project company's contribution to growth in the economy). Here, the incremental effect of the project on the key economic stakeholder groups is assessed on a with vs. without-project basis, or before vs. after-project basis.

iii) **Environmental and Social Effects**. IFC considers environmental and social sustainability as an important component of development outcome quality. The XPSR's assessment of Environmental and Social Effects should comprise both the project's Environmental and Social Performance and its Environmental and Social Impacts. E&S has its own self-evaluation and IEG validates it by its E&S specialist, and IEG conducts field validation covering all category A projects, and some other selected projects.

iv) **Private Sector Development** addresses to what extent the project company has developed into a corporate role model and whether the project has contributed to IFC's purpose by spreading the benefits of growth of productive private enterprise beyond the project company or financial intermediary. In making the assessment, it is important to distinguish the project's effects on wider private sector development from those extraneous factors that may have affected the project. In particular, the indicator considers whether the project met its appraisal and Board-approved objectives to achieve private sector developmen effects in the stakeholder groups.

11. **Separately, XPSR rates also the following**, with the four-points scale:

i) **IFC's Additionality,** defined as the benefit or value addition IFC brings that a client would not otherwise have. The rating of IFC's additionality considers IFC's value proposition in providing support to the project and fulfilling the operating principles. The XPSR should assign a rating for IFC's additionality based on an assessment of delivery against the additionality claims specified in the Board report.

ii) **IFC'S Investment Outcome** assesses the extent to which IFC has realized to date, and that it expects to realize over the remaining life of the investment, the loan income and/or equity returns that were expected at approval.

12. **Additional Ratings**. The following three ratings are optional for XPSRs, but IEG validates if rated, or adds them if not rated for the EvNotes:

i) **IFC's Work Quality** is an assessment of its operational performance, including in relation to E&S aspects, with respect to pre-commitment work in screening, appraising and structuring/underwriting, and its supervision following project approval by the Board and subsequent commitment. This assessment should be made independently of the ratings for development outcome and investment outcome.

ii) **Screening, Appraisal and Structuring**. IFC's operating policies and procedures, as well as its Credit Notes set the standard of what IFC considers an appropriate professional standard of project execution quality.

iii) **Supervision and Administration**.

## IFC - The Anticipated Impact Measurement and Monitoring (AIMM) system.[52]

1. In 2017, IFC piloted a new, ex-ante project impact assessment tool—the Anticipated Impact Measurement and Monitoring (AIMM) system. The AIMM system is supposed to enable IFC to measure the development impact of its investment and advisory projects.

2. The AIMM system is supposed to provide a robust operational framework that:

- i) Improves IFC's ability to select and design projects that maximize its development reach.
- ii) Sets ambitious targets and the incentives to achieve them.
- iii) Strengthens IFC's capacity to deliver an optimal mix of projects that deliver both high development impact and solid financial returns.

3. The AIMM system is supposed to assess a project's development impact along two dimensions:

- i) **Project Outcomes:** A project's direct effects on stakeholders (including employees, customers, suppliers, government, and the community); the direct, indirect, and induced effects on the economy and society overall; and the effects on the environment.
- ii) **Market Outcomes:** A project's ability to catalyze systemic changes that go beyond those effects brought about by the project itself.

4. For market-level outcomes, the AIMM system will assess the degree to which an intervention improves the structure and functioning of markets by promoting one or more of the following objectives:

- i) **Competitiveness:** Competitive markets are those where firms can effectively enter, exit, and compete. These markets also support product or process innovation, improved management practices, and/or lower product cost.
- ii) **Resilience:** This objective refers to improving the depth, structure, regulation, and governance of markets to help them withstand shocks.
- iii) **Integration:** Promoting this objective involves the enhancement of physical and/or financial connectivity, within and across markets.
- iv) **Inclusiveness:** Markets that are more inclusive are those that support fair and full access by marginalized groups to goods and services, finance, and economic opportunities.
- v) **Sustainability:** Markets are more sustainable when firms and consumers adopt environmental and social sustainability technologies and practices.

5. To deliver one or more of the market objectives, IFC's interventions will identify channels through which systemic changes occur. These channels refer to actions that put in place frameworks that enable markets to deliver market objectives, promote competition that causes other market players to up their game, provide demonstration effects, replication, and more generally the spillover of ideas while creating new productive networks, or build capacity and skills that open new market opportunities and support market objectives.

6. The AIMM system will help IFC to maintain a line of sight from its intermediate (including market) objectives to the World Bank's twin goals and the SDGs. To this end, under AIMM it will evaluate its incremental impact by comparing the direct and indirect outcomes of an IFC intervention to a scenario without one. By design, the AIMM system measures project-level and systemic outcomes ("creating markets") against objectives that are associated with the SDGs.

[52] Source: IFC information material (details to be included in final version).

7. The second "M" stands for "monitoring," an essential component of the AIMM system. Each development outcome claim in IFC projects will be explicitly tied to one or more monitoring indicators, regularly tracked during portfolio supervision. By tracking these indicators, the AIMM system is supposed to link project ratings with real-time results measurement findings. There will also be a structural link between front-end diagnostics and ex-post evaluation functions.

8. The AIMM system thus is supposed to connect diagnostics to ex-ante project selection/scoring, which is tied to results measurement during portfolio supervision, and ultimately, to ex-post evaluation. In total, the system comprises two critical pillars (project ratings and results measurements).

### IFC – Self-Evaluation of Advisory Services[53]

1. At completion of IFC Advisory Services (AS) projects, IFC produces PCRs (Project Completion Reports), which IEG validates in EvNotes. All projects are required to conduct a PCR, unless the project qualifies for one of the exclusion criteria approved by IFC management.[54]

2. The process of conducting PCRs includes:

i) Self-evaluation by project teams, which are responsible for the quality of PCRs and providing all required documentation and evidence.

ii) Independent validation by IEG. IEG conducts an in-depth desk review of a sample of projects selected by IEG based on internal data and independent research, which may also be supplemented by a field visit

iii) For PCRs selected for validation, IEG prepares an Evaluative Note (EvNote), which records IEG's independent assessment and ratings. IEG first circulates a draft EvNote to the PCR team inviting their comments on aspects related to factual accuracy. IEG then assesses the comments and revises the EvNote, if needed and as appropriate, and records it as final.

iv) PCR teams' comments will be reflected in an IFC Comments section of the EvNote.

3. The PCR includes ratings for Development Effectiveness (DE) and for IFC's Role and Contribution. IEG validates these ratings using the same definitions and methodology. In addition, IEG in the EvNotes also rates the IFC Work Quality.

4. Development effectiveness reflects the extent to which an intervention achieved its intended development results and private sector development. This is based on a synthesis (not an average) of ratings of five dimensions: (1) Strategic Relevance; (2) Output Achievement; (3) Outcome Achievement; (4) Impact Achievement;[55] and (5) Efficiency.

5. While DE is based on a synthesis of all underlying dimensions, achievement of overall project objectives is important for rating overall DE. Rating scale: Highly Successful, Successful, Mostly Successful, Mostly Unsuccessful, Unsuccessful, Highly Unsuccessful. Not Applicable can be used if they meet exclusion criteria. The five underpinning aspects are rated Excellent, Satisfactory, Partly Unsatisfactory,

[53] Mostly taken from the IFC Guidance document: Advisory Services – Project Completion Reports for Ratings.
[54] Examples of excepted projects: No advisory services provided. No development results expected. Humanitarian assistance to regions affected by natural disasters (delivery of basic supplies to regions affected by natural disasters).
[55] Impact may not necessarily be reflected since for many projects it may be too early.

Unsatisfactory, or Not Applicable if they meet exclusion criteria. For Impact Achievement, additional ratings of Too Early to Judge or Cannot be Verified can be assigned.

6. IFC Role and Contribution is the extent to which IFC made a special contribution to the client. This indicator thus asks to what extent IFC brought additionality or made a special contribution to the AS. Rating scale: Excellent, Satisfactory, Partly Unsatisfactory, Unsatisfactory.

7. IFC Work Quality is an assessment of project design, supervision and implementation. This dimension (only rated by IEG) assesses the extent to which services provided ensured quality at entry and supported effective implementation through appropriate supervision and execution, towards the achievement of DOs. This dimension does not affect AS project's DE. IEG assesses the IFC's Overall Work Quality by assessing and rating separately two dimensions: (1) Project Preparation and Design and (2) Project Implementation, and then arriving at an overall synthesis rating. Uses the same four-point scale as for IFC role and contribution (above).

## ADB - Public Sector Projects

1) **Results Framework.**

ADB has a results framework. Management sees as an advantage that ADB has strong results framework and design and monitoring matrix. IED in collaboration with ADB management has launched a knowledge platform to share findings, lessons and recommendations from past projects. Goal is to provide easy access to information that can be used to inform and improve quality of design and implementation of new projects.

2) **Evaluation Policy**.

There is an evaluation policy for IED, but ADB does not have a free-standing evaluation policy – operational procedures are contained in Operational Manual, including independent and self-evaluation. The 2018 External Review of the Independent Evaluation Department (IED) recommended therefore to develop an ADB-wide evaluation policy. To this end, in their view, management, the Board and its Development Effectiveness Committee (DEC), and IED would all need to be engaged in drafting the policy through a highly engaged process designed to generate consensus on what all of the parties involved want out of evaluation, and defining those purposes in terms of a series of outcomes that can be measured and reported on, while preserving the essential features of independence and objectivity. Appendix 1 summarizes this aspect of the External Review in more detail. IED agrees with this recommendation.

3) **Self-Evaluation and Independent Evaluation**

- Role of self-evaluation in ADB along with independent evaluation, to improve the design and execution of future operations and activities and to revise policies and business processes.
- Operational departments conduct self-evaluations and notify management and Executive Board. Management submits PCRs to the Board, and IEG then submits validation reports when ready. For CPSs, IED submits both CAPEs (country program evaluations) and final reviews to the Board.
- Outputs: Operations: Project Completion Reports (PCRs) are prepared for all public sector, non-sovereign, and technical assistance operations (called Technical Assistance Completion Reports). Country Partnership Strategy Final Review prepared for all country programs unless IED is doing a CAPE. Other self-evaluations may be undertaken on an ad hoc basis.
- Management view: For proper and systematic PCRs: Need a template, and staff instructions, with timing expectations. Important to emphasize that PCRs are for the whole projects including parts not financed by ADB. If management takes PCRs seriously, so will staff.
- Self-evaluation reports should assess operations and evaluate the adequacy of preparation, design, appraisal, and implementation arrangements as well as the performance of consultants, contractors, suppliers, borrower, client, executing agency and the ADB overall.
- Management view: Project self-assessment: Very useful learning exercises for new staff, especially if undertaken together with staff that have been responsible for implementation. When having to use some consultants, staff should stay involved and be responsible. Final PCR should reflect their views, not the views of consultants. Also valuable references for staff starting new projects – what happened to previous similar operations? So review existing self-evaluations.
- Management emphasized need for project teams to keep in mind throughout implementation the PCR information and analytical requirements needed at the end. Think about PCRs from the start.
- Management view: Need for PCRs to show all financial costs/investments, also those carried by the government or other partners. This can be complicated since governments may not have same motivation for a PCR. So have to get executing agencies involved. With staff: Important that they understand – this is not an evaluation of them or their efforts, but of the projects.

- Spends $30-50,000 all inclusive on PCRs – done carefully as part of supervision work.
- Encourage PCR missions to e.g. visit as many subprojects as possible – not just those conveniently located.
- PCRs go through several layers of management review, often several rounds. Received by the Board but don't think has ever been a Board meeting on a specific PCR. Several instances of Board members referring to PCR findings "next time around".
- PCRs: Accountability and learning, but still primarily an accountability tool. Important part of project cycle.
- IED's Annual Evaluation Review is based on IED's evaluations and validations. May limit this exercise to projects. Issues may include life cycle (project administration), learning, and organization for delivery.
- PCRs projects: Now two years after closing. Will become one year after financial closure (which is later than project completion).
- Management: IED now validates all PCRs – useful. At worst: a second opinion is always useful. IED also comments as peer reviewer on draft project (concept paper and the draft report) and policy operations documents.

4) **<u>Harmonization of Evaluation Systems, Instruments and Methods</u>:**

- For projects and country programs: Self-evaluations and validations/independent evaluations follow same methodologies.
- Operational departments undertake self-evaluations of projects and other activities, and IED validates all self-evaluations of projects and country programs and also undertakes independent evaluations for selected operations.
- From IED Self-Evaluation 2017: Nature and scope of validation work not entirely clear to IED and ADB staff. IED's role should be to use a standardized template and guidelines to validate the assessments in the PCRs, XARRs (extended annual review reports) and CPSFRs (country partnership strategy final reviews), based on evidence provided to support the self-evaluations by operational units. But IED often finds itself in the position of having to defend the quality of the evidence that underpins its validations. This is an inversion of what ought to be standard practice. Distinguishing the responsibilities of IED from those of ADB operational units should be a high priority.
- Issues for projects: Typically re ratings – where disconnects happen (variance of self assessment and IED evaluation). This usually stems from lack of PCR quality and lack of sufficient information (evidence) to support ratings. Criteria have been harmonized.
- New Initiative: TA self-evaluations. The TA PCRs to be delivered soon after completion - expect within six months.
- IED has worked with management on roll-out of TA self-evaluation. Completion reports are now just two-pagers. Have piloted – and new guidelines – now more systematic. Used workshops and provided capacity building. IED will validate samples of those TAs over $250,000. Light touch below. Guidelines have been developed but are not yet publicly available.
- ADB requires that all public sector projects that have incurred ADB expenditures and are regarded as completed are subjected to self-evaluation by regional departments. IED does not participate in the preparation of PCRs or TCRs.

- Evaluation of projects by IED is undertaken at two levels; (i) through a desk review, or validation, undertaken for approximately 80% of PCRs and reported in project completion report validations (PVRs); and (ii) through field-based evaluations, undertaken for 10%–20% of projects, and reported in project performance evaluation reports (PPERs).

- For a smaller proportion of TA projects, TA performance evaluation reports (TPERs) are prepared, usually for a group of such projects at a time. TPERs will continue as an IED product; they involve a field mission and provide more detailed evaluative assessments.

- PVRs, PPERs, and TPERs (independent evaluation reports) are made available to the public through the IED website upon their approval by the Director General.

- PVRs and PPERs differ in their depth and breadth of assessments because of the amount of time spent, the analytical resources utilized, and the type of evidence collected in their preparation. PVRs rely on a rapid assessment of project performance based mainly on desk reviews and cross-checking of the PCR, the report and recommendation of the President (RRP) and associated documents. Preparation of a PVR usually takes 2–3 weeks.

- Although the PVR process is heavily dependent on the PCR and RRP for information on project design and performance, new information submitted by the regional department and information gained independently by IED can add to the PVR and influence ratings. The PVR therefore stands as an independent review of the project.

- PPERs are in-depth assessments of projects, based on evidence from documentation and files, as well as field visits, occasional surveys, and interviews with ADB staff, government, and other stakeholders. Preparation of a PPER typically requires 3 months of work from an evaluation specialist on full-time basis, with more intermittent help from an IED officer and assistant. Consultants are often also used for PPERs. A project can be the subject of both a PVR and (subsequently) a PPER. If that happens then the PPER ratings supersede the PVR ratings in ADB and IED reporting systems.

- A new ICR validation is being developed and piloted. Until now, IED has not validated individual TCRs, although it might review a group of TA projects and their TCRs for a thematic or corporate evaluation study. Likewise, a number of relevant TA projects and TCRs may be reviewed for a country assistance program evaluation.

**5) <u>Ratings and evaluation principles</u>**

The overall success rating of a public sector operation is based on the assessments of: (i) relevance, (ii) effectiveness, (iii) efficiency, and (iv) sustainability. Each core criterion is weighed evenly when calculating the overall success rating.

- <u>The relevance assessment</u> addresses the extent to which: (i) the intended outcomes of the project were strategically aligned with the country's development priorities (considering both what is included in the project and what ought to be included) and did not duplicate the project work of other development partners. (ii) the intended outcomes were aligned with ADB's country and sector strategies, including one or more of the ADB corporate priorities with which the project was explicitly associated (themes, strategic agendas, or drivers of change); and (iii) the project design was appropriate for achieving the intended outcomes, i.e., competent analysis was carried out, lessons were applied, the right financing instrument or modality was chosen, innovation and transformative effects were given attention, and the indicators and targets at various levels were laid down well and lent themselves to measurement.

  Each rating uses a four-point scale (3 to 0): 3 (e.g., *highly relevant*) is equivalent to a better than expected result; 2 (e.g., *relevant*) is equivalent to an expected result; 1 (e.g., *less than relevant*) is a less than an expected result; and 0 (e.g., *irrelevant*) is no result or a poor result.

- The effectiveness assessment looks at whether the project's intended outcomes were achieved or were expected to be achieved at the time of observation (i.e., at completion or later), and whether any unintended outcomes had inadvertently reduced the value of the project. A project can be rated: highly effective, effective, less than effective, or ineffective.

  - The outcomes are evaluated against the baselines and targets listed in the DMF (design and monitoring framework) at the outcome level, if the DMF is fully in line with the main project document and the baselines and targets themselves are not deficient.
  - Effectiveness requires reliable data. Data on outputs and outcomes need to be derived from credible and documented sources.
  - The effectiveness assessment will consider the results of the implementation of safeguard related plans and gender action plans to the extent that specific interventions were identified, especially if they were negative. This is the case even if safeguard or gender outcomes were not identified in the project DMF. Projects can have unintended adverse effects on people if social and environmental risks are not dealt with.
  - Assessments that influence the effectiveness rating can include whether: (i) safeguard plans and monitoring reports, if any, were adequate, disclosed in time, and executed in time; (ii) safeguard monitoring was adequate and timely; (iii) any changes in safeguard measures were well justified and were made in reaction to a changed scope; (iv) potentially negative environmental and social impacts of the project were avoided, minimized, or mitigated, and affected people were properly compensated and not made worse off; (v) there are remaining or unresolved safeguard-related issues and complaints; and (vi) opportunities materialized for affected or displaced people to benefit from the project.

- The efficiency of a project is a measure of how well it used resources to achieve its outcome(s). It indicates whether the project used resources efficiently for the country and/or society (not merely for the operating entity) on a whole-of-life basis. An evaluation is expected to include a recalculation at completion of the EIRR (economic internal rate of return) used at the appraisal stage, based on updated actual costs and benefits, if such a calculation was undertaken in the RRP. A project can be rated highly efficient, efficient, less than efficient, or inefficient. In the case of most ADB-supported projects where the benefits are measurable and quantifiable, the EIRR is the standard efficiency measure for a project.

- The sustainability assessment focuses on the likelihood that project outcomes and outputs will be maintained over the economic life of the project (for investment operations) or over a meaningful timeframe, demonstrating the persistence of results from the policy supported and institutional actions taken (for policy-based operations). A project can be rated most likely sustainable, likely sustainable, less than likely sustainable, or unlikely sustainable.

- The development impacts assessment is focused on long-term, far-reaching changes to which the project has plausibly contributed. A project's development impacts can be rated highly satisfactory, satisfactory, less than satisfactory or unsatisfactory. If long-impacts are uncertain, the rating can refer to likely impacts.

- Separate ratings are provided for the performances of ADB and cofinanciers. Ratings should be assigned as follows: (i) Highly satisfactory (ii) Satisfactory iii) Less than satisfactory. (iv) Unsatisfactory.

- Borrower and Executing Agency Performance. This assessment focuses on the adequacy of ownership and assumption of responsibility by all participating government entities during the project cycle. Ratings of the performance of the government (the borrower) and the executing and implementing agencies should be assigned as follows: (i) Highly satisfactory. ii) Satisfactory. (iii) Less than satisfactory. There was a major shortfall in at least one key performance area. (iv) Unsatisfactory.

- Overall Assessment: The steps and considerations for the overall rating are as follows: (i) Weights for core criteria. Each of the core criteria has an equal weight of 25%. Fixed whole number scores are used to assign appropriate overall ratings (highly successful, successful, less than successful, or unsuccessful. The overall rating is determined by aggregating the ratings for the four core criteria of relevance, effectiveness, efficiency, and sustainability. Under each core criterion, the four descriptor ratings translate into a whole number score between 0 and 3. The overall project assessment rating is a weighted average of the values of the four core criteria ratings and therefore ranges between 0 and 3. The ratings should be assigned as follows: (i) Highly successful. (ii) Successful. (iii) Less than successful. (iv) Unsuccessful.

6) **Some Points and Issues pertaining to Self-Evaluation**:

- ADB used to do institution-wide Quality at Entry assessments every two years, in the beginning with IED involvement. Now this is done at the level of operational departments.
- ADB has strengthened its self-evaluation system as part of efforts to improve development effectiveness and has put greater emphasis on knowledge and learning from operations.
- However, a few years ago it was considered that results frameworks and monitoring remained weak. Content of completion reports somewhat superficial. Learning from self-evaluation reports was uneven. PCRs served more for accountability, although some regions also focused on learning. Better record in learning from country strategy implementations because it was mandatory to take account of the lessons from the previous country strategy when preparing a new one.
- IED provides training for PCRs and recognizes and awards good quality PCRs. IED also provides comments for new projects at concept stage, mostly limited to adoption of lessons and monitoring framework.
- IED provides comments on appraisal reports, typically re use of lessons and logframe. Capacity development function. Governments to provide a completion report.
- Annual reports: AER (IED): DER (management). Management also Annual Portfolio Performance report. Overlap: IED ratings. AER has three parts: Performance – IED ratings, a thematic chapter, and implementation of evaluation recommendations.

## ADB Country Assistance Program Evaluations and Country Partnership Strategy Final Review Evaluations

- IED has been preparing country assistance program evaluations (CAPEs) since 1998. In 2006, it began validating ADB country partnership strategy final review (CPSFR) reports, which are prepared by operations departments.
- CPSFR validations mostly cover only one CPS period. They will generally also include an assessment of implementation and results of projects approved before the validation period but that was ongoing or completed over that period.
- The CPSFR validation checks the final review of the operations department, and may expand on it, as needed. A CAPE does not require such a final review. The processes for preparing a CAPE and a CPSFR validation are the same, unless specified otherwise.
- A CAPE, or a CPSFR validation, is prepared before each new CPS for the country so it can provide an input to the new CPS by assessing the program, identifying lessons, and making recommendations.
- In cases where IED decides not to prepare a CAPE for the country, the operations department prepares a CPSFR, which is then validated by IED. CAPEs and CPSFR validations should be completed before the informal board seminar on the new CPS. Plans for new CPSs need to be announced in time for the evaluations to be carried out properly.
- Typically, the total preparation time required by IED will be 12 months for a CAPE and a minimum of 5 months for a CPSFR validation.
- Country- level evaluations: Validation of Country Partnership Strategy Final Review (CPSFR) takes one mission. Full IED evaluation of country programs (CAPE) includes sector work and may take several missions. Validation of the CPSFR is for a concluded country program (with one programming cycle), except where a CAPE is pipelined; as opposed to a CPSFR, the CAPE evaluates more than one programming cycle. The final reviews (FRs) are done towards the end of a cycle, so that IED's validation of the FR will be available before the new CPS, as per guidelines.
- IED's country-level evaluations normally consider all types of ADB operations in a country and how these activities are coordinated to contribute to maximum development impact. The evaluations should review and analyze not only ADB's public sector operations and technical assistance (TA) and policy dialogue, but also ADB's non-sovereign operations (NSO), including NSO with public sector entities in the country concerned.
- The purpose of the CAPE and CPSFR validation is to provide ADB's Board of Directors and Management with an independent assessment of past operational performance in a country; and to articulate issues, lessons, and recommendations.
- Standard evaluation questions should be similar in all country-level evaluations: (i) Relevance, (ii) Effectiveness, (iii) Efficiency, (iv) Sustainability, (v) Development impacts, (vi) ADB performance, (vii) Borrower performance, (viii) Issues, lessons, and recommendations.
- The evaluation should pay close attention to analyzing whether the CPS objectives are strategically relevant and whether and how ADB has contributed to their achievement.
- Given the complexities of country-level evaluation, the limitations of the CAPE and CPSFR validation methodologies should be acknowledged clearly in the reports (e.g., they should explain factors impinging on the robustness of the assessment, the breadth and depth of the evidence upon which the assessment is drawn, and the inability to precisely measure and/or attribute impact at sector and country levels when recent data are absent). For CPSFRs and their validations in particular, assessments of the effectiveness, sustainability and development impacts of new operations may be premature. They may need to concentrate on operations approved before the current CPS period but ongoing and/or completed over that period.
- Five evaluation criteria are applied: (i) relevance, (ii) effectiveness, (iii) efficiency, (iv) sustainability, and (v) development impacts. A weight of 20% is assigned to each criterion, although, in exceptional cases

the weights assigned to each criterion can be adjusted where the country conditions and the nature and modality of ADB's interventions are significantly different from normal.

- A CAPE or CPSFR validation rates the performance of ADB's strategy and program for the country based on the actual value of the sum total of ADB financial interventions ongoing and completed during the evaluation period (the CPS program comprises sovereign and non-sovereign loan and grant operations, and TA operations). It first assesses the relevance of the CPS objectives based on the various CPSs under review. Then it assesses the performance of the selected programs using the effectiveness, efficiency, sustainability and development impacts evaluation criteria. A weighted average of the ratings of various sector-level assessments of the five criteria is provided to arrive at overall ratings for the CPS program. After weighing the rating scores for the five criteria equally, the overall success rating is obtained.
- Ratings: Relevance (20%): The rating categories are: *highly relevant*, *relevant*, *less than relevant*, and *irrelevant*. Effectiveness (20%). The rating categories are *highly effective*, *effective*, *less than effective*, and *ineffective.* Efficiency (20%). The rating categories are *highly efficient*, *efficient*, *less than efficient*, and *inefficient.* Sustainability (20%). The rating categories are *most likely sustainable*, *likely sustainable*, *less than likely sustainable*, and *unlikely sustainable*. Development impacts (20%). The rating categories are *highly satisfactory*, *satisfactory, less than satisfactory*, and *unsatisfactory*.
- IED prefers quantitative assessments, relying on survey data and economic and financial analysis. In the absence of these, the evaluation team needs to apply more qualitative approaches to arrive at a performance rating. The two approaches need to be combined. A numerical scoring system is used to aggregate the various ratings of the success of the CPS, for communication, and recording across countries, sectors, and time periods. The CAPE or CPSFR validation ratings are used by ADB Management for corporate reporting.
- IED's scoring system at the country program level is consistent with its scoring at the project level in terms of scoring and scaling: (i) a score for a criterion goes from 0 points (lowest) to 3 points (highest); (ii) the score is multiplied by the weighting percentage, e.g., relevance rated 1 (less than relevant) is multiplied by 20% = 0.20; (iii) the success rating score is derived by adding up the various individual criterion rating scores; (iv) a final rating is assigned by comparing the score with the following IED success rating thresholds: *highly successful* (HS) = 3.00–2.50, *successful* (S) =2.49–1.60, *less than successful* (LS) = 1.59–0.80, and *unsuccessful* (US) = 0.79–0. The thresholds for identifying whether a strategy is HS, S, LS, or US are as follows: 2.5 =< HS <= 3.0; 1.6 =< S < 2.5; 0.8 =< LS < 1.6; 0.0 =< US < 0.8.
- Disclosure. All CAPE and CPSFR validation reports are put in the public domain by way of the ADB and IED websites.

## ADB Non-sovereign Operations

- Abbreviations: XARR – extended annual review report (self-evaluation). XVR – extended annual review validation report.
- A nonsovereign operation (NSOs) is defined as an ADB-financed transaction in the form of a guarantee, loan, or equity investment with a subsovereign, state-owned enterprise, other public-private entity, or private sector entity as obligor or investee, normally without direct sovereign indemnity.
- Self-evaluations are presented in extended annual review reports (XARRs). Independent evaluations of NSOs presented in project performance evaluation reports (PPERs) prepared by IED.
- PPERs are not prepared by IED for all projects, and IED validates all XARRs.
- The guidelines apply to all PPERs for NSOs prepared by IED, as well as all extended annual review validation reports (XVRs).
- Performance evaluation uses the adopted ECG standards involves assessing and rating NSO projects on the basis of (i) development results, (ii) ADB additionality, (iii) ADB investment In addition, PPERs include an assessment of (i) identified lessons and their dissemination, and (ii) follow-up actions and recommendations.
- The evaluating team rates the project's performance for the first four dimensions, and assigns an overall performance rating. The evaluation dimensions and rating standards are identical for PPERs, XARRs, and XVRs.
- Selection of NSOs ready for evaluation is defined as having reached early operating maturity (EOM). Evaluation will be selected on the basis of purposeful sampling to serve specific evaluative needs such as the potential for learning, input into other high-level evaluation studies, the high profile of a transaction, credit and other risks, the likelihood of replication, or the desirability of balanced country and sector coverage.
- The final PPER will be bar coded and circulated to ADB's Board of Directors (the Board) and senior management team, with a cover memorandum noting confidentiality and restricted distribution requirements.
- IED validates all NSO XARRs where no PPER is being prepared for that project. The XVR, which is a non-Board information paper, is prepared after an independent review by IED
- using the evaluation and rating methodology of these guidelines.
- NSO contributions to high-level development impacts are assessed along a probable results chain considering the actual achievement of relevant project outputs and outcomes. For private sector NSO, this assessment reflects how well the investment contributes towards PSD and other targeted development impacts, and ultimately, towards inclusive growth, environmentally sustainable growth or regional cooperation. For public sector NSO, an assessment of investment contributions to private sector development is optional, if there was no related project objective stated at the time of approval.
- The overall rating of the NSO project is derived from the underlying ratings applied to the four dimensions of an NSO evaluation: (i) development results, (ii) ADB investment profitability, (iii) ADB work quality, and (iv) ADB additionality. The overall ratings matrix assigns ratings ranging from *unsatisfactory* to *excellent* for each of the four NSO dimensions, and derives an overall summary result. Fixed weights are not applied to each of the four dimensions to derive the overall rating, and the relative significance of each measure will depend on the project context and the importance of various project objectives.

## IFAD Public Sector Projects

1) **Results Framework**

IFAD has a Development Effectiveness Framework (DEF) that shows planned improvements and their status. It also has a Results and Impact Management System (RIMS) – updated in 2017 with a new list of core indicators to assess project performance. The core indicators have also been integrated into IFAD's Operational Results Management System (ORMS), which can now be updated and monitored continuously throughout the project life-cycle.

2) **Evaluation Policy**.

IFAD has a free-standing Executive Board-approved evaluation policy. The revised policy was approved in 2011, and some human resource-related items were amended in 2015 – the date of the policy on the IFAD website.

- The policy focuses primarily on independent evaluation, but also includes self-evaluation. The evaluation policy calls for self-evaluation reviews of all country programs and projects financed by IFAD.
- The evaluation policy specifies that self-evaluation is essential to facilitate independent evaluations by IFAD's Independent Office of Evaluation (IOE).
- The policy also calls for IOE to assess the design and functioning of the self-evaluation system and for IOE to comment on the RIDE

3) **Self-Evaluation and Independent Evaluation**

- The purpose of the self-evaluation system is to ensure performance assessment and generation of lessons.
- Management is responsible for the self-evaluation system with a designated Department taking the lead (the Programme Management Department).
- Management produces an annual report – the Report on IFAD's Development Effectiveness (RIDE).
- The RIDE draws on the self-evaluation outcome ratings, but also discusses other aspects of IFAD's development effectiveness. It is discussed by the Evaluation Committee and the Executive Board, and is publicly available.
- IOE also produces its separate annual report – Annual Report on Results and Impact (ARRI) of IFAD Operations. This is based on the findings from independent evaluations and validations, and is discussed as the RIDE by the Evaluation Committee and the Executive Board. It is also made publicly available.

4) **Harmonization of Evaluation Systems, Instruments and Methods**:

- There is full harmonization of evaluation criteria between self-evaluations and independent evaluations, reflected in harmonization agreements between IFAD and IOE – the latest (still in effect) is from 2017. It focuses on the evaluation criteria.
- The OECD/DAC evaluation criteria form the basis for the IFAD evaluation systems, with additional criteria arising from the evolution of the evaluation function and interactions with the Executive Board.
- Both self-and independent evaluations use the same six-point rating scale (next para), and will also rate the overall rural poverty impact to enable comparison and the tracking of trends over time.
- At the project level, this harmonization applies for independent project completion validations (PCRVs), project performance evaluations (PPEs), impact evaluations (IEs), and self-evaluation

project completion reports (PCRs). PCRVs, PPEs and IEs continue to calculate and show the average difference (or disconnect) between self- and independent evaluation ratings and to assess the quality of PCRs.
- As appropriate, the evaluation and validation reports will make an ex-post assessment of the validity of the assumptions at the design stage, and will examine a project's theory of change in greater detail.

5) **Rating Scale and Evaluation Principles**

i) **Rating scale and methods**

- The six-point rating scale: Highly Satisfactory – Satisfactory – Moderately satisfactory – Moderately Unsatisfactory – Unsatisfactory – Highly Unsatisfactory.
- On a binary scale the first three ratings are Satisfactory and the last three Unsatisfactory. IFAD moved from a four to a six-point scale in 2005.
- Evaluators are generally expected to exercise their own judgments, based on the available evidence. They are also encouraged to make full use of all the points on the rating scale.
- All ratings will be from 6 to 1 – corresponding to the ratings – and in round numbers with no decimals. No weights will be assigned to ratings to determine the final rating for project performance or overall project achievement.
- The project performance is an arithmetic average of the ratings on relevance, effectiveness, efficiency and sustainability (so it can have decimals). In contrast, the rating for overall project achievement is an integer number and it is not an average of all other ratings. So it represents an independent judgment, drawing on the ten other ratings (without any numerical averaging).

ii) **Evaluation principles**

- **Rural poverty impact**: Impact is defined as the changes that have occurred or are expected to occur in the lives of the rural poor (whether positive or negative, direct or indirect, intended or unintended) as a result of development interventions. There are four impact sub-domains of this rating (not to be rated individually):
  - Household income and net assets: Household income provides a means of assessing the flow of economic benefits accruing to an individual or group, whereas assets relate to a stock of accumulated items of economic value. The analysis will assess to the extent possible the trends in equality over time.
  - Human and social capital: An assessment of the changes that have occurred in the empowerment of individuals (particularly among vulnerable groups) and their capacity to organize themselves and act collectively.
  - Food security and agricultural productivity: Changes in food security relate to availability, stability, affordability and access to food and stability of access, whereas changes in agricultural productivity are measured in terms of yields; nutrition relates to the nutritional value of food and child malnutrition.
  - Institutions and policies: Changes in the quality and performance of institutions, policies and the regulatory framework that influence the lives of the poor.

- **Project performance:** This is an average of the ratings for relevance, effectiveness, efficiency and sustainability of benefits.
  - Relevance: The extent to which the objectives of a development intervention are consistent with beneficiaries' requirements, country needs, institutional priorities and policies. It also

entails an assessment of project design, coherence in achieving its objectives, and relevance of targeting strategies adopted.
- Effectiveness: The extent to which the development intervention's objectives were achieved, or are expected to be achieved, taking into account their relative importance.
- Efficiency: A measure of how economically resources/inputs (funds, expertise, time, etc.) are converted into results.
- Sustainability of benefits: The likely continuation of net benefits from a development intervention beyond the phase of external funding support. It also includes an assessment of the likelihood that actual and anticipated results will be resilient to risks beyond the project's life.

- **Other performance criteria**:
  - Gender equality and women's empowerment: The extent to which IFAD interventions have contributed to better gender equality and women's empowerment, for example, in terms of women's access to and ownership of assets, resources and services; participation in decision making; work load balance and impact on women's incomes, nutrition and livelihoods.
  - Innovation: The extent to which IFAD development interventions have introduced innovative approaches to rural poverty reduction.
  - Scaling up: The extent to which IFAD development interventions have been (or are likely to be) scaled up by government authorities, donor organizations, the private sector and other agencies.
  - Environment and natural resources management: The extent to which IFAD development interventions contribute to resilient livelihoods and ecosystems. The focus is on the use and management of the natural environment, including natural resources defined as raw materials used for socio-economic and cultural purposes, and ecosystems and biodiversity – with the goods and services they provide.
  - Adaptation to climate change: The contribution of the project to reducing the negative impacts of climate change through dedicated adaptation or risk reduction measures.

- **Overall project achievement**: Overarching assessment of the intervention, drawing on (but with no numerical averaging) the analysis and ratings of rural poverty impact, relevance, effectiveness, efficiency, sustainability of benefits, gender equality and women's empowerment, innovation, scaling up, as well as environment and natural resources management, and adaptation to climate change.

- **Partners' performance** (IFAD and the Government). This criterion assesses the contribution of partners to project design, execution, monitoring and reporting, supervision and implementation support, and evaluation. The performance of each partner will be assessed on an individual basis with a and to be rated.

- **PCR quality ratings**: Scope, quality, lessons learned, candor.

**Some points and issues pertaining to self-evaluation:**

- PCRs are formally done by governments (although often with IFAD support). In the past these reports took a few years to complete after project closing. They are now supposed to be completed within six months, and there are efforts to try and have them ready by project closing and at least two months thereafter – a significant improvement, but there are still some delays. There is no strict time limit for the PCRV, but it is generally conducted within six months from the date when it is received by IOE.

- All completed projects have PCRs, as well as PCRVs or a PPE or an Impact evaluation. IOE normally does not conduct a PPE or an impact evaluation when they already have a PCRV. However, in the past they have done so in a few cases, based on the IOE selection framework.
- Project monitoring is under the relevant governments, and often gets modest or low priority.
- IFAD management has been pushing for improvements in M&E, including with training and assistance from CLEAR. The program is called PRIME (Program in Rural M&E).
- Impact evaluations: Management did a little in the past, improvements were needed. There is an impact assessment initiative run by the Research and Impact Assessment Division within the Strategy and Knowledge Department of IFAD. It assesses about 15% of the closed projects and report systematically on attributable impact.
- PCRVs are publicly available – about 25-30 per annum.
- The IFAD budget for a PCR is $20,000, but that does not include government efforts.
- Monitoring improved with the ORMS – everything on line, with ratings and flags, and monitoring of timely reporting.
- Inside IFAD all staff have access to monitoring and self-evaluations. However, PCRs are formally by the governments, so can only disclose if borrowers do not object. Last year about half were disclosed.

## IFAD - Country Program and Strategies

- At the country level, IFAD has country strategic opportunities programmes (COSOPs) and self-evaluation COSOP completion reviews (CCRs). The CCRs focus on assessing the relevance and effectiveness of the COSOP (as opposed to the project level) and provide a rating for it. Evaluation criteria at the country level are: Relevance, effectiveness, non-lending activities. For independent evaluations there are country strategy and project evaluations (CSPEs) that will adhere to the criteria and definitions in the Evaluation Manual (second edition – December 2015) – to assess portfolio performance according to project evaluation criteria, with special emphasis on the assessment of the relevance and effectiveness of both– programs prepared by management units – and non-lending activities.
- CSPEs will also apply the portfolio-level criteria – not done for the COSOP completion reviews (self-evaluations).
- The regional divisions will provide an ad hoc self-assessment as an input to the CSPE unless a CCR is already available.
- Now have 74 active COSOPs/country strategy notes (CSNs) (roughly 50-50). The latter are lighter than COSOPs and for shorter duration or interim periods.

## Annex 6: Interviews Summary

## I) EDs' Offices

| | Performance | Accountability | Learning |
|---|---|---|---|
| **EDs offices**<br><br>Overall:<br>The QA evaluation being carried out by IDEV is very relevant.<br><br>The three outcomes proposed by the evaluation (performance, accountability and learning) are good<br><br>Quality and results will be key for next replenishment | Supervision and SESP fill different purposes. The impression is that they are seen as a box-ticking exercise. The Board doesn't get much information on how the SESP operates.<br><br>Staff turn to "satisfactory" very quickly, even when issues are present; often the narrative and the ratings do not match. There is an issue of trust. CSPs are always satisfactory. The methodology for CSP self-evaluation should change. The Board does not focus much on PCRs but mostly on CSPs. They are however concerned about the disconnect mainly between CSPs and CPPRs<br><br>Not too concerned about the SESP at the corporate level (DER, RMF), even if not always clear whether the RMF is using unvalidated results/ratings for Board reporting? More concerned about project level and logical frameworks. The logical framework is a key design and monitoring tool but of variable quality and it is not used consistently to engage with counterparts and to measure success. The main problem is the (intermediate and final) outcome level.<br><br>The PCR is mostly done by consultants and not much attention is paid to it, so the Board focuses more on the validation by IDEV.<br><br>Provide incentives and reward staff by the quality of logical frameworks; give the most difficult and problematic projects to the best staff; build a reputation of "project fixers". | Could be a good idea that staff does not provide ratings in the PCR but just IDEV during the validation. IDEV may consider doing fewer PCR evaluations but increase field validation rather than just desk review; fewer but better. There should also be a way to validate the task managers' self-evaluation during supervision<br><br>There is a need for some external review of ratings to keep the system honest (regional and country perspective). There is a need to eliminate the perception that the project performance is associated with staff performance<br><br>The lending culture is very strong. What are the incentives to deal with project results rather than only lending.<br><br>Accountability towards the Board is a problem. The Board does not get much information on how decisions get made. Issue of trust in the system. There seem to be little compliance with rules and established procedures or not implemented consistently as the Board has no means to oversee compliance of project cycle SESP processes (IPRs and PCRs).<br><br>It is not clear where the sources of information feeding in the RMF are coming from. This is aggravated by the lack of candour of the SESP which is aggregated at the corporate level.<br><br>The Board needs guidance on IDEV and Bank's rating | There is no learning culture in the Bank and often the Bank's response is defensive ("we are doing it already").<br><br>Evaluation is about learning not about ratings as it should stimulate the sharing of knowledge and best practices.<br><br>Management does not take learning seriously. The fact that certain problems are recurrent is the indication that not much learning takes place.<br><br>Resources should be made available for learning events. Board members could participate in learning events<br><br>The lending culture drives behaviours and undermines learning on what works and what does not work<br><br>Learning and sharing lessons should also be for Governments and project beneficiaries.<br><br>If there is no sharing of information on best practices and lessons learned, the impression is that the Bank is lagging behind other similar institutions |

| | | | |
|---|---|---|---|
| | Board openness to review the definition of the ratings and the 4 points scale again and examine whether just having Satisfactory or Unsatisfactory could lead to a bias towards the better rating. There is a need to revisit the rating methodology<br><br>Some indicators are too abstract and there is an attribution issue.<br><br>Use budget coefficients also for problem projects to motivate and support corrective action.<br><br>Revisit the CSP evaluation methodology esp. the overlap with previous CSPs and on-going portfolio. | systems and rating methodology (4 point scale?) particularly in APPR, CPPR.<br><br>Procedures, guidance and rules are in place, the main problem is their implementation and lack of accountability. The Board's comments are recurrent, but nobody seems accountable for addressing them.<br><br>Does not think that there is push back by management but often they don't know how candid the system is. Often giving staff the benefit of the doubt.<br><br>There is clearly a conflict of interest in having task managers preparing, supervising and then self-assessing their own work. Clarify who should do what. | |

## II) SESP General interviews

| | Performance | Accountability | Learning |
|---|---|---|---|
| **RDVP** | The evaluation reports on Quality at Entry and Quality of Supervision were mostly on assessing compliance and reporting and did not discuss organization and institutional issues specifically the role of the country portfolio officer and the regional program officer's role in Country Portfolio Performance and in Country Program Improvement Plans. Their role is to provide a proactive support to complement the feedback (learning) loop from portfolio implementation issues.<br><br>The project TM and the CPO have different mandates in terms of project management and no coordination is occurring so as to timely discuss project and portfolio day-to-day implementation issues. This is the reason why the Implementation Support Unit was established in each region since last year (2018) to ensure coordination and monitoring.<br><br>The implementation Support Unit Manager is responsible to address systemic issues after project approval at the Regional and Country level to enhance project implementation including Borrower institutional capacity, fiduciary, financial management issues and ensure smooth closing of projects funded under special accounts and timely PCR preparation.<br>Its role is to bring project actors together while the sector manager will be in charge of business development.<br><br>For the private sector projects (NSO), portfolio management is still centralized. Although KPIs in terms of lending (business development) are mostly for the region, the NSO portfolio management is still not decentralized. Investment officers report to PINS (Centralized at HQ) but | If the CSP-CR has to be reviewed and its quality validated by IDEV, this should be a real time tool to inform the new strategy. Timing is crucial. Actually, the CSP-CR has no rating but should be in the dash-board, transparent and informative to help assist in revising the CSP and to inform the new CSP. In that regard, the whole ecosystem of the CSP should be looked at to strengthen the accountability system and learning from past experience and also to prepare for the next country assistance strategy<br><br>With regard the CSP-CR, Board complained about the flawed methodology and utility of addressing at the same time the portfolio management issues and the country assistance strategy in one single shot. Contradiction of satisfactory country strategy and an unsatisfactory CPPR has created confusion with the Board requesting an appropriate methodology of the CSP-CR and CPPR instruments. Parallelism between CSP performance and CPPR performance should be discussed and methodology agreed with the Board<br><br>Accountability mechanisms should be reviewed. Dual accountability or at least some oversight function could be performed by the Implementation Support Unit Manager (New). More of a team approach could improve candour and rigor.<br><br>The implementation Support Unit would now be capable to engage with the sector divisions and provide better harmonization across sectors for better accountability. There will be a resources issue though since there are too many projects for the staff | Lessons learning from portfolio management does happen at the region but is diffused for project management where no single location to learn (CPO, RPO, Sector and Task managers). The role of the Implementation Support Unit is to all for enhanced and coordinated learning at all these levels.<br><br>CSPs are not rated by Mgt. nor by IDEV. However, the narrative raises issues as needed. Better to keep it that way so as not to create un-necessary tension around ratings and improve learning opportunities.<br><br>Ratings are already available for the projects that are part of the CSP portfolio but CSPs evaluations by IDEV are not done at the right time.<br><br>There are specific sections already in the CSP about lessons learned including those drawn by IDEV<br><br>The PCR should not be seen as a tool to judge the TM, otherwise it will undermine candour in ratings and learning opportunities. An external look over the PCR is always good. Letting IDEV provide the PCR rating in consultation with the TM could be a |

| | | | |
|---|---|---|---|
| | will be increasingly decentralized.<br><br>Compliance is a problem. The Bank has changed a lot and many new people came on Board who are not familiar with procedures and do not know how to manage and restructure projects. Operations Academy will address this issue.<br><br>TMs are overwhelmed. They are the jack of all trades and they don't get enough support. Playing catch-up game all the time. Most incentives are towards approval rather than results.<br><br>On candour, there is a need to separate the performance of a project with the performance of staff. The TM should not be abandoned to him/herself but provided support as needed, while understanding that the Borrower is actually the one implementing the project.<br><br>Issues should be discussed with the Borrowers but not the ratings. Also, ratings are too categorical; S or U does not reflect the reality. The methodology needs revisiting with more flexibility. IPRs don't seem particularly useful to address issues, MTRs are much more seen as a way tool for corrective action.<br><br>Continuous change in Government priorities is a big issue that has consequences on the required staff profile and the SESP methodology to deal with this issue has to be harmonized with IDEV<br><br>On strategy papers, the process has just been overhauled with the Chief Economist in charge of the diagnostic work and CSPs being guided by a stronger analytical base and rigor.<br><br>The RISP is not really a strategy document per se but more a way to rationalize how to allocate regional resources to | available. It is not entirely clear who should approve the ratings now, the sector Manager or the Implementation Manager.<br><br>Ratings make the discussion unnecessarily contentious. Let the outcome speak for themselves. Also, there is a need to harmonize methodologies with IDEV otherwise we will be reporting different things to Management and to the Board<br><br>Record management has moved to electronic but not in place yet. If you go to SAP you will not find the ratings. Lag in producing various reports is still important. Actually, IPRs are uploaded in Sharepoint software while project data (old system before 2011 approval) are on SAP. No centralized database. | good option. The TMs cannot evaluate their own work. |

| | | | |
|---|---|---|---|
| | regional projects. It would rely on CSPs as strategy documents. RISPs overlap several CSPs in the same country | | |
| **SNVP** | SESP is an institutional issue.<br><br>Focus on implementing recommendations rather than re-opening the same issues<br><br>Looking at the reliability of the ratings through the dashboard.<br><br>Portfolio dashboard inserts flags according to verifiable performance indicators and can categorize projects in problem status even if the task managers thinks that all is well<br><br>Incentivize project restructuring with support and money. Reconcile speed of delivery with quality<br><br>IPR timeliness and validation is not captured by the dashboard, but it could.<br><br>There is an issue about candour and staff overload<br><br>President Directive 02-2015 on design, implementation cancellation of Bank group's sovereign operations is a good directive and should be complied with.<br><br>Portfolio ratings do not distinguish between public and NSO. It then aggregates everything in one average but the nature of the business is very different. Same for budget support operations<br><br>No incentive for TM to give a bad rating. IPRs are not on-line and half of them are not validated by Managers. Currently ratings are not linked to SAP but this will happen under the RRS. Definition of DO and IP not the same in SAP and IPRs including the rating system. | The Portfolio dashboard (and quarterly portfolio flash) is an accountability mechanism that ensures better compliance and provides a framework for enhanced proactivity on corrective action.<br><br>The dashboard report is in the Performance Agreement of Sr Managers with the President. It could be made to cascade to Managers.<br><br>The system will allow fixing issues before it is brought to the attention of Sr Mgt.<br><br>Accountability could be more shared with the country team; but SESP and ratings should not involve the Borrower. A results dashboard (RRS) will complement the actual delivery dashboard<br><br>Roles should be different between design, implementation and validations otherwise conflict of interest. Implementation support is more of a Borrower's issue. Some arm-length review is necessary for IPRs, like what is happening for the private sector operations. The role of the new Implementation support unit manager could help address this issue.<br><br>CPO involvement in rating system may displace accountability<br><br>May need a KPI on supervision. Proactivity and realism are there, but it does not really work. Accountability issue may be clarified when moving from supervision to implementation support activity. Borrowers may be part of the process with independent reviewers.<br><br>The role of OPSCOM is to ensure high quality assurance | Learning is very different ball game for NSO.<br><br>XSR format is not adapted. Mostly done by financial experts, not equipped when it comes to results and impact<br><br>Review of sector reports especially by IDEV could be an opportunity for learning. Could be done more by the Borrowers through RTA.<br><br>No real SESP on sector reports, but IDEV does review some<br><br>The preparation of strategy papers (country or thematic) are more conducive to learning. PCRs do not always allow to extract a credible story that can be used or replicated somewhere else.<br><br>Need to do better PCRs or change the format so that it is not seen as a box-checking at the end of the cycle |

| | | | |
|---|---|---|---|
| | Ratings on DO are composed of a number of sub-indicators which generate an average but people do not know where the average came from. The range provided for the average rating (from 2.5 to 3.49 is satisfactory) leads to "gaming" the system.<br>Ratings assessment methodology needs review<br><br>RRS will allow to integrate data (on Sharepoint before) and provide a platform for accessing all project data.<br><br>The rating scale may need to be reviewed again to accommodate for a more nuanced view of projects performance rather than just S or U. This may allow for a more candid reflection of reality and avoid an induced bias towards better ratings. | mechanisms before Board approval but little on SESP (consistency of logframes and compliance with safeguard guidelines, PD 02.2015 etc..)<br><br>The Board is overall happy with the RMF even if it may come across as too rosy but has more issues with the results framework of individual projects, whose readiness is questionable.<br><br>There is a need to disconnect staff performance with project performance as an incentive for candour.<br><br>On strategy papers there has been lots of improvements on accountability with a number of tools. RMF three color-coded system seems to work.<br><br>Major strategy papers get self-assessed and then validated by an independent source before going to the Board with a new strategy | |
| **ECVP** | Evaluation of processes is part of good governance.<br><br>The formats and templates do not take into account the different business models of the private sector and the budget support instrument. Many sections of the IPRs are not adequate as well as the reporting frequency for PBOs esp. when they only have 1 or 2 disbursement tranches. When the project goes to the Board, conditions have already been met. Much more importance needs to go to design. There is no space for IPRs. PCRs get made for a cycle (2-3 years) to account for policy dialogue<br><br>No standard criteria and guidelines for BTORs/Aide memoires. So it is quite easy to "game the system". Aide memoires are sometimes redundant.<br><br>Results-based finance could be the solution | The system is largely paper-based and there is no IT-supported central system.<br><br>In the RMF, core governance indicators are only referring to the CPIA and the Mo Ibrahim Index.<br><br>Also on safeguards, standards are required and the work should be cleared by safeguard experts at arm-length from operations<br><br>The best staff should be assigned to the most difficult projects (problem solvers) to incentivise corrective action and balance the excessive focus on the lending side. De-link project performance from staff performance | The key question is how to continue policy dialogue and learning after the very short duration of PBOs projects.<br><br>The paragraphs on learning have become generic and perfunctory |

| | | | |
|---|---|---|---|
| | Missing gap in the use of resources and delivery. Link delivery to the budget.<br><br>ADOA uses its manual to provide its own ratings at origination stage. Ratings are independent from the origination Unit. The Board receives three documents: i) the project, ii) the ADOA Note and iii) the Credit Note approved by the Credit review Committee (CRC). | Once the operation is approved the ADOA team is out of the picture as ADOA does not get involved in supervision; maybe it should. There is a need to harmonize the templates so that the two systems can talk to each other. Once the operation is approved, information does not get collected to inform the ADOA framework any more (follow up on ADOA ratings). | |
| **PEVP** | Public sector projects are now managed by Regional Offices (Nairobi, Abidjan for energy) and they are in charge of supervision under the new Delegation of Authority framework.<br><br>Staff tend to over-rate their projects but the Dashboard is now raising flags on-line in the system which makes it more difficult to distort portfolio performance. However, this mostly focuses on disbursement and procurement which do not provide a complete picture of project performance. In any event the Dashboard seems to be more useful than the IPR system.<br><br>Most of the time, performance issues stem from initial delays and over-estimation of existing capacity. This seems to point to readiness issues and reduced time for project preparation which is then difficult to catch up with through supervision.<br><br>Formal restructuring is rare, both because the transaction cost is too high and the Borrower is not always forth-coming, fearing to add delays. There is a need to revisit restructuring procedures and make them more flexible. Also, necessary to realign the evaluation methodology applicable to revised targets.<br><br>IPRs are not really helping to take corrective actions. They | IPRs are of little value, Management approves them automatically, they are not used as a management tool and there is no system to disseminate information. This does not help accountability mechanisms.<br><br>PCR should be done independently. IDEV may just take responsibility for preparing and/or rating the PCR. But the may need additional resources to better consult with the TMs and obtain better knowledge of the field situation. | PCRs are good tools for learning especially in the same country for follow up projects, CSP and CPPRs. However there is a need for more creativity in terms of dissemination and sharing lessons with other countries on the same sector or similar specific issues.<br><br>IDEV should take the initiative to organize learning events including their Evaluation Notes or corporate evaluation studies and add new dimensions to sharing lessons (rather than having to read bulky reports). This could also be done jointly with Management. |

| | | | |
|---|---|---|---|
| | are often not filed in a timely manner and they are more of box-ticking exercise. MTRs on the other hand seems a more dedicated and opportune moment to look at mid-term adjustments. Ratings are not necessary to do that. | | |
| **Audit** | There is no driver (KPI) for quality, therefore compliance goes down and candour becomes a factor.<br><br>Look at the ratio of Number of projects per TM to assess the quality of supervision and of the SESP<br><br>They do audit PCR compliance but not IPRs<br><br>CSP is a key control with IDEV piloting validation | For accountability one needs some arm-length review process over the ratings, similar to the private sector where the IO and the portfolio mgt are different.<br><br>Sector people need to do a better job, bringing CPOs in for rating validation will not resolve the issue<br><br>IDEV could validate IPRs on a pilot basis | CSP do take learning into consideration more than projects |
| **Task Managers** | The Implementation Progress and Result (IPR/EER) Report is useless as it is developed outside SAP in Sharepoint drive with no connection to SAP where all project data is located. It is normally prepared after the aide memoire and BTORR which are annotated by the Manager who will not review the IPR nor validate the rating if any.<br><br>The rating system may or may not reflect the reality of the project implementation performance depending on the seriousness of the task manager who, by nature, would present an overoptimistic view of the situation.<br><br>Countries change priorities all the time but typically projects are not restructured to reflect the changes and the TM cannot proceed without the Borrower's agreement<br><br>The project supervision of the sample project did highlight crucial issues but were not timely addressed or acted upon specifically in the absence of a performing project coordination unit. The country did not have the capacity to implement the project nor to provide the right direction for a smooth implementation with a weak steering committee. | Incentives are not there to encourage staff to produce timely IPR and PCR with evidenced performance and enhanced learning from experience.<br><br>The IPR and PCR were not considered as a decision-making tool but rather a compliance tool to report on issues with no sufficient accountability or attribution (administrative requirement).<br><br>It is not the role of the Task Manager to prepare the PCR as a conflict of interest is always there mixing up the performance of the project with the TM's own performance. Generally, PCR is done by the consultant under the supervision of the TM.<br><br>The delivery dashboard is currently what exerts the most pressure to ensure accountability. If the performance is publicly visible, the Managers call the TM with questions | The learning part of the IPR may be helpful (if it is in the system) as IPRs are built on previous ones prepared.<br><br>The self-Evaluation system as used inside the Bank is worthless and does not lead to enhanced learning by TM and sector managers. It seems that the IPR and PCR tools are mechanical and do not constitute an opportunity for learning specifically in the softer part of the project.<br><br>Post implementation and achievement of project should be done after 5-6 years of completion to ensure the sustainability of outcomes and probably undertake an impact assessment of project results.<br><br>Projects typically offer very good lessons which are not capitalized. The new CSP could make use of the |

| | | | |
|---|---|---|---|
| | In this case, IPR ratings were agreed with the Government. This may increase ownership but reduce flexibility as Managers will have even less of an incentive to question the ratings.<br><br>The performance of project outcomes was 50-50, so the project was given a Satisfactory status. The project was downgraded by IDEV's validation but it was considered that IDEV's PCREN is useless because too theoretical and does not include field work. | | program to inform new projects in the country. |
| **CPOs** | CPOs are not necessarily the solution to improving the performance of individual projects. That function should rest more with the TM, the sector experts and the safeguard/fiduciary teams that are part of the supervision team. CPO has more of a role in following up with the Government on implementation issues such as disbursement and procurement and portfolio performance when it comes to the CPPR<br><br>In some cases, CPOs are associated to the project supervision but more often they are not. It is very contextualized and it depends on the CPO's work load.<br><br>In general CPOs seem sceptical in taking up a more active role on project supervision and they would rather keep some distance when it comes to prepare CPPRs/CSPs.<br><br>CPOs do not really intervene in the ratings of the projects. Sometimes they provide comments for the Aide Memoires and the IPR ratings. In one case the CPO advised the Task Managers that the IPR ratings were over-optimistic and did not reflect the reality but the TM entered a "S" rating in SAP anyway.<br><br>Increasing the credibility of the ratings will have to rely on having stronger teams, better staffed with the right expertise, and possibly some additional oversight, but | There is not a clear guidance and expectation on roles and responsibility between the TM and the CPOs. It is very ad hoc and it depends on where the TMs and CPOs are located. CPOs mostly responds to Country Managers (CM) in the way they monitor portfolio issues, disbursement, procurement, delays, dealing with flags raised by the Dashboard, but they cannot systematically participate to project supervision. Their participation would not necessarily raise the credibility of the ratings unless there was a process of dual accountability in place, but they doubt that CPOs have time for that.<br><br>However, the introduction of the new Implementation Manager could help keeping an eye on the project, raising issues and make the TM and Sector Managers aware about problems that need to be dealt with including possibly on ensuring compliance with requirements and candour of the ratings being provided. A better division of labour and guidance is still needed in terms of expectation from each player.<br><br>Currently IPRs are approved automatically in the system without the need to provide a comment by the Manager (like in other institutions).<br><br>Better defined KPI may help redressed some of the | When it comes to CSPs/CPPRs they take the ratings that are posted in the system (IPRs/XSR and PCRs) but they are not aware of IDEV's Evaluation Notes (PCRENs/XSRENs) and of the different ratings which are often more critical.<br>CSPs completion only looks at completed projects not on-going projects.<br><br>Therefore, opportunity for learning from an independent source of evaluation is minimized and CPPRs and CSPs often start with over-optimistic assumptions and credibility issues.<br><br>IPRs/MTRs/PCRs are useful document for learning especially for the teams working in the same country. Probably less for staff in other regions as they may think that lessons are either too generic and not actionable or too specific and not replicable.<br><br>For CSPs one needs to consider i) the pipeline, ii) analytical work and iii) |

| | | | |
|---|---|---|---|
| | often the resources are not available for that. The question is "what is the incentive structure for objectivity".<br><br>The usefulness of the IPR is not obvious, since there seem to be an issue of compliance and staff tend to vote with their feet. There may be scope to simplify the IPRs just stating the fact and providing the ratings, but there is no evidence that even that will be seen as a trigger for corrective action.<br><br>If issues cannot be fixed through IPRs' ratings, what's the use of it ? The role of the Borrower is predominant in the capacity to restructure a project.<br><br>However, clearly CPOs are an important stakeholder since at some point the projects will end up in their portfolio that they are responsible for carrying the CPPRs on.<br><br>CPPRs have to rely on the quality of project supervision and use the indicators being provided by the system. The CPPR does not have an overall rating but includes individual ratings on specific indicators as found in SAP or the Dashboard.<br><br>On the CPPRs, only the PCR ratings are considered and not the CPREN ratings provided by IDEV. This excludes therefore a more independent and often critical view of the quality of the projects. This will eventually also affect the CSPs that are in large part influenced by the quality of the CPPR. All this will eventually escalate in how indicators will be aggregated and presented to the Board through the ADER.<br><br>RBFs are not used for M&E purposes. The quality of the results-based framework at project of CSP/RISP levels needs to be reviewed for better assessment of results. However, the time horizon of CSPs and portfolio reviews are different which provide a distorted picture of | incentives that are now skewed towards lending rather than supervision and achieving results. However, even in that case, teams would require more support from a dedicated supervision support unit.<br><br>It is felt that having the same person being in charge of designing the project, supervising it and eventually rate its performance is riddled with conflict of interest and not conducive to candour.<br><br>It was argued that the Bank is a financial institution which does not implement projects and therefore it is unfair to look at accountability only from the perspective of the Bank's team rather than the Government's teams, where accountability is very poor. Therefore the project performance should not reflect poorly on the staff performance, as it is often the case.<br><br>In any case project teams require more support for supervision when needed, staff workloads need to be revisited and some shift of priority from lending to supervision.<br><br>The new arrangement with a shift of accountability towards the Implementation Management Unit is still un-tested and creates some confusion as clearly there is an issue of capacity and the Implementation Manager (who covers all sectors) cannot substitute the technical knowledge of the Sector Managers.<br><br>Also, the IM should be involved at quality at entry as well, otherwise they will just inherit problems that will be more difficult to fix. The sector managers sometimes are pushing to accelerate delivery through the Board even if the project is not ready and then the TMs suffer the consequences. | results.<br>The preparation of a separate Diagnostic Note under the responsibility of the chief economist may help. Not clear how the evaluation of these two products will work but the CDN should be assessed in its capacity to inform the CSP and whether we can better learn from them (CDN and CSP).<br><br>SESP help assist in operational learning but little at strategic level.<br><br>Building capacities on new processes for RBM is crucial including on safeguards and cross cutting issues. |

| | | | |
|---|---|---|---|
| | performance between the portfolio and the country/regional strategies. Portfolio reviews/analysis should be aligned to CSPs/RISPs SESP timeline.<br><br>Untimely achievement of development objectives reduces the validity of IPR ratings. CPPRs should be synchronized with CSP/RISP timeline and PCREN should be feedback to staff and considered for official ratings in CPPRs. RMF should also be cascading to allow for meaning IPR reports on results/Dos (results matrices to be adapted to country and region context with M&E reporting systems). Results frameworks lack smart indicators mainly on the social sector (poverty assessment, vulnerability, gender disparities, ect..). CSPs ad RISPs pillars alignment to the High Five priorities I an issue (7sectors) whereas one pillar may make a difference.<br><br>Results framework should be flexible and realistic to ease monitoring with effective data collection methodologies | The role of CPOs is fundamental in assessing portfolio performances which with the assistance of the delivery unit allows for overall assessment of performance at country and regional levels. However, projects alignment to the strategic objectives and pillars is fundamental. There is need to increase the budget and means for TM to enhance the quality of IPR, PCR, CPPRs.<br><br>IDEV CPEs evaluation reports findings are reflected in new design of CSPs and RISPs but there is need for a joint work with vice presidencies front offices across regions to enhance the credibility and explore new solutions for high quality SESP.<br><br>The role of the implementation support division is crucial in ensuring day to day support and monitoring project portfolio performances, procurement, disbursement etc… | |

## III) Private sector NSOs

| | Performance | Accountability | Learning |
|---|---|---|---|
| **PINS** | The Portfolio Mgt and Monitoring unit provides support (quality portfolio management support) to ensure bankability and structuring to the origination team with some level of compliance through the annual portfolio review.<br><br>Budget KPIs are still skewed towards lending and new transactions. Very little money for XSRs<br><br>Very important to ensure quality at entry; garbage in garbage out principle. Supervision and XSR new templates are being designed and will be rolled out soon to take into account the new eco-system (corporate portfolio management advising sector managers and origination team to comply with Bank policies, business model during the project cycle, new operations guidelines).<br><br>Active monitoring is done in joint venture between the various departments and PINS who decide on the active monitoring with intelligence on the ground and preparation of PSR/XSR with particular attention to environmental and social impact of NSO projects.<br><br>New templates of PSR and XSRs will be put in place to better track development impact with detailed processes, standards, training on norms to enhance compliance mainly for reporting and review of NSO performance results. New guidelines are also prepared for valuation of Bank assets in the books. All in all, there is a paradigm shift of how NSO projects are designed and managed with a corporate portfolio repository | The Portfolio Mgt and Monitoring unit ensures watch-listing and flagging projects and submission to the special operations unit (work out division). Flagged projects are subject to meetings on a monthly basis and quarterly monitoring report on risk management (liaison with CRC on the headroom)<br>An on-line platform where all documents are loaded provides visibility and transparency. It went life last year.<br><br>It is not in the interest of the Investment Officer to avoid candidness because if something goes wrong he/she will be the first one to get exposed.<br><br>There is no measurement of the disconnect with IDEV as no thorough discussion is held on early maturity (XSR due date) and XSR-EN preparation.<br><br>The department is underway to harmonize ADOA process (additionality) and development outcomes measurement<br><br>NSO portfolio management is actually decentralized with portfolio officers at HQ et in the region under sector managers or PAIFD or Energy Department. Consolidated portfolio supervision report for agric, industry and services and transport, is prepared by PINS every 6 months based on PSRs and related reports (aide memoirs or BTORR).<br><br>Accountability of NSO portfolio management lies under the Credit and Risk Management Committee who ensures a strict follow up based on self-assessment and monitoring of NSO projects in the various departments (evolution of the portfolio, risk capital, average capital consumption, etc…).<br><br>NSO Watch List meetings are held under PINS to decide on the ratings and NSO operations to be managed by the Special Operations Unit (‘SOU’) under the Senior Vice President office of the AfDB. The SOU | Lessons drawn in NSO PSRs/XSRs are checked by sector managers in departmental meeting during which private sector development officers would normally attend to enhance the feedback loop. However, there is no consolidated LL database to be accessible by NSO development and portfolio management officers.<br>A new system on capturing LL from PSR and XSR with new portfolio management report. |

| | | | |
|---|---|---|---|
| | system making PINS the stop shop for data on projects from inception until independent evaluations (XSRENs). | provides specialized knowledge and services to resolve Non-Performing Loans in the Non-Sovereign Operations of the Bank ('NSO') including distressed and underperforming investments and assists with the identification and management of investments on the Bank's watch list. The SOU's main responsibility is assessing underperforming and problem investments and leading the restructuring, workout, turnaround or in some cases re-capitalization of companies.<br><br>XSRs due dates are set without strict follow up of early maturity dates but with appropriate judgment on the maturity date based on targets (no. of XSRs by year) to be achieved (Nota Bene: although guidelines for XSR preparation are in splace, there is a cap on budget resources to be used for XSRs preparation. List of XSRs to be prepared is shared with sector managers.<br>XSRs are discussed by the whole ecosystem and validation of XSR ratings is done by sector managers and PINS.<br><br>Due to the heavy administrative workload of NSO development and portfolio management officers (10-15 projects by PO), quality of QaE and SESP is suffering. Decision underway to restructure the structuring procedures by creating a new economic division including ADOA and quarterly reporting on results | |
| **Regional NSO portfolio officers** | Lessons learned are not captured by investment officers in new deals. The NSO SESP favour personal learning as NSO learning is not institutionalized.<br>Logframes and ADOA reports are not used as M&E tool. ADOA, although a principal path for NSO approval, is not a good predictor of results (only few outcomes are captured and followed up by clients). SESP for NSO should be results driven and not only operationally driven (mainly at PSR). | Portfolio managers are too much busy with loan administration and supervision and is not better placed to undertake XSRs (no interest). These are generally done by consultants. Third party assessments (independent, external peer review) may provide more accurate and candid assessment od what has worked or not as the role of NSO management does not encourage for candour and validity of ratings | The learning is not taking place in NSO activities. Actual NSO SESP do not allow for cumulative knowledge (lack of continuity of focus on results), do not provide value addition to portfolio managers' work (loan administration with no interaction with investments officers to convey lessons of experience) and are not part of the Bank's institutional memory (not used for knowledge management purpose) |
| **Consolidated Portfolio** | Active monitoring is done in joint venture between the various departments and PINS who | NSO portfolio management is actually decentralized with portfolio officers at HQ et in the region under sector managers or PAIFD or | Lessons drawn in NSO PSRs/XSRs are checked by sector managers |

| | | | |
|---|---|---|---|
| **Management Division. Regional Hubs (Manager and Officers)** | decide on the active monitoring with intelligence on the ground and preparation of PSR/XSR with particular attention to environmental and social impact of NSO projects.<br>New templates of PSR and XSRs will be put in place to better track development impact with detailed processes, standards, training on norms to enhance compliance mainly for reporting and review of NSO performance results. New guidelines are also prepared for valuation of Bank assets in the books. All in all, there is a paradigm shift of how NSO projects are designed and managed with a corporate portfolio repository system making PINS the stop shop for data on projects from inception until independent evaluations (XSRENs). | Energy Department. Consolidated portfolio supervision report for agric, industry and services and transport, is prepared by PINS every 6 months based on PSRs and related reports (aide memoirs or BTORR).<br>Accountability of NSO portfolio management lies under the Credit and Risk Management Committee who ensures a strict follow up based on self-assessment and monitoring of NSO projects in the various departments (evolution of the portfolio, risk capital, average capital consumption, etc…).<br>NSO Watch List meetings are held under PINS to decide on the ratings and NSO operations to be managed by the Special Operations Unit ('**SOU**') under the Senior Vice President office of the AfDB. The **SOU** provides specialized knowledge and services to resolve Non-Performing Loans in the Non-Sovereign Operations of the Bank ('NSO') including distressed and underperforming investments and assists with the identification and management of investments on the Bank's watch list. The **SOU's** main responsibility is assessing underperforming and problem investments and leading the restructuring, workout, turnaround or in some cases re-capitalization of companies.<br>XSRs due dates are set without strict follow up of early maturity dates but with appropriate judgment on the maturity date based on targets (no. of XSRs by year) to be achieved (Nota Bene: although guidelines for XSR preparation are in splace, there is a cap on budget resources to be used for XSRs preparation. List of XSRs to be prepared is shared with sector managers.<br>XSRs are discussed by the whole ecosystem and validation of XSR ratings is done by sector managers and PINS. Due to the heavy administrative workload of NSO development and portfolio management officers (10-15 projects by PO), quality of QaE and SESP is suffering. Decision underway to restructure the structuring procedures by creating a new economic division including ADOA and quarterly reporting on results | in departmental meeting during which private sector development officers would normally attend to enhance the feedback loop. However, there is no consolidated LL database to be accessible by NSO development and portfolio management officers.<br>A new system on capturing LL from PSR and XSR with new portfolio management report. |

# IV) CSP and RISP Case Study

| | Performance | Accountability | Learning |
|---|---|---|---|
| Regional lead Economists | CSP/RISPs SESP do not apply a numerical rating system for assessing performances but play a fundamental role in addressing the pitfalls and redirecting the strategic objectives specially at mid-course or in designing the new country or regional strategies. The CSPs/RISPs SESP provide the basis for the analysis of strategies' implementation risks related to performance and results. Disconnect in ratings with CPPRs and other mid-term reviews is due to the fact that CSPs/RISPs preparation teams work in silos. CSP/RISP teams involved in SESP should be normally aware of the discrepancies with portfolio ratings. To solve the disconnect issue, supervision and mis-term reviews should be systematically and regularly undertaken with realistic portfolio improvement plans.<br>Rating systems should be more insightful based on analytical assessment and provide more explanation on why results were not delivered. This is due to the ack of robust logframes or results frameworks.<br>RECS and implementation agencies and beneficiaries' involvement in SESP preparation may increase their validity based on agreed KPIs and deliverables. Capacity development of RECS (including in statistics and impacts assessments) is necessary to enhance the quality of design and follow up of strategic development objectives. | Regional Directorates play a fundamental role in adopting an operational approach to CSP/RISPs SESP by engaging ADB staff in adapting their strategies to the country regional context. The regional delivery unit (see regional program coordinators) helps assist the regional directorate in reviewing what has worked or not in terms of country or regional strategic achievements and results including the risk factors around them. | Peer reviews and independent review process encourage the learning from CSPs/RISPs SESP and improve their quality together with the Economic and Sector Work that the CISPs/RISPs SESP recommend to recommend to enhance their evaluability and results. The involvement of regional economists in the design and implementation phases of CSPs/RISPs are minimal but contribute mainly to the knowledge. This was corrected by the new CSP/RISP guidelines |
| Sector Managers (Regional Directorate) | CSPs/RISps SESP wrongly attribute country or regional achievements results to the Bank without highlighting the country or regional capacities in implementing CSPs/RISPs. Also, the comparison of Bank's achievemetns with those of other financiers is self-defeating (auto-flagellation). SESP should clearly cover | The lack of staff in sector divisions does not allow for continued strategic and operational dialogue and for increased supervision and just-in time actions to address project or program implementation issues. The SESP at project level is restricted to operational issues and only mid-term reviews provide an opportunity to | The high turnover of CPOs and TMs creates a caveat for continuous learning and building the Bank's institutional memory. i |

| | | | |
|---|---|---|---|
| | attribution/contribution analysis.<br>Disregarding numerical ratings in assessing implementation performances may encourage better dialogue and focus on strategic objectives.<br>IPR template should be reviewed to include sufficient data on performance and consultations/coordination with implementing agencies should be encouraged to increase the SESP validity. | reflect on the adequacy of Bank's response to country/region strategy.<br>The Bank should recognize that the SESP are only for addressing implementation issues and allowing mistakes to be corrected for the good of the country of the region.<br>There is no real incentives for TMs candour and objectivity. | |
| Country CPO | RBFs are not used for M&E purposes. The quality of the results-based framework at project of CSP/RISP levels needs to be reviewed for better assessment of results. However, the time horizon of CSPs and portfolio reviews are different which provide a distorted picture of performance between the portfolio and the country/regional strategies. Portfolio reviews/analysis should be aligned to CSPs/RISPs SESP timeline. Untimely achievement of development objectives reduces the validity of IPR ratings. CPPRs should be synchronized with CSP/RISP timeline and PCREN should be feedback to staff and considered for official ratings in CPPRs. RMF should also be cascading to allow for meaning IPR reports on results/Dos (results matrices to be adapted to country and region context with M&E reporting systems). Results frameworks lack smart indicators mainly on the social sector (poverty assessment, vulnerability, gender disparities, ect..). CSPs ad RISPs pillars alignment to the High Five priorities I an issue (7sectors) whereas one pillar may make a difference.<br>Results framework should be flexible and realistic to ease monitoring with effective data collection methodologies | The role of CPOs is fundamental in assessing portfolio performances which with the assistance of the delivery unit allows for overall assessment of performance at country and regional levels. However, projects alignment to the strategic objectives and pillars is fundamental. There is need to increase the budget and means for TM to enhance the quality of IPR, PCR, CPPRs. IDEV CPEs evaluation reports findings are reflected in new design of CSPs and RISPs but there is need for a joint work with vice presidencies front offices across regions to enhance the credibility and explore new solutions for high quality SESP.<br>The role of the implementation support division is crucial in ensuring day to day support and monitoring project portfolio performances, procurement, disbursement etc… | SESP help assist in operational learning but little at strategic level.<br>Building capacities on new processes for RBM is crucial including on safeguards and cross cutting issues. |
| Regional Coordinator for climate change | Cross-cutting issues such as climate change, gender, social and environmental safeguards are not well analysed at inception nor at SESP during implementation | Regional coordinators are few and cannot cover all countries and regions. They are not systematically involved in field missions or meetings of CSP/RISP | TMs mostly value the interaction on climate change and environmental and social safeguards. Undertaking special |

| | | | |
|---|---|---|---|
| and safeguards | phases of project and country/regional strategies. Risks to adaptation to climate change not reflected or embedded in assessing performances or National Determined Contributions (country commitments). The M&E framework with indicators (2016-2020 update) provide an operational guidance for reporting on climate change issues that need to be addressed at project or CSP/RISP levels. CSP/RISP SESP should also report on processes to deliver sector results on social and environmental safeguards. These processes should be described and documented at inception (early engagement of safeguards during post approval) with adequate risk assessment. | meetings although they are custodian of the Bank's climate change action plan (2011-2015). The new 2016 action plan has entered into force with no tangible improvement although guidance for mainstreaming climate change was issued. Only 2 paragraphs are allocated to these issues which may devote an annex to CSP/RISP reports and complete reporting on achievements in SESP. Climate change, social and environmental safeguards should be used as support team to project and CSP/RISP preparation and implementation and involved in early stage of project/CSP/RISP design (pre-board stages) for better recognition of the subject matter.<br>The CSP/RISP SESP should report on countries commitments (NDCs) and capacities in CC adaptation and resilience in safeguards. This would help countries' accreditation for green climate fund.<br>There is need to increase interaction with ECON complex and build capacity to adequate and sufficient staff address E&S safeguards at country or regional level. Countries and RECs should be involved in SESP to report on E&S safeguards with adequate reporting guidelines on E&S management/mitigation plan.<br>Status of E&S safeguards in portfolio flagship reports is missing. | supervision solely on these issues is too late, not productive or influencing. Training of TMs on integrating these issues should be continuous because of the staff turnover and procedures simplified for better mainstreaming. There is need to increase learning from CSP/RISP case studies on E&S safeguards and CC adaptation and also on SDG and mainstreaming guidelines of S&E should reach out countries and RECs. Administrative budget for supervision should be increased to allow for more knowledge generation of E&S safeguard results at project, country and region. |
| Regional Implementation Support Division | CPR and CSP SESP are strengthened through CPIP which increases their relevance. However, difficult to attribute results to the Bank due to the lack of valid information on country results and project data. Incentives to staff to provide candid assessments are missing. Supervision should be mostly focusing on operational and development risks. The dashboard is mostly oriented towards procedures and not focused on risk factors of implementation and outcome achievement. | Aide memoirs and BTORRs are the main documents used for project SESP. Few sector division managers validate IPR ratings (although in the workflow). Training TMs on M&E is essential. Also, involving country officials in rating performances at supervision, mid-term and completion help increase the validity and objectivity of ratings.<br>Implementation support division is well placed to enhance the credibility of IPR, CSP-MTR in providing comments on the implementation results and anticipating portfolio implementation issues. However, there is an overlap of responsibilities with sector | Little knowledge sharing out of supervision and IPRs or mid-term reviews and completion reports. Need for thematic reviews of IPRs and SESP for better sharing and dissemination of results. |

| | | | |
|---|---|---|---|
| | | divisions. | |
| Regional Gender Coordinator | Gender issues are not sufficiently covered in projects, CSPs or RISPs from design to completion due to weak methodology in anticipating and assessing results (lack of outcome indicators) as gender experts are not systematically involved during project cycle and CSP/RISP mid-term reviews or RISP/CSP-CRs. Some countries have benefitted from thorough assessments of gender profile and helped in designing gender policy strategies and action plans. Gender profiles should be systematically done to provide the right outcome indicators in projects, CSPS and RISPS results frameworks.<br>Gender issues were not dealt with correctly in regional integration due to lack of resources and insufficient design of governance programs which did not systematically include gender dimensions. | Few gender specialists are located in regions. Gender policy guidelines have to be reviewed as no clear plan of action is in place to report on results in SESP. No dedicated budget is in place to cover specifically gender issues during preparation, negotiation down to completion.<br><br>A categorization system of gender issues in the country or region, if put in place based on gender profiles, will ease follow up and supervision and increase the credibility of SESP projects or CSPs/RISPs in report on gender.<br><br>The institutional framework should be reviewed to enhance gender experts' involvement in CSPS and RISP assessments, policy dialogue, validations and corrective actions mainly of safeguards (example compensation in resettlements). The implementation support divisions should be well staffed with gender experts. | Lessons learned on gender achievements and outcomes are not covered although IDEV provides from time to time gender evaluations. Training of TMs and facilitation on gender should be continuously done to implement action plans |
| Chief Financial Management | Financial management is not rated at IPR and FM experts are not contacted for IPR missions. Aide memoirs and BTORR are the only way of reporting on FM issues with no detailed rating. The implementation support division may improve the situation but there is need to share findings among these divisions for problem anticipation and timely corrections.<br>PFM and fiduciary risk assessment in governance programs are well dealt with quality SESP at mid-term and CR. | Extensive resources are used to fiduciary risk assessment and implementation of mitigation plans and in building capacities | Lessons learnt are well captured at CSP/RISP levels with regard FM due to development partners consultations. Lessons from implementation of PFM programs and support of PFM agenda are well incorporated in CSPs. This is not the case for project SESP |

## Annex 7: CSP and RISP Case Studies

### A. CSPs Case Study

# CSP SESP ASSESSMENT TEMPLATE

**NARRATIVE ASSESSMENT OF CSP SESP DIMENSIONS**

**CSP Basic Data**

| | |
|---|---|
| **Country: COTE D'IVOIRE** | |
| **Countries Status: Fragile State/ADF/ADB** | |
| **CSP Program: (CSP 2013-2017 and 2013 CPPR** | **Date of Reports: October 2013** |
| **CSP – MTR and CPPR:** | **Date of Report: 2015(?)** |
| **CSP-CR: 2013-2017 and 2017 CPPR** | **Date of Report: 13 April 2018** |

**NEW CSP 2018-2022 and 2018 CPPR**

| | | | |
|---|---|---|---|
| **Report date** | 28 September 2018 | | |
| | Mission date ? | | |

## Overall CSP – CSP-MTR – CSP-CR AND CSP-E

| SESP Instruments | Narrative Assessment |
|---|---|
| CSP RESULTS FRAMEWORK MONITORING MATRIX | The 2013-2017 CSP was in support of the economy recovery of the country after the civil strife, the results framework was built on the 2012-2015 National Development Plan (NDP) and lessons learned from the 2011-2012 country strategic note. Expected long term and outputs of the 2013-2017 and potential interventions were presented results monitoring matrix showing expected outputs and outcomes at mid-term (2015) and at the end of the CSP period 2017. Long term and medium-term outputs and outcomes were quantified with reference to the Government NDP and monitorable **indicators**. **However, the results framework was reviewed based on major difficulties in achieving national development plan (NDP) development outcomes, and the new 2016-2018 NDP, with revised expected outcomes and major expected outputs at end of CSP period (2017) with reference to Bank interventions and pillars**. There was no CSP-MTR available despite reference to it in documentation. The linkage with financed programs and projects (loans, TA and grants) and indication of expected outputs and outcome achievements of the 2013-2017 CSP were also specified together with CSP contribution to the High 5s during the CSP period (2013-2017). |
| RESULTS MONITORING | Results monitoring was done with specification of each project to the High 5s |
| PROGRAMME IMPLEMENTATION EVALUATION | The 2013-2017 CSP completion report presented an assessment of the achievements based on the results monitoring matrix with few recommendations put forward the new 2018-2022 CSP. More importantly, the IDEV independent program evaluation (CSPE) submitted in December 2017, covered the period 2006-2016 and informed the new 2018-2022 CSP. CSPE was not conducted as a validation but an evaluation outputs and outcomes of Bank interventions during the 2006-2016 period. It was |

| | |
|---|---|
| | based on evaluation criteria such as relevance, effectiveness, efficiency and sustainability. The only satisfactory criterion was sustainability of results. Fragility and vulnerability risk factors were assessed and provided lessons and recommendations that were considered in the 2018-2022 CSPs (for which a management response was provided together with follow up actions to be undertaken within the new 2018-2022 CSP. |
| PORTFOLIO ALIGNMENT WITH THE BANK'S HIGH 5S | The portfolio alignment with the High 5s strategic objectives was conducted **within the** 2013-2017 CSP results framework and put forward in the new 2018-2022 CSP. |
| IMPLEMENTATION STATUS OF THE COUNTRY PORTFOLIO PERFORMANCE IMPROVEMENT PLAN (CPIP) | An assessment of achievements under the 2013 and 2017-2018 CPIP was described with clearly identified actions, monitoring indicators, authorities in charge for each generic portfolio performance issue. |
| **BANK GROUP PORTFOLIO PERFORMANCE REVIEW (CPPR)** | CPPRs were reviewed at 2013-2017 CSP-CR and the new 2018-2022 CSP with scoring of projects according to overall DOs performance ratings. Average scores of supervision reports highlighted development outcomes (DO) at 2016-2017 (2.53; 2.67 and 2.5 out of 3. 2016-2017 CPIP highlighted actions and monitoring results to improve the portfolio (quality at entry, project implementation, procurement, disbursement and audit, monitoring and evaluation). Recommendations were focused both on the portfolio performance and on the new CSP strategy alignment and coherence. CPPRs at CR analyzed the Bank as well as the Government performance in managing the portfolio with lessons and improvements for the future, mainly lessons to be incorporated into the 2018-2022 strategy. Measurable actions and monitoring results indicators were also provided in 2016-2017 CPIP and for future portfolio management**.** |
| **CSP-MTR** | Not available |
| **CSP-CR** | Based on the country context and development constraints, the 2013-2017 CSP-CR described the Bank's positioning as well as the Bank resource allocations and CSP implementation status at completion with indication of the achievements (outputs and outcomes) under each pillar and for each finance operation. The CSP-CR identified the main lessons and recommendations with orientation of the new strategic pillars for the CSP period (2018-2022) based on the IDEV 2006-2016 CSPE findings and recommendations as well as a review of 2013-2017 portfolio performance review. The 2013-2017 CSP-CR provided also an assessment of the implementation of previous 2015-2017 CPPRs, recommended measures in strengthening the quality of the portfolio, and proposed a new strategic framework for the 2018-2022 CSP with more alignment to the High 5s and to the State national development plan. |

**Overall CSP SESP Instruments Assessments**

The 2013-2017 CSP results frameworks encapsulated in CSP results monitoring matrix provided main qualitative as well as quantitative result indicators for each project under each pillar. However, the lack of clear data reporting systems with identified sources of information and data collection methodology reduced the effectiveness of the SESP in assessing and managing performances particularly in assessing country assistance strategy outcomes. The quality of the results-based framework at CSP level provided performance indicators and measurement. Also, lessons and recommendations in CSP-CR were based on strategy implementation findings as well on portfolio management. This has contributed to identify lessons that were considered into the new 2018-2022 CSP. The assessment of critical risk factors of the strategy implementation was done and taken into consideration in the new 2018-2022 CSP.
The assessments at 2013-2017 CSP-CR was made without categorization (rating) of the achievements of results as well as the overall effectiveness of Bank's strategy in achieving the stated strategic objectives or Bank's own

performance. DOs in CPPRs were generously rated for projects which distorted CSP results considering delays and landmark changes in achieving national development results due to the particular fragile situation of the country (at the end of the civil strife) and its high potential for regional integration. Cross-cutting issues such as climate change, social and environmental safeguards were analysed and operations were covered.

| **1- Performance Management** | Narrative Assessment |
|---|---|
| 1. (a) Is the CSP considered a strategy document or more of a programmatic document for country projects; or both?<br>(b) When evaluating the CSP, what weight do you give to its different components: strategy vs. portfolio?<br><br>2. To what extent are country priorities identified in the CSP reflected at the operational level?<br><br>3. To what extent do the CSP SESP outputs rely on the performance of the individual projects that it covers and to what extent have the CPPRs been taken into account in the preparation of the CSP MTR and the CR?<br><br>4. Are the CSP SESP outputs (MTR-CPPR-CR) aligned with:<br>• Main operational policy documents?<br>• High 5s?<br>• Environmental and social safeguards such as: gender, fragility, safeguards?<br>• Fiduciary & governance policy?<br><br>5. CPPRs and CPIPs are factored in the CSP and include a rating on a scale of 1 to 4 which is based on a simple average reported performance of public sector operations. How well were other non-lending activities such as ESWs, policy dialogue and analytical notes considered as part of the CSP self-evaluation?<br><br>6. Is the recent decision to separate the Country Diagnostic Note (CDN) from the CSP per se, likely to | CSP was considered both as a strategic and programming tool of Bank assistance after the social insurrection of 2011. This was designed after interim strategies between 2009 and 2009 (global country strategy for 2009-2010 and interim CSP for 2011-2012). The 2013-2017 CSP was designed following a recovery economic plan encapsulated in the national development plan 2013-2015. The selection of projects and programs were in complete alignment to the national NDP based on Bank's medium term assistance strategy for the period. More weight was given to the strategy components of Bank assistance in response to the country priorities and economic and social recovery.<br>Country priorities during the transitional period after the civil strive (2009-2012) were reflected at the operational level in 2013-2017 CSP following a review of the 2009-2010 global strategy and interim strategy 2011-2012 for which monitoring results matrix was developed.<br>CSP SESP tools were useful to demonstrate the achievement of Bank assistance strategy outputs and outcomes but have lacked precision and clarity with the absence of rigorous data reporting systems, data collection methodology and identifiable sources of information. This has reduced the effectiveness of the CSP SESP in assessing Côte d'Ivoire development constraints with the potential to regional integration. Risks to development including fiduciary risks were also assessment. The economic and social recovery as well as rehabilitation of local infrastructure were considered as a priority. Development issues such as regional integration, youth employment and governance issues were assessed and provided a basis for new program in the following 2018-2022 CSP.<br>The time horizon of the CSPs and portfolio reviews were in parallel which provided parallel actions to enhance the performance of the strategy implementation and portfolio management. Furthermore, CPPRs and CPIPs were not only focused on operational issues, but also on stragic alignment and coherence issues in light of the national development plans (2013-2015 and 2016-2018). The alignment of the portfolio management to strategy implementation was straightforward. Project performance ratings based on portfolio performance (DOs) were generous for 2.5, 2.67 and 2.53 (highly satisfactory out of a 3-point scale which has led to a distorted picture of the CSP results.<br>The 2013-2017 CSP was set both for operational and strategic levels in Côte d'Ivoire with Results frameworks and monitoring results matrices that were used to assess the achievements within the implementation of the strategy. The CSP results-based framework described the strategic objectives during the economic |

| | |
|---|---|
| facilitate the alignment with the Bank's key strategic documents and the robustness of the CSP self-evaluation process in the MTR and the CR?<br><br>7. (a) Should the SESP of the CSP (mainly the CSP-MTR and CSP-CR) cover both the CDN and the CSP document or only the latter?<br>(b) How could the CSP-MTR and CSP-CR better reflect the portfolio performance at the country level?<br>(c) What about private sector operations?<br><br>8. Are the SESP main outputs (CSP-MTR, CR, CPPRs, Annual Portfolio Reviews) adequate to ensure quality and proper delivery of the CSP?<br><br>9. Are SESP main outputs (CS¨-MTR and CSP-CR, Annual Portfolio Review) geared towards addressing:<br><br>(a) strategic issues?<br>(b) policy dialogue?<br>(c) Partnership?<br>(d) knowledge management?<br><br>10. Was IDEV validation of the project PCRs, CSPs and CSP-Es factored-in when discussing the performance of the country portfolio and of the CSP? Does external validation help improve the candour of the self-assessment? Should the CSP-CR be rated? | and social recovery period with alignment to the country urgent needs as highlighted in the NDPs.<br>CSP SESP were strengthened through the various CPPRs (2013 and 2016) and CPIP put in place (2013 and 2016) which have increased the CSP relevance. Furthermore, it was possible to attribute results to the Bank due to the strategic alignment to the country own strategy. Supervision missions were also focused on operational and development and fiduciary risks to achieving strategic outcomes and financial management.<br>Project M&E systems are crucial to report on results in CSP SESP and CPPRs which were reviewed at the 2016 CPIP. However, PCREN were not considered as official ratings in CPPRs. CSP pillars alignment to the High Five priorities was done during the 2013-2017 CSP completion report.<br>In the case of Côte d'Ivoire, non-lending activities (ESWs, policy dialogue and analytical notes) were thoroughly considered as part of the 2013-2017 CSP due to the particular situation of the country (economic and social recovery) with more budget support, and TAs. The new 2018-2022 strategy included more investment operations (with few budget support, however) and mainly non-sovereign operations with clear results projected indicators at mid-course and end of the CSP period. Private sector operations were also assessed in CSP-CR.<br>The Ivorian CSP SESP provided a strong basis for a risk analysis of strategy implementation particularly with few lessons from portfolio management, strategy implementation and IDEV independent evaluation of the 2006-2016 program. The new 2018-2022 strategy directions, risks and mitigation measures were identified.<br>Côte d'Ivoire has also benefitted from a global assessment made by donors of the main impediments to economic and social recovery taking into consideration the high potential for regional integration and social development.<br>CSP SESP should have also reported on processes to deliver sector results on social and environmental safeguards. These processes were not documented at early engagement of safeguards in projects or programs with adequate risk assessment.<br>Policy dialogue, coordination and results management were also discussed in feedback workshops around CPPR and CPIP results and specific workshops in designing the new 2018-2022 country strategy.<br>IDEV Independent program evaluation (2006-2016); although not considered as a validation exercise were thoroughly factored in the new 2018-2022 CSP design and discussion of the performance of the country portfolio. Management response was provided with clear follow up actions and inserted in the new country strategy. Validation of portfolio performance was not based on IDEV validations (PCREN). Discrepancies ere mainly discovered on the disbursement ratio (24% according to APPR but 32.7% according to the CSP-CR. CSP SESP do not apply a numerical rating system for assessing performances. IDEV independent program evaluation |

| | |
|---|---|
| | based on evaluation criteria such as relevance, effectiveness, efficiency was deemed satisfactory. The only unsatisfactory criterion was Sustainability. There are no discrepancies in Bank performance and country performance which were narratively assessed by IDEV and also in the CSP-CR. External validation would have helped in ensuring candour of the self-assessment but rating Bank and country performances are not difficult to apprehend according to country economist and CPO. |
| **OVERALL SESP PERFORMANCE MANAGEMENT DIMENSION ASSESSMENT**<br><br>In overall, CSP SESP assessed CSP implementation and portfolio performance management and offered a strong linkage to the achievement of strategic outcomes due to the strong alignment to the national development plans. The major difficulty is the lack of M&E systems both at CSP strategy (Results frameworks) and portfolio management (project results M&E frameworks) which were identified in the next CPIP 2017. De-linking CPPRs from CSP CR may have allowed more focus on strategic issues in providing more explanation of how and why results were not delivered but there is no strong position form the country team. Portfolio reviews were well conducted but the main issues remain the candor and of the portfolio assessments which rely on IPRs. If well analyzed and validated, portfolio IP and DO performances would have been a good predictor to the achievement of development objectives based on the likely achievements of project results. The sustainability of strategic outcomes is a real issue to be addressed in portfolio management. | |
| 2- **Accountability** | |
| 1- (a) Are roles and responsibilities sufficiently clear in the preparation, conduct, review, sign-off, follow-up for the various steps of the CSP SESP instruments (country team, regional directorate, country managers and chief economist)?<br>(b) Do Regional DG and Country managers see the CSP SESP outputs as a relevant accountability tool?<br><br>2. Do the SESP of the CSP outputs provide a relevant perspective on the results achieved and communicate overall performance in a credible way?<br><br>3. Is the Country, Regional and HQ Management exerting leadership over the correct implementation of the CSP SESP outputs and lines of accountability?<br><br>4. Are TMs of country program and portfolio being held accountable for timely submission and proper | Roles and responsibilities are clear in the preparation of the CSP SESP instruments with the location of the Bank HQ. The Regional Directorate at HQ plays a fundamental role in adopting an operational approach to Côte d'Ivoire CSP SESP by engaging Bank staff in adapting the strategy to the country and regional context (migration, security issues at the borders, difficult situation of neighbor countries, and high potential of Côte d'Ivoire in regional integration, etc.). The country context influenced Bank's strategy with more economic and social recovery components and governance for a stabilized economy. The regional delivery unit helps assist the regional directorate in reviewing what has worked or not in terms of country strategic achievements and results including the risk factors.<br>CSP SESP outputs did not rely on strong M&E system with data collection methodology and identification of sources of information as well as the analytical methodology of country program results but offered a credible accountability framework due to the clarity of the program assistance. The CSP team did recognize that the CSP SESP are not only to address implementation issues but also to enhance the alignment and coherence to country national development objectives encapsulated in the NDPs (social and economic recovery, reduction of social disparities, climate change, environment).<br>Portfolio reviews and CPIPs analysis provided some credibility to CSP-CR results although not validated by IDEV which submitted at an independent program evaluation of the 2006-2016 program. CPPRs and CPIPs have also included strategic issues of operational activities with the CSP and considered in the new 2018-2022 CSP. |

| | |
|---|---|
| implementation of the CSP SESP outputs?<br><br>5. To what extent do CSP SESP outputs rely on strong M&E system and offer a credible accountability framework?<br><br>6. Is attribution documented as a factor doe determining the degree of Bank's accountability?<br><br>7. To what extent TMs discuss country program and portfolio performance and results during staff performance conversations?<br><br>8. Does SESP aggregation of CSP products provide a relevant and cost-effective reporting of results through the RMF, ADER and other reporting tools (dashboard, etc..)? | The country team was able to work in a more stabilized environment after the return to HQ and feedback events inside HQ and at the country level were organized. Intensive policy dialogue within the new CSP guidelines made it clear how the CSP is the main strategic document where CPPRs should complement the assessment of strategy outcome assessment.<br>There are no real incentives for TMs candour and objectivity in IPRs and PCRs which raised few issues discussed at CODE during the CSP discussions. However, projects alignment to the strategic objectives and pillars is considered as fundamental. As for Morocco, and Tunisia, there is need to increase the budget and means for TM to enhance the quality of IPR, PCR, CPPRs. IDEV program evaluation findings were mostly reflected in the design of the new CSP but there is need for a joint work with vice presidency and DGRW front offices to enhance the credibility and explore new solutions for high quality CSP SESP.<br>The role of the implementation support division is crucial in ensuring day to day support and monitoring project portfolio performances, procurement, disbursement and also project M&E (2017 CPIP), etc… Social and environmental safeguards and climate adaptation issues were well discussed and featured in the new program 2018-2022. Involving subject matter experts at the regional hub (mainly gender, social and environmental safeguards) would support the CSP team in the CSP preparation and implementation at early stage of project/CSP design (pre-board stages). This would provide better recognition of their importance to country development outcomes, particularly Côte d'Ivoire which is facing social and regional disparities issues besides the fragile/vulnerable situation.<br>Sector division managers indicated that they have validated IPR ratings (in the workflow). However, training TMs and Support Implementation Division staff on M&E systems, data collection and analysis methodologies is essential. Also, involving country officials in rating performances at supervision, mid-term and completion help increase the validity and objectivity of ratings including at the CSP SESP levels.<br>Implementation support division is well placed to enhance the credibility of IPR and portfolio management. CSP-CR should provide room for highlighting strategy implementation issues and anticipating portfolio management issues. However, there is an overlap of responsibilities with sector divisions which may be solved with clear division of labor.<br>Few gender specialists located in regional hubs were systematically involved in CSP design and SESP. No dedicated budget is in place to cover specifically gender issues, climate change adaptation and E&S safeguards during preparation, negotiation down to completion. |
| **OVERALL SESP ACCOUNTABILITY DIMENSION ASSESSMENT**<br><br>As for Morocco and Tunisia, accountability dimension is well recognized in Côte d'Ivoire CSPs with the dual responsibility at country and regional level of TMs and subject matter experts (gender, climate change and | |

| |
|---|
| E&S safeguards, etc..). Team work for CSP preparation, follow up and completion with portfolio manager is crucial to enhance the accountability system around the CSP SESP preparation. The role of CPO is fundamental in assessing portfolio performances with the assistance of the delivery unit. An overall assessment of performance at country strategy level increases the reliability and validity of the CSP SESP. Country authorities and the Bank enforced coordination and partnership with other donors which were involved in supporting the country. The institutional framework of CSP SESP was reviewed to enhance the economic and social program with continued policy dialogue, validation and corrective actions mainly of safeguards (example regional and gender disparities, decentralized and local social development programs including infrastructure). Implementation support divisions should be also well staffed with subject matter experts covering the main cross-cutting issues as highlighted in the Ten-year strategy and the High 5s. |

| **3. Learning** | |
|---|---|
| 1. To what extent have the CSP SESP outputs for country program and portfolio reviews been used as a source of learning and knowledge management?<br><br>2. To what extent have the CPPRs and CPIPs used for the preparation of the CSP-MTR and CSP-CR?<br><br>3. Are CSP SESP outputs based on appropriate analysis and lessons learnt from previous engagements in CSP Completion, midterm reviews, country portfolio performance reviews, country portfolio improvement plan (CPIP), IDEV evaluations?<br><br>4. (a) Did feedback loop sessions and lessons learned discussions take place after CSP- MTR or CSP-CR?<br>5. (b) Have these feedback sessions led to better procedures for country program, restructuring, differentiation according to country situations (fragile context, MIC, non-lending country,… )<br><br>6. To what extent are recommendations and lessons learned described in CSP-MTR and CSP-CR actionable and | Peer reviews and independent review processes encouraged the learning from Côte d'Ivoire CSP-SESP and improved their quality together with the Economic and Sector Work that the CSP SESP covered. This has also enhanced the evaluability of the new 2018-2022 CSP pillars.<br>The involvement of regional economists in the design and implementation phases of CSPs was important but would have contributed more mainly on the potential that Côte d'Ivoire has in regional integration (as the2nd economy). The regional context (security issues at the borders as well as regional integration as one dimension among others) are important. These were correctly covered in the new 2018-2022 CSP.<br>CPPRs and CPIPs were also used for the preparation of the new CSP design, with recommendations for better selection of projects. The extension to non-sovereign operations at CSP-CR was also undertaken as part of portfolio performance assessment.<br>IDEV evaluation of the 2006-2016 Bank assistance helped in addressing strategic issues and provided recommendations for the new 2018-2022 CSP. Feedback sessions and lessons learned discussions took place after CSP-CR and at CPPRs and CPIPs. The enhancement of the alignment and coherence with the country strategy was ensured. Country portfolio performance and CSP strategic objectives and programs were aligned although CODE members felt that there is more to be done in solving portfolio management issues.<br>CSP CR lessons of experience were actionable and useful and were taken into consideration in the next strategy implementation and portfolio management. The design of the new CSP was more based on IDEV findings and recommendations with few recommendations from CSP-CR<br>Increased administrative budget for supervision should be allocated for more knowledge generation of lessons learned at country strategy, portfolio performance and E&S safeguard results. There is no feedback and capitalization workshops from CSP SESP providing success-failure stories based on assessment of achievements of country performance indicators. IDEV evaluations helped in capturing lessons and generating knowledge that enabled new strategic orientations notably of cross-cutting issues such as climate change, gender and regional disparities and other |

| | |
|---|---|
| useful to prepare the next CSP? Has any of it been used?<br><br>7. What incentives could change behaviours in terms of documenting learning evidence of proactivity for corrective action, better practices of ratings, follow up, awards for innovation, increase of value of the knowledge created, learn for failure, create opportunities for mining lessons and knowledge?<br><br>8. To what extent will the new CDN strengthen the Bank's knowledge of country situation issues and their linkages with national issues, leading to better alignment of the Bank's country operational programmes?<br><br>9. Are there concerns over ratings and disconnects between CSP-MTR and CSP-CR and CSPE that could distract from learning?<br><br>10. To what extent were leadership signals perceived that learning and knowledge management are key outcome of CSP SESP outputs? | real sectors as well as the investment climate and competitiveness and also private sector operations.<br>Lessons learnt are well captured by the new 2018-2022 CSP.<br>Since there is no rating system in CSP-SESP, IDEV ratings of evaluation criteria (relevance, coherence, efficacy, effectiveness, impact and sustainability) were useful and should have been used in CSP-SESP. However, according to the regional directorate, learning and knowledge management are key outcome of CSP SESP outputs but should be effectively recognized in policy dialogue missions as well as CPPRs and CPIPs discussions. Narrative assessment of Bank and country performance was seen as more important than a numerical rating. |
| **OVERALL SESP LEARNING DIMENSION ASSESSMENT**<br><br>Thematic reviews of IPRs and CSP SESP for better sharing and dissemination of results are needed to enhance accountability and learning. Training of TMs and facilitation on E&S and climate change in particular for country in "vulnerable" situation like Côte d'Ivoire are of utmost importance. These should be continuously done to implement social development action plans, youth employment and local infrastructure and regional integration. Training of TMs on integrating these issues should also be continuous due to the staff turnover. Simplified procedures for mainstreaming and implementation are also important. There is need to design information/reporting systems to assess performance and achievements during project, CSP cycle and help in knowledge management. Increasing the number of knowledgeable country and regional experts and building capacities on new processes for RBM including on safeguards and cross cutting issues would enhance the quality of CSP SESP.<br>Evaluation criteria such as relevance, coherence, efficacy, efficiency, impact and sustainability at CSP-SESP increase the usefulness of the self-assessment exercise and help in introducing an independent validation of | |

| |
|---|
| CSP achievements. Harmonization of evaluation criteria and ratings would increase the reliability of the assessments and learning from CSP experience. Finally, However, an assessment framework based on critical success factors or conditions of what has worked or not worked and why should have been identified to enhance the implementation of the strategy. |
| **CONCLUSIONS**<br><br>CSP-SESP in combination with CPPR-CPIP assessments offered a solid base for enhanced portfolio performance management, accountability and learning. CSP SESP tools were useful to demonstrate the achievement of Bank assistance strategy outputs and outcomes but have lacked precision and clarity with the absence of rigorous data reporting systems, data collection mand analysis methodology and M&E systems at project level. This has reduced the effectiveness of the CSP SESP in assessing and managing portfolio performance particularly in assessing country performances in managing program implementation.<br>The dual accountability dimension at country level of TMs, continued team work for CSP design, MTR and CR feedback and interaction with stable portfolio manager (CPO) have enhanced the CSP SESP quality. However, the quality of the results-based framework at CSP level needs to be enhanced with clear monitoring and measurement methods. Lessons and recommendations in CSP CR were both on strategy implementation and portfolio performance and were taken into consideration in the new 2018-2022 CSP. An assessment of critical success (or failure) factors of the strategy implementation would have enhanced the quality of its preparation.<br>The assessments at CSP-CR was made without categorization (rating) of the achievements of results of Bank assistance strategy. Ratings are considered as a panacea to enhance their quality, reliability and credibility. |

# CSP SESP ASSESSMENT TEMPLATE

## NARRATIVE ASSESSMENT OF CSP SESP DIMENSIONS

## CSP Basic Data

| | |
|---|---|
| **Country: MOROCCCO** | |
| **Countries Status: Fragile State/ADF/ADB** | |
| **CSP Program: 2012-2016** | **Date of Report: 16 March 2012** |
| **CSP – MTR 2012-2016 and CPPR 2014** | **Date of Report: 28 August 2014** |
| **CSP-CR 2012-2016 and CPPR 2016** | **Date of Report: 24 August 2016** |

## NEW CSP 2017-2021

| | | | |
|---|---|---|---|
| **Report date** | **9 March 2017** | | |
| | Mission date | | |

## Overall CSP – CSP-MTR – CSP-CR AND CSP-E

| SESP Instruments | Narrative Assessment |
|---|---|
| CSP RESULTS FRAMEWORK MONITORING MATRIX | The 2012-2016 CSP results framework is a narrative statement of expected achievements by the various on-going or newly approved programs. **There are no objectively verifiable indicators and entities' accountability for results**. This framework was reviewed at MTR with better linkage to financed programs and projects (loans, TA and grants) and clear indication of expected outputs and outcome achievements for the remaining CSP period (2014-2016). It was also reviewed at CSP completion and presented more accurately the output and outcome achievements of the various financed programs or projects approved or completed within the CSP period |
| RESULTS MONITORING | Results monitoring was done at MTR and presented the status (on-going/not attained/attained or partially attained) of the mid-term targets It states also the monitorable actions to be conducted for the remaining period of the CSP. |
| PROGRAM IMPLEMENTATION EVALUATION (CSPE) | The CSPE covers the period 2004-2014 and served as a basis of the 2017-2021 CSP. Agreed recommendations formed the management response to the CSPE. |
| PORTFOLIO ALIGNMENT WITH THE BANK'S HIGH 5S | The portfolio alignment with the High 5s strategic objectives was conducted **with no strategic actions to increase the alignment of the portfolio**. The 2017-2021 offers more indication on the alignment. |
| IMPLEMENTATION STATUS OF THE COUNTRY PORTFOLIO PERFORMANCE IMPROVEMENT PLAN (CPIP) | An assessment of achievements under the 2013 CPIP was described with clearly identified actions, monitoring indicators, authority in charge, schedule and trends up to the next period of the CPPR. However, **it does not reflect strategic issues of portfolio management improvement based on CSP output and outcome achievements**. |
| **BANK GROUP PORTFOLIO** | CPPRs were reviewed at CSP-MTR and CSP-CR with scoring of projects |

| | |
|---|---|
| **PERFORMANCE REVIEW (CPPR)** | according to SAP method (for projects approved before 2011) and to IPR method (after 2011). Average scores of supervision reports highlighted implementation performance (IP) and development outcomes (DO). However, there is **no indication to CSP strategic objectives and alignment to CSP strategy**. CPIPs were reviewed followed by specific feedback workshop on the 2014 portfolio review. However, recommendations were more focused on the portfolio performance rather than on the CSP strategy. CPPRs at MTR and CR Bank as well as the Government performances in managing the portfolio with lessons and improvements for the future **with no linkage to the country future strategy (alignment to pillars and policy dialogue).** |
| **CSP-MTR** | Based on the country context and development constraints, the CSP-MTR described the Bank's positioning as well as the allocation of resources and CSP implementation status at MTR with clear indication of the achievements (outputs and outcomes) under each pillar. The CSP-MTR identified the main lessons and recommendations with orientation on the strategic pillars for the remaining CSP period (2014-2016) and new commitments taking into account the lessons learnt with **more quantitative indicators in the results framework based** on the new lending program (2014-2016) and scenarios. The CSP-MTR provided recommendation on selection of projects that fall under the new orientation of the portfolio (more non-sovereign projects) but **with no action plan and portfolio restructuring (or re-alignment) measures and more generally on the "how to"** ensure more flexibility in project selection under the CSP strategic pillars. |
| **CSP-CR** | Based on the country context and development constraints, the CSP-CR describes the Bank's positioning as well as the Bank resource allocations and CSP implementation status at CR with clear indication of the achievements (outputs and outcomes) under each pillar. The CSP-CR identified the main lessons and recommendations with orientation on the new strategic pillars for the CSP period (2017-2017) based on the IDEV CSPE results and recommendations with indication on how IDEV findings were dealt with in the design of the new strategy. The CSP-CR provided also an assessment of the implementation of CSP-MTR recommended measures (more non-sovereign projects for example) and proposed a new strategic framework for the 2017-2021 CSP with more alignment to the High 5s. However, it does not offer an accountability framework on what has worked or not worked and why and what are the risk factors to a successful implementation of the strategy. |
| **Overall CSP SESP Instruments Assessments**:<br><br>The CSP results frameworks lacked precision and clarity and needed data reporting systems which reduced the effectiveness of the SESP in assessing and managing performances particularly in assessing country assistance strategy for which innovative cross-cutting and thematic issues should be assessed and considered in new CSPs. The quality of the results-based framework at CSP level lacked precision on performance indicators and measurement. Lessons and recommendations in CSP MTR or CR are more based on portfolio assessment with no assessment of what has worked and what has not worked, based on assessment of critical success (or failure) factors of the strategy implementation.<br>The assessments at MTR or CR is made without categorization (rating) of the achievements of results (rating) as well as the overall effectiveness of Bank's strategy in achieving the stated strategic objectives or Bank's own | |

| |
|---|
| performance. CPPRs ratings are generous for projects which distort CSP results considering delays in achieving country development results. Timeliness of CPPRs is of importance to address implementation as well as strategic issues. Cross-cutting issues such as climate change, social and environmental safeguards were analysed as the country is championing the climate change discussions at international level (COP21-22). However, although gender issues were covered in CSP SESP, CSP MTR and CR reports should have demonstrated better processes for delivering higher results on all social and environmental safeguards. The 2004-2014 IDEV CSPE (which is not a validation) covered more more than one CSP and focused the assessment on evaluation criteria such as relevance, effectiveness, efficiency, sustainability. IDEV did not rate the bank and Government performance but narratively assessed Bank's contribution/pefformance, leverage and partnership. Overall assessment was moderately satisfactory. The management response agreed on all recommendations and provided follow up actions to be incorporated in 2017-2021 CSP. |
| **CSPE 2004-2014 Date: 2 September 2016** |

| **1- Performance Management** | Narrative Assessment |
|---|---|
| 1. (a) Is the CSP considered a strategy document or more of a programmatic document for country projects; or both?<br>(b) When evaluating the CSP, what weight do you give to its different components: strategy vs. portfolio?<br><br>2. To what extent are country priorities identified in the CSP reflected at the operational level?<br><br>3. To what extent do the CSP SESP outputs rely on the performance of the individual projects that it covers and to what extent have the CPPRs been taken into account in the preparation of the CSP MTR and the CR?<br><br>4. Are the CSP SESP outputs (MTR-CPPR-CR) aligned with:<br>• Main operational policy documents?<br>• High 5s?<br>• Environmental and social safeguards such as: gender, fragility, safeguards?<br>• Fiduciary & governance policy?<br><br>5. CPPRs and CPIPs are factored in the CSP and include a rating on a scale of 1 to 4 which is based on a simple average reported performance of public sector operations. How well were other non-lending activities such as ESWs, policy | CSP was considered both as a strategy document and programming tool of Bank assistance with the selection of projects or programs that support Bank assistance strategic objectives. More weight was also given to strategic components with likely attribution/contribution of Bank financed projects and programs.<br>Country priorities were reflected at the operational level at CSP MTR and CR following a review of CPIP and thorough project selection process.<br>CSP SESP tools were useful to demonstrate the achievement of Bank assistance strategy outputs and outcomes but have lacked precision and clarity with the absence of rigorous data reporting systems. This has reduced the effectiveness of the CSP SESP in assessing and managing portfolio performance particularly in assessing Morocco development results. More development issues (for example, youth employment or anti-corruption issues) should have been more thoroughly discussed and assessed in new CSP.<br>The time horizon of CSPs and portfolio reviews are different which provide a distorted picture of performance between the portfolio and the country strategy. Furthermore, CPPRs are more focused on operational issues and less on strategy implementation. Project performance ratings (based on the SAP and IPR methods) were generous for IP and DO which have distorted the CSP results. The IDEV CSPE provided moderately satisfactory performance to effectiveness, efficiency and sustainability but satisfactory on relevance and alignment of CSP strategy to Bank and country priorities.<br>CSP were not based on clear theory of change (TOC) both on operational and strategic levels in Morocco.<br>CSP Results frameworks were not used for M&E purposes. The quality of the results-based framework at project level (logframe) does reflect the strategic objectives and alignment with country development objectives and Bank policies and sector strategies. Also, results frameworks at project level |

| | |
|---|---|
| dialogue and analytical notes considered as part of the CSP self-evaluation?<br><br>6. Is the recent decision to separate the Country Diagnostic Note (CDN) from the CSP per se, likely to facilitate the alignment with the Bank's key strategic documents and the robustness of the CSP self-evaluation process in the MTR and the CR?<br><br>7. (a) Should the SESP of the CSP (mainly the CSP-MTR and CSP-CR) cover both the CDN and the CSP document or only the latter?<br>(b) How could the CSP-MTR and CSP-CR better reflect the portfolio performance at the country level?<br>(c) What about private sector operations?<br><br>8. Are the SESP main outputs (CSP-MTR, CR, CPPRs, Annual Portfolio Reviews) adequate to ensure quality and proper delivery of the CSP?<br><br>9. Are SESP main outputs (CS¨-MTR and CSP-CR, Annual Portfolio Review) geared towards addressing:<br><br>(e) strategic issues?<br>(f) policy dialogue?<br>(g) Partnership?<br>(h) knowledge management?<br><br>10. Was IDEV validation of the project PCRs, CSPs and CSP-Es factored-in when discussing the performance of the country portfolio and of the CSP? Does external validation help improve the candour of the self-assessment? Should the CSP-CR be rated? | lacked smart indicators mainly on the social sector (poverty assessment, youth employment, gender disparities, etc..). CPPRs and CSP SESP are strengthened through the various CPIP put in place (2013-2014 and 2015) which have increased the CSP relevance. However, it was difficult to attribute results to the Bank due to the lack of valid information on country outcomes. Supervision missions were not focused on operational and development risks to achieving strategic outcomes. The dashboard is mostly oriented towards procedures (disbursement, procurement, time to effectiveness, etc..) but not focused on risk factors of implementation and outcome achievement at country level.<br>Project M&E systems are crucial to report on results in CSP SESP or CPPRs. But, un-timely achievement of development objectives reduced the validity of IPR DO ratings as highlighted by IDEV CSPE. CPPRs should be synchronized with CSP timeline and PCREN should be fed to staff and considered for official ratings in CPPRs. CSP pillars alignment to the High Five priorities is an issue (7 sectors) but a good selection of sectors where Bank's comparative advantage is demonstrated should guide Bank assistance strategy.<br>In the case of Morocco CSP, non-lending activities (ESWs, policy dialogue and analytical notes) were thoroughly considered as part of the CSP self-evaluation and also for the design of the new 2017-2021 strategy. Private sector operations were not thoroughly discussed in CSP-MTR and CR although both independent evaluation and MTR reviews highlighted the need to increase the volume of lending to private sector.<br>The Moroccan CSP SESP provide a strong basis for an analysis of strategy implementation risks. However, the "discrepancy" in assessing strategy results with CPPRs at mid-term review or completion is due to the fact that CSP preparation teams work in silos. CSP team involved in CSP SESP preparation should have normally dealt with discrepancies of portfolio ratings and strategy results frameworks.<br>Morocco has benefitted from a thorough assessment of gender profile which helped in designing gender policy and strategy that have been translated in budgetary resource allocations and reporting system of line ministries in charge of implementation of gender action plans. However, achievement of outcome indicators at project level were not aggregated at CSP results frameworks.<br>Cross-cutting issues such as climate change, gender, social and environmental safeguards are not well analysed at inception nor at SESP during implementation phases of project and country/regional strategies. Risks to adaptation to climate change not reflected or embedded in assessing performances or National Determined Contributions (country commitments). CSP SESP should have also reported on processes to deliver sector results on social and |

| | |
|---|---|
| | environmental safeguards. These processes should have been documented at early engagement of safeguards at projects or programs with adequate risk assessment.<br>Finally, Morocco CSP SESP have dealt with strategic issues, policy dialogue, partnerships and knowledge management in feedback workshops around CPPR results and specific workshops in designing the new 2017-2021 country strategy. Independent evaluation of CSP findings were factored in new CSP design and discussion of the performance of the country portfolio for which management response was provided with clear follow up actions. However, validation of projects DOs were not based on IDEV validations (PCREN) nor increased or improved the candour of the self-assessment. Bank and country performance were not independently rated by IDEV but performance was narratively assessed (attribution/contribution to country results). CSP SESP do not apply a numerical rating system for assessing performances and in addressing the pitfalls and redirecting the strategic objectives specially at mid-course or in designing the new country strategy. |
| **OVERALL SESP PERFORMANCE MANAGEMENT DIMENSION ASSESSMENT**<br><br>In overall, CSP SESP assessed portfolio performance management but lacked a strong linkage to the achievement of strategic outcomes. The major difficulty is the lack of M&E systems both at CSP results frameworks and projects results frameworks (logframes) which lack alignment, coherence and convergence. De-linking CPPRs from CSP MTR or CR may have allowed more focus on strategic issues in providing more explanation of how and why results were not delivered. However, if well conducted, portfolio reviews would have been considered as a good tool to predict the achievement of development objectives based on achievements of projects/programs IP and DOs. | |
| 2- **Accountability** | |
| 1- (a) Are roles and responsibilities sufficiently clear in the preparation, conduct, review, sign-off, follow-up for the various steps of the CSP SESP instruments (country team, regional directorate, country managers and chief economist)?<br><br>2- Do Regional DG and Country managers see the CSP SESP outputs as a relevant accountability tool?<br><br>3- Do the SESP of the CSP outputs provide a relevant perspective on the results achieved and communicate overall performance in a credible way?<br><br>4- Is the Country, Regional and HQ Management exerting leadership over | Roles and responsibilities are clear in the preparation of the CSP SESP instruments with the creation of regional directorate and country offices. Regional Directorates play a fundamental role in adopting an operational approach to CSP SESP by engaging ADB staff in adapting their strategies to the country (or the regional context). The regional delivery unit helps assist the regional directorate in reviewing what has worked or not in terms of country strategic achievements and results including the risk factors. The SESP at project level is restricted to operational issues and only mid-term reviews provide an opportunity to reflect on the adequacy of Bank's response to country/region strategy.<br>CSP SESP outputs did not rely on strong M&E system nor offered a credible accountability framework. Also, the CSP team di recognize that the CSP SESP are not only for addressing implementation issues but should have promoted a strong alignment and coherence to country development objectives.<br>Portfolio reviews in Morocco CPPR analysis provided some |

| | |
|---|---|
| the correct implementation of the CSP SESP outputs and lines of accountability? | credibility to CSP MTR and CR results. However, CPIPs should have also included strategic issues of misalignment and lack of coherence between operational activities with the CSP strategy for enhanced credibility. |
| 5- Are TMs of country program and portfolio being held accountable for timely submission and proper implementation of the CSP SESP outputs? | The country team concept is no longer working and information sharing events are now organized inside the regional directorates at CPPR assessments and CSP policy dialogue. New CSP guidelines did not make it clear which from CSP is the main strategic document or the subsidiary and how they complement each other. |
| 6- To what extent do CSP SESP outputs rely on strong M&E system and offer a credible accountability framework? | There are no real incentives for TMs candour and objectivity. However, projects alignment to the strategic objectives and pillars is fundamental. There is need to increase the budget and means for TM to enhance the quality of IPR, PCR, CPPRs. |
| 7- Is attribution documented as a factor doe determining the degree of Bank's accountability? | IDEV CPEs evaluation reports findings are reflected in new design of CSPs but there is need for a joint work across regions to enhance the credibility and explore new solutions for high quality CSP SESP. |
| 8- To what extent TMs discuss country program and portfolio performance and results during staff performance conversations?<br><br>9- Does SESP aggregation of CSP products provide a relevant and cost-effective reporting of results through the RMF, ADER and other reporting tools (dashboard, etc..)? | The role of the implementation support division is crucial in ensuring day to day support and monitoring project portfolio performances, procurement, disbursement etc…<br>Only 2 paragraphs were allocated to safeguards and climate adaptation issues which may devote an annex to CSP reports and complete reporting on achievements in CSP-SESP. Climate change, social and environmental safeguards should be used as support team to project and CSP preparation and implementation and involved in early stage of project/CSP design (pre-board stages) for better recognition of their importance to country development outcomes.<br>Country authorities were involved in CSP SESP but did not report on E&S safeguards with adequate reporting guidelines on E&S management/mitigation plan.<br>Few sector division managers validate IPR ratings (although in the workflow). Training TMs on M&E is essential. Also, involving country officials in rating performances at supervision, mid-term and completion help increase the validity and objectivity of ratings including at the CSP SESP levels but need to be independently validated.<br>Implementation support division is well placed to enhance the credibility of IPR, CSP-MTR in providing comments on the implementation results and anticipating portfolio implementation issues. However, it may appear that there is an overlap of responsibilities with sector divisions.<br>Few gender specialists are located in regions. Gender policy guidelines have to be reviewed as no clear plan of action is in place to report on results in SESP. No dedicated budget is in place to cover specifically gender issues during preparation, negotiation down to completion. |
| **OVERALL SESP ACCOUNTABILITY DIMENSION ASSESSMENT**<br><br>Accountability dimension is well recognized with the dual responsibility at country and regional level of TMs | |

| |
|---|
| and subject matter experts (gender, climate change and E&S safeguards, etc..). However, team work from CSP design, MTR and CR and interaction with portfolio managers (CPOs) are crucial to enhance the accountability system in CSP SESP. However, the lack of staff in sector divisions does not allow for continued strategic and operational dialogue and for increased supervision and just-in time actions to address project or program strategic implementation issues. The role of CPOs is still fundamental in assessing portfolio performances which with the assistance of the delivery unit allows for overall assessment of performance at country strategy level. Also, a partnership agreement with a performance management platform that translates both commitments (country authorities and the Bank) to enforce coordination and in assessing attribution and/or contribution is crucial to enhance mutual accountability. The institutional framework of CSP SESP should be reviewed to enhance gender and other subject matter experts' involvement in CSP assessments, policy dialogue, validations and corrective actions mainly of safeguards (example compensation in resettlements). Implementation support divisions at the regional hubs should be well staffed with subject matter experts covering the main cross-cutting issues highlighted in the ten year strategy and the High 5s. if well conducted, CSP SESP aggregation of CSP outputs and outcomes products may have provided a relevant and cost-effective reporting of results through the ADER and other reporting tools (dashboard, etc..) which is not actually the case. |

| **3. Learning** | |
|---|---|
| 1. To what extent have the CSP SESP outputs for country program and portfolio reviews been used as a source of learning and knowledge management? | Peer reviews and independent review processes encouraged the learning from Morocco CSP-SESP and improved their quality together with the Economic and Sector Work that the CSP SESP covered. This has also enhanced the evaluability of CSP pillars. |
| 2. To what extent have the CPPRs and CPIPs used for the preparation of the CSP-MTR and CSP-CR? | The involvement of regional economists in the design and implementation phases of CSPs were minimal but would have contributed more mainly on the knowledge of the regional context as regional integration is one dimension among others. This was corrected by the new CSP guidelines. |
| 3. Are CSP SESP outputs based on appropriate analysis and lessons learnt from previous engagements in CSP Completion, midterm reviews, country portfolio performance reviews, country portfolio improvement plan (CPIP), IDEV evaluations? | CSP-SESP helped in assessing operational issues but with little learning at CSP strategic level. As such, CPPRs and CPIPs were not totally used for the preparation of the new CSP design, although there was a recommendation for better selection of projects which was translated into concrete action. The extension to non-sovereign operations at CSP-MTR and CSP-CR was not emanating from portfolio performance assessment. |
| 4. (a) Did feedback loop sessions and lessons learned discussions take place after CSP- MTR or CSP-CR?<br>(b) Have these feedback sessions led to better procedures for country program, restructuring, differentiation according to country situations (fragile context, MIC, non-lending country,… ) | IDEV evaluation of the 2004-2014 strategies helped in addressing strategic issues and helped in preparing the new 2017-2021 CSP. However, the exercise was more on the accountability side rather than on the learning although critical risk factors were assessed. Also, feedback sessions and lessons learned discussions took place after CSP- MTR or CSP-CR and at CPPRs but CPIPs did not include the enhancement of alignment and coherence between country portfolio performance and CSP strategic objectives and programs. |
| 5. To what extent are recommendations and lessons learned described in CSP-MTR and CSP-CR actionable and useful to prepare the next CSP? Has any of it been used? | CSP-MTR and CR lessons of experience were actionable and useful but mostly on strategy implementation. The design of the new CSP was more based on IDEV findings and recommendations.<br>Increased administrative budget for supervision should be allocated for more knowledge generation of lessons learned |

| | |
|---|---|
| 6. What incentives could change behaviours in terms of documenting learning evidence of proactivity for corrective action, better practices of ratings, follow up, awards for innovation, increase of value of the knowledge created, learn for failure, create opportunities for mining lessons and knowledge?<br><br>7. To what extent will the new CDN strengthen the Bank's knowledge of country situation issues and their linkages with national issues, leading to better alignment of the Bank's country operational programmes?<br><br>8. Are there concerns over ratings and disconnects between CSP-MTR and CSP-CR and CSPE that could distract from learning?<br><br>9. To what extent were leadership signals perceived that learning and knowledge management are key outcome of CSP SESP outputs? | at country strategy, portfolio performance and E&S safeguard results.<br>Lessons learned on gender achievements and outcomes are not covered although IDEV provides from time to time gender evaluations. There is no feedback loop from CSP SESP that provide success-failure stories and results based on assessment on achievements in reaching development indicators. IDEV evaluations helped in capturing lessons and generating knowledge that enabled new strategic orientations notably of cross-cutting issues such as climate change, gender and regional disparities.<br>Lessons learnt are well captured at CSP levels with regard Financial Management due to development partners consultations mainly around the PFM agenda. Lessons from implementation of PFM programs and support of PFM agenda are well incorporated in CSPs in Morocco. This is not the case for project IPR and PCRs.<br>Since there is no rating system in CSP-SESP, IDEV ratings of evaluation criteria (relevance, coherence, efficacy, effectiveness, impact and sustainability) were useful but should be well digested by CSP team before using in CSP-SESP.<br>According to the regional directorate, learning and knowledge management are key outcome of CSP SESP outputs but should be effectively recognized by organizing capitalization workshops on CSP SESP together with policy dialogue, CPPRs and CPIPs. |
| **OVERALL SESP LEARNING DIMENSION ASSESSMENT**<br><br>Need for thematic reviews of IPRs and SESP for better sharing and dissemination of results. Training of TMs and facilitation on gender should be continuously done to implement action plans. Training of TMs on integrating these issues should be continuous because of the staff turnover and procedures simplified for better mainstreaming. There is need to design information/reporting systems that assess performance and achievements during project, CSP cycle and help in knowledge management. There is need to increase the number of knowledgeable country as well as regional experts and build capacities on new processes for RBM including on safeguards and cross cutting issues.<br>Lessons learned should be institutionalized during project life cycle and documented for capitalization of lessons learned at project and country levels. In fact, including evaluation criteria such as relevance, coherence, efficacy, efficiency, impact and sustainability at CSP-SESP increases the usefulness of the self-assessment exercise and helps in introducing an independent validation of results achievements. Harmonization of evaluation criteria and ratings would increase the reliability of the assessments and learning from CSP experience. | |
| **CONCLUSIONS**<br><br>CSP-SESP in combination with CPPR assessments do not offer a solid base for enhanced performance management, accountability and learning. CSP SESP tools were useful to demonstrate the achievement of Bank assistance strategy outputs and outcomes but have lacked precision and clarity with the absence of rigorous data reporting systems and M&E systems at project level. This has reduced the effectiveness of the CSP SESP in assessing and managing portfolio performance particularly in assessing Morocco development results. | |

Although, there is a dual accountability dimension at country level of TMs, the team work for CSP design, MTR and CR feedback and interaction with portfolio managers (CPOs) should be enhanced in CSP SESP.
The quality of the results-based framework at CSP level lacked precisions on performance indicators and measurement. Lessons and recommendations in CSP MTR or CR are more based on portfolio assessment with no assessment of what has worked and what has not worked, based on assessment of critical success (or failure) factors of the strategy implementation.
The assessments at MTR or CR was made without categorization (rating) of the achievements of results of Bank assistance strategy. CPPRs ratings were generous for projects which distorted CSP results considering delays in achieving country development results.

# CSP SESP ASSESSMENT TEMPLATE

## NARRATIVE ASSESSMENT OF CSP SESP DIMENSIONS

### CSP Basic Data

| | |
|---|---|
| **Country: TUNISIA** | |
| **Countries Status: Fragile State/ADF/ADB** | |
| **CSP Program: (i) Interim CSP 2014-2015[56]<br>(ii) Interim CSP 2014-2015 extended to 2016** | **Date of Reports: 12 February 2014<br>22 December 2015** |
| **CSP – MTR and CPPR: NA** | **Date of Report: NA** |
| **CSP-CR: 2014-2016 and CPPR 2016** | **Date of Report: 3 October 2016** |

### NEW CSP 2017-2021

| | | | |
|---|---|---|---|
| **Report date** | 30 May 2017 | | |
| | Mission date | | |

## Overall CSP – CSP-MTR – CSP-CR AND CSP-E

| SESP Instruments | Narrative Assessment |
|---|---|
| CSP RESULTS FRAMEWORK MONITORING MATRIX | The 2014-2015 Interim CSP results framework was built in a logical framework specifying the country main development constraints, expected long term outcomes, end of the I-CSP period (2015) expected outcomes, outputs of the 2014-2015 I-CSP and potential interventions and technical assistance operations to be implemented under each pillar during the I-CSP period (2014-2015). Few long term and I-CSP outcomes were quantified (mainly for infrastructure projects) but with **no clear objectively verifiable indicators. However, the results framework (intervention logic) was revised in 2015 after specifying the major difficulties in achieving development outcomes, new long term expected outcomes (till 2018), revised expected outcomes and major expected outputs at end of I-CSP period (2016) with more quantifiable indicators and responsible entities**. There was no I-CSP-MTR due to the short timeframe of the I-CSP (2014-2015 extended to 2016). The linkage to financed programs and projects (loans, TA and grants) and indication of expected outputs and outcome achievements of the 2014-2016 I-CSP were also specified. The results monitoring matrix was reviewed at I-CSP completion (2016) and presented more accurate output and outcome achievements of financed programs/projects approved during the I-CSP period (2014-2016). |
| RESULTS MONITORING | NA |
| PROGRAMME IMPLEMENTATION EVALUATION | The I-CSP completion report presented an assessment of the achievements based on the results monitoring matric with few recommendations put forward the new 2017-2021 CSP. More importantly, the IDEV independent program evaluation (CSPE) submitted September 2016, covered the period |

[56] Due to the difficult situation of the country, the Bank designed interim CSPs (I-CSPs) for the following periods: 2012-2013; 2014 and 2015 extended to 2016 for which no CSP-MTR was prepared. The ICSP-CR was prepared in 2016 and served the basis for a full-fledged CSP 2017-2021

| | |
|---|---|
| | 2004-2015 and informed the new 2017-2021 CSP. Management response to the CSPE has provided follow up actions that were taken into consideration in the new CSP. |
| PORTFOLIO ALIGNMENT WITH THE BANK'S HIGH 5S | The portfolio alignment with the High 5s strategic objectives was conducted **within the** 2017-2021 CSP results framework. |
| IMPLEMENTATION STATUS OF THE COUNTRY PORTFOLIO PERFORMANCE IMPROVEMENT PLAN (CPIP) | An assessment of achievements under the 2013 and 2015 CPIP was described with clearly identified actions, monitoring indicators, authority in charge for each generic portfolio performance issue. However, **it did not reflect on strategic issues of the portfolio improvement plan based on I-CSP results.** A CPIP was also developed in 2016 and inserted in 2017-2021 CSP. |
| **BANK GROUP PORTFOLIO PERFORMANCE REVIEW (CPPR)** | CPPRs were reviewed at I-CSP-CR and the new 2017-2021 CSP with scoring of projects according to IPR performance ratings. Average scores of supervision reports highlighted implementation performance (IP) and development outcomes (DO) at 2015 and 2016 CPPRs. However, there is **no indication to CSP strategic objectives and alignment to CSP strategy**. Portfolio performance indicators were reviewed for 2013 and 2015 CPPRs. Feedback workshop was organized to discuss CPIP and CPPR findings in 2015. However, recommendations were more focused on the portfolio performance rather than on the I-CSP strategy alignment and coherence. CPPRs at MTR and CR analyzed Bank as well as the Government performance in managing the portfolio with lessons and improvements for the future, however, **with no linkage to the country future strategy (alignment to pillars and coherence with the country development dimensions).** It should be noted that CODE made observations on the 2014-2016 I-CSP-CR and CPPR findings for which Management provided a response as an annex to the 2017-2021 CSP |
| **CSP-MTR** | NA |
| **CSP-CR** | Based on the country context and development constraints, the 2014-2016 I-CSP-CR described the Bank's positioning as well as the Bank resource allocations and I-CSP implementation status at completion with indication of the achievements (outputs and outcomes) under each pillar and for each finance operation. The CSP-CR identified the main lessons and recommendations with orientation of the new strategic pillars for the CSP period (2017-2021) based on the IDEV CSPE findings and recommendations as well as a review of 2014-2016 portfolio performance review. The 2014-2016 CSP-CR provided also an assessment of the implementation of previous 2014-2015 I-CSP, recommended measures in strengthening the quality of the portfolio, and proposed a new strategic framework for the 2017-2021 CSP with more alignment to the High 5s. However, it did not offer an accountability framework on what has worked or not worked and why although risk factors and mitigation measures were identified to enhance the implementation of the strategy. |
| **Overall CSP SESP Instruments Assessments**<br><br>The I-CSP results frameworks encapsulated in I-CSP logical frameworks provided main qualitative as well as quantitative result indicators for each project under each pillar. However, the lack of clear data reporting systems with identified sources of information and data collection methodology reduced the effectiveness of the SESP in assessing and managing performances particularly in assessing country assistance strategy outcomes. The quality of the results-based framework at I-CSP level provided performance indicators and measurement. However, lessons and recommendations in I-CSP-CR were more based on portfolio assessment. | |

The assessment of critical success (or failure) factors of the strategy implementation was done and inserted in the new 2017-2021 CSP.
The assessments at I-CSP-CR was made without categorization (rating) of the achievements of results as well as the overall effectiveness of Bank's strategy in achieving the stated strategic objectives or Bank's own performance. IDEV independent CSPE assessed the performance of the I-strategies 2004-2015 based on evaluation criteria such relevance, effectiveness, efficiency, sustainability and contribution to development outcomes. It was not a validation exercise. However, in its Management response, Management agreed on all recommendations and presented follow up actions to be considered in the 2017-2021 CSP. CSPE did not rate the Bank and country performance (only narrative assessment) based on contribution, coordination and partnerships. CPPRs ratings were generous for projects which distorted I-CSP results considering delays in achieving country development results due to the particular political situation of the country (political transition). Cross-cutting issues such as climate change, social and environmental safeguards were analysed and operations were inserted to reduce regional disparities and promote climate change adaptation. This was rated as moderately unsatisfactory by IDEV CSPE

| **1- Performance Management** | Narrative Assessment |
|---|---|
| 1) (a) Is the CSP considered a strategy document or more of a programmatic document for country projects; or both?<br>(b) When evaluating the CSP, what weight do you give to its different components: strategy vs. portfolio?<br><br>2) To what extent are country priorities identified in the CSP reflected at the operational level?<br><br>3) To what extent do the CSP SESP outputs rely on the performance of the individual projects that it covers and to what extent have the CPPRs been taken into account in the preparation of the CSP MTR and the CR?<br><br>4) Are the CSP SESP outputs (MTR-CPPR-CR) aligned with:<br>a. Main operational policy documents?<br>b. High 5s?<br>c. Environmental and social safeguards such as: gender, fragility, safeguards?<br>d. Fiduciary & governance policy?<br><br>5) CPPRs and CPIPs are factored | I-CSP was considered more as programming tool of Bank assistance after the social insurrection of 2011. An I-CSP was designed for the 2012-2013 period followed but a second one for the period 2014-2015 which was extended to 2016 due to the delays in implementing the program. The selection of projects or programs that support Bank short term assistance strategic objectives was done with more weight given to operational components of Bank financed projects and programs.<br>Country priorities during the political transitional period of 3 years (2012-2014) were reflected at the operational level in 2014-2016 I-CSP following a review of 2014-2015 I-CSP results and 2013-2015 CPPRs and CPIPs.<br>I-CSP SESP tools were useful to demonstrate the achievement of Bank assistance strategy outputs and outcomes but have lacked precision and clarity with the absence of rigorous data reporting systems, data collection methodology and identifiable sources of information. This has reduced the effectiveness of the I-CSP SESP in assessing Tunisia development constraints and risks to development outcomes. The stabilization of the economic and social situation was considered as a priority. Development issues such as regional decentralization, geographic and gender disparities, youth employment and/or anti-corruption) were assessed in the following I-CSP 2014-2015 extended to 2016 and few operations were part of the Bank assistance program.<br>The time horizon of the I-CSPs and portfolio reviews were in parallel which provided parallel actions to enhance the performance of the interim strategy implementation and the portfolio management. However, CPPRs and CPIPs were more focused on operational issues. The alignment of the portfolio management to the interim strategy implementation was |

| | |
|---|---|
| in the CSP and include a rating on a scale of 1 to 4 which is based on a simple average reported performance of public sector operations. How well were other non-lending activities such as ESWs, policy dialogue and analytical notes considered as part of the CSP self-evaluation?<br><br>6) Is the recent decision to separate the Country Diagnostic Note (CDN) from the CSP per se, likely to facilitate the alignment with the Bank's key strategic documents and the robustness of the CSP self-evaluation process in the MTR and the CR?<br><br>7) (a) Should the SESP of the CSP (mainly the CSP-MTR and CSP-CR) cover both the CDN and the CSP document or only the latter?<br>(b) How could the CSP-MTR and CSP-CR better reflect the portfolio performance at the country level?<br><br>8) What about private sector operations?<br><br>9) Are the SESP main outputs (CSP-MTR, CR, CPPRs, Annual Portfolio Reviews) adequate to ensure quality and proper delivery of the CSP?<br><br>10) Are SESP main outputs (CS¨-MTR and CSP-CR, Annual Portfolio Review) geared towards addressing:<br>a) strategic issues?<br>b) policy dialogue?<br>c) Partnership?<br>d) knowledge management?<br><br>11) Was IDEV validation of the project PCRs, CSPs and CSP-Es factored-in when discussing the performance of the country portfolio and of the CSP? Does external validation help improve the candour of the self-assessment? Should the CSP-CR be rated? | straightforward due to the short period of the I-CSP. Project performance ratings based on IPRs were generous for IP (highly satisfactory: 3.53) and DO (satisfactory 3.22) which led to a distorted picture of the I-CSP results. Although the IDEV CSPE independent evaluation was not a validation, evaluation criteria were moderately satisfactory and moderately unsatisfactory for the sustainability criterion due to the lack of durable results of environmental and social components of the program. On CPPRs, issues were raised by CODE on the quality of the portfolio.<br>The 2 I-CSPs were set both on operational and strategic levels in Tunisia. However, I-CSP Results frameworks were not used for M&E purposes. The I-CSP results-based framework described the strategic operational objectives during the transitional period with alignment to the country urgent needs.<br>I-CSP SESP were strengthened through the various CPPRs (2013 and 2015) and CPIP put in place (2013 and 2015) which have increased the I-CSP relevance. However, it was possible to attribute results to the Bank due to the short-term program designed for the country. Supervision missions were also focused on operational and development risks to achieving strategic outcomes. The dashboard comprised risk factors of implementation and outcome achievement at country level for the new 2017-2021.<br>Project M&E systems are crucial to report on results in I-CSP SESP and CPPRs. CPPRs were synchronized with I-CSP timeline but PCREN were not considered as official ratings in CPPRs. I-CSP pillars alignment to the High Five priorities was done in the new 2017-2021 CSP but needs a good selection of sectors where Bank's comparative advantage should be demonstrated.<br>In the case of Tunisia, non-lending activities (ESWs, policy dialogue and analytical notes) were thoroughly considered as part of the I-CSP due to the particular situation of the country (political transitional) with more budget support, and TAs. The new 2017-2021 strategy included more investment operations (with few budget support, however) with clear results projected indicators for 2019 and 2021 (year 2015 was the baseline).<br>Private sector operations were not thoroughly discussed in I-CSP CR but were considered in the new strategy.<br>The Tunisian CSP SESP provided a strong basis for a risk analysis of strategy implementation particularly with few lessons for the new 2017-2021 strategy and mitigation measures were identified. I-CSP team involved in I-CSP SESP preparation worked in harmony on portfolio management and strategy results frameworks, due to the short timeline of the interim strategy and the stability of the team. The I-CSP and portfolio were strictly followed due to the political risk and security issue in the country.<br>Tunisia has benefitted from a thorough assessment of the main impediment to economic and social stabilization particularly the local and social development (removing local and gender |

| | |
|---|---|
| | disparities for priority remote regions) as well as a series of sector notes on industry, agriculture, social which were developed for the new 2017-2021 CSP. An attempt was also made to aggregate project outcomes to the CSP results frameworks.<br>I-CSP SESP should have also reported on processes to deliver sector results on social and environmental safeguards which led to unsatisfactory evaluation made by IDEV in the CSPE. These processes were not documented at early engagement of safeguards in projects or programs with adequate risk assessment.<br>Finally, Tunisia I-CSP SESP have dealt with strategic issues, policy dialogue, partnerships and knowledge management in feedback workshops around CPPR and CPIP results and specific workshops in designing the new 2017-2021 country strategy. Independent evaluation of CSP findings were factored in new CSP design and discussion of the performance of the country portfolio for which management response was provided with clear follow up actions. However, validation of project DOs was not based on IDEV validations (PCREN) nor improved the candour of the self-assessment. CSP SESP do not apply a numerical rating system for assessing performances which would have played a fundamental role in addressing the pitfalls and redirecting the strategic objectives specially at the extension of the 2014-2015 I-CSP. |
| **OVERALL SESP PERFORMANCE MANAGEMENT DIMENSION ASSESSMENT**<br><br>In overall, I-CSP SESP assessed portfolio performance management and offered a strong linkage to the achievement of strategic outcomes due to the short period of the Bank interim. The major difficulty is the lack of M&E systems both at CSP strategy (every 2 years, the last one extended by one year). Results frameworks and projects results frameworks were alignment and strictly followed due to the political situation and security issues. De-linking CPPRs from I-CSP CR may have allowed more focus on strategic issues in providing more explanation of how and why results were not delivered. However, portfolio reviews were well conducted but the main issues remain the candor and of the portfolio assessments which rely on IPRs. If well analyzed and validated, portfolio IP and DO performances would be good indicators to predict the achievement of development objectives based on their likely achievements. | |
| **Accountability** | |
| 1- Are roles and responsibilities sufficiently clear in the preparation, conduct, review, sign-off, follow-up for the various steps of the CSP SESP instruments (country team, regional directorate, country managers and chief economist)?<br>2- Do Regional DG and Country managers see the CSP SESP outputs as a relevant accountability tool? | Roles and responsibilities are clear in the preparation of the I-CSP SESP instruments with the creation of the regional directorate in Tunis. The Regional Directorate plays a fundamental role in adopting an operational approach to Tunisia I-CSP SESP by engaging ADB staff in adapting the interim strategy to the country and regional context (migration, security issues at the borders, difficult situation of neighbour countries, etc.). The country context influenced Bank's strategy with more social components and economic stabilization program. The regional delivery unit helps assist the regional directorate in reviewing what has worked or not in terms of country strategic achievements and results including the risk factors. The SESP at |

| | |
|---|---|
| 3- Do the SESP of the CSP outputs provide a relevant perspective on the results achieved and communicate overall performance in a credible way?<br><br>4- Is the Country, Regional and HQ Management exerting leadership over the correct implementation of the CSP SESP outputs and lines of accountability?<br><br>5- Are TMs of country program and portfolio being held accountable for timely submission and proper implementation of the CSP SESP outputs?<br><br>6- To what extent do CSP SESP outputs rely on strong M&E system and offer a credible accountability framework?<br><br>7- Is attribution documented as a factor doe determining the degree of Bank's accountability?<br><br>8- To what extent TMs discuss country program and portfolio performance and results during staff performance conversations?<br><br>9- Does SESP aggregation of CSP products provide a relevant and cost-effective reporting of results through the RMF, ADER and other reporting tools (dashboard, etc..)? | project level is however restricted to operational issues and by only real-time reviews (for the specific case of Tunisia). This has provided an opportunity to reflect on the adequacy of Bank's response to country/region situation.<br>I-CSP SESP outputs did not rely on strong M&E system with data collection methodology and identification of sources of information as well as the analytical methodology of country program results but offered a a credible accountability framework due to the short period of the I-CSP (2 years each, the last extended by one year). Also, the CSP team did recognize that the CSP SESP are not only for addressing implementation issues but have promoted a strong alignment and coherence to country strategic priorities (reducing social disparities, climate change, environment, infrastructure and economic stabilization support).<br>Portfolio reviews in Tunisia in CPPRs and CPIPs analysis provided some credibility to I-CSP-CR results although not validated by IDEV which submitted at the same time the independent evaluation of the 2004-2015 Bank assistance. Although constraints and impediments to high portfolio performance were assessed, CPPRs and CPIPs should have included strategic issues of operational activities with the I-CSP and also the new 2017-2021 CSP strategy for enhanced credibility.<br>The country team was stable in working together and sharing information including organization of feedback events inside the regional directorate and in the country including during the intensive policy dialogue. New CSP guidelines did not make it clear which from CSP or the CPPR is the main strategic document to be considered for validation and assessment.<br>There are no real incentives for TMs candour and objectivity in IPRs which raised few issues discussed at CODE in a special session dedicated to the 2014-2016 I-CSP-CR and 2016 CPPR. However, projects alignment to the strategic objectives and pillars is fundamental. As for Morocco, there is need to increase the budget and means for TM to enhance the quality of IPR, PCR, CPPRs. IDEV CSPEs evaluation reports findings are mostly reflected in new design of the new CSP but there is need for a joint work with vice presidency and DGRN front offices to enhance the credibility and explore new solutions for high quality CSP SESP.<br>The role of the implementation support division is crucial in ensuring day to day support and monitoring project portfolio performances, procurement, disbursement etc…specially for Tunisia which shows some difficulties in complying with Bank rules and procedures due to the high turnover of the actual administration, newly in office. Social and environmental safeguards and climate adaptation issues were well discussed and featured in the new program 2017-2021. Complete reporting on achievements in I-CSP-SESP should have been done to enrich the discussion of the climate change adaptation, social and environmental safeguards using the available subject matter experts at the regional hub to support the CSP team in the CSP |

| | |
|---|---|
| | preparation and implementation at early stage of project/CSP design (pre-board stages). This would provide better recognition of their importance to country development outcomes, particularly Tunisia which is facing environmental and social issues besides the political and security situation.<br>Country authorities were involved in I-CSP SESP and did report on E&S safeguards with adequate reporting and ownership. Guidelines on E&S management/mitigation plan were also followed.<br>Sector division managers indicated that they have validated IPR ratings (in the workflow). However, training TMs and Support Implementation Division staff on M&E systems, data collection and analysis methodologies is essential. Also, involving country officials in rating performances at supervision, mid-term and completion help increase the validity and objectivity of ratings including at the CSP SESP levels.<br>Implementation support division is well placed to enhance the credibility of IPR, CSP-CR in providing comments on the implementation results and anticipating portfolio implementation issues. However, there is an overlap of responsibilities with sector divisions which may be solved with clear division of labor.<br>Few gender specialists are located in regional hubs. No dedicated budget is in place to cover specifically gender issues, climate change adaptation and E&S safeguards during preparation, negotiation down to completion. |
| **OVERALL SESP ACCOUNTABILITY DIMENSION ASSESSMENT**<br>As for Morocco, accountability dimension is well recognized in Tunisia I-CSPs with the dual responsibility at country and regional level of TMs and subject matter experts (gender, climate change and E&S safeguards, etc..). Team work for I-CSP design, and I-CSP-CR and interaction with portfolio manager (CPO) were crucial to enhance the accountability system in Tunisia I-CSP SESP. The role of CPO is fundamental in assessing portfolio performances which, with the assistance of the delivery unit. An overall assessment of performance at country strategy level increases the reliability and validity of the I-CSP SESP. In the case of Tunisia, a partnership agreement with performance management platform was set at the country level and helped to a quick response to the country social and economic needs and priorities during the political transition period. Country authorities and the Bank enforced coordination and partnership with other donors heavily involved in supporting the country and helped assess the interim results and ensured continuity of aid assistance. This has also enhanced mutual accountability. The institutional framework of I-CSP SESP was reviewed to enhance the economic and social program with continued policy dialogue, validation and corrective actions mainly of safeguards (example regional and gender disparities, decentralized and local social development programs including infrastructure). Implementation support divisions at the regional hubs should be also well staffed with subject matter experts covering the main cross-cutting issues as highlighted in the Ten-year strategy and the High 5s. I-CSP SESP aggregation of I-CSP outputs and outcomes products provided a relevant and cost-effective reporting of results in Tunisia which served in the new 2017-2021 CSP formulation. | |
| **3. Learning** | |
| 1. To what extent have the CSP SESP outputs for country program and portfolio reviews been used as a source of learning and knowledge | According to the regional directorate, learning and knowledge management are key outcome of CSP SESP outputs but should be effectively recognized in policy dialogue missions as well as CPPRs and CPIPs discussions. |

| | |
|---|---|
| management?<br><br>2. To what extent have the CPPRs and CPIPs used for the preparation of the CSP-MTR and CSP-CR?<br><br>3. Are CSP SESP outputs based on appropriate analysis and lessons learnt from previous engagements in CSP Completion, midterm reviews, country portfolio performance reviews, country portfolio improvement plan (CPIP), IDEV evaluations?<br><br>4. (a) Did feedback loop sessions and lessons learned discussions take place after CSP- MTR or CSP-CR? (b) Have these feedback sessions led to better procedures for country program, restructuring, differentiation according to country situations (fragile context, MIC, non-lending country,… )<br><br>5. To what extent are recommendations and lessons learned described in CSP-MTR and CSP-CR actionable and useful to prepare the next CSP? Has any of it been used?<br><br>6. What incentives could change behaviours in terms of documenting learning evidence of proactivity for corrective action, better practices of ratings, follow up, awards for innovation, increase of value of the knowledge created, learn for failure, create opportunities for mining lessons and knowledge?<br><br>7. To what extent will the new CDN strengthen the Bank's knowledge of country situation issues and their linkages with national issues, leading to better alignment of the Bank's country operational programmes? | Peer reviews and independent review processes encouraged the learning from Tunisia I-CSP-SESP and improved their quality together with the Economic and Sector Work that the I-CSP SESP covered. This has also enhanced the evaluability of the new 2017-2021 CSP pillars.<br>The involvement of regional economists in the design and implementation phases of I-CSPs was minimal but would have contributed more mainly on the knowledge of the regional context (security issues at the borders as well as regional integration as one dimension among others). This was covered in the I-CSP SESP and correctly dealt in the new CSP.<br>I-CSP-SESP helped in assessing operational issues but with little learning at CSP strategic level. As such, CPPRs and CPIPs were not totally used for the preparation of the new CSP design, although there was a recommendation for better selection of projects which was translated into concrete action. The extension to non-sovereign operations at CSP-MTR and CSP-CR was not emanating from portfolio performance assessment.<br>IDEV evaluation of the 2004-2015 Bank assistance helped in addressing strategic issues and provided recommendations for the new 2017-2021 CSP. It has provided an assessment of critical success factors and conditions for high CSP performance. Feedback sessions and lessons learned discussions took place after I-CSP-CR and at CPPRs and CPIPs as well as on the CSPE. The enhancement of the alignment and coherence with the I-strategy was easily ensured due to the short timeline of the I-CSPs and I-CSP SESP. Country portfolio performance and CSP strategic objectives and programs are aligned although CODE members felt that there is more to be done in addressing portfolio performance issues.<br>I-CSP CR lessons of experience were actionable and useful but mostly on strategy implementation and portfolio management. The design of the new CSP was more based on IDEV findings and recommendations with few recommendations from I-CSP SESP<br>Increased administrative budget for supervision should be allocated for more knowledge generation of lessons learned at country strategy, portfolio performance and E&S safeguard results.<br>There is no feedback and capitalization workshops from I-CSP SESP providing success-failure stories based on assessment of achievements of country performance indicators. IDEV evaluations helped in capturing lessons and generating knowledge that enabled new strategic orientations notably of cross-cutting issues such as climate change, gender and regional disparities and other real sectors as well as the investment climate and competitiveness.<br>Lessons learnt are well captured by the new CSP with regard Financial Management issues due to development partners continued consultations. Lessons from implementation of PFM program and support of PFM agenda are well incorporated in |

| | |
|---|---|
| 8. Are there concerns over ratings and disconnects between CSP-MTR and CSP-CR and CSPE that could distract from learning?<br><br>9. To what extent were leadership signals perceived that learning and knowledge management are key outcome of CSP SESP outputs? | Tunisia CSP.<br>Since there is no rating system in I-CSP-SESP, IDEV ratings of evaluation criteria (relevance, coherence, efficacy, effectiveness, impact and sustainability) were useful although Bank and country performances were narratively assessed. This should have been used in I-CSP-SESP to ease future validation exercise. |
| **OVERALL SESP LEARNING DIMENSION ASSESSMENT**<br>Thematic reviews of IPRs and CSP SESP for better sharing and dissemination of results are needed to enhance accountability and learning. Training of TMs and facilitation on E&S and climate change in particular for country in transition like as Tunisia are of utmost importance. These should be continuously done to implement local development action plans, youth employment and local infrastructure disparities. Training of TMs on integrating these issues should also be continuous due to the staff turnover. Simplified procedures for mainstreaming and implementation are also important. There is need to design information/reporting systems to assess performance and achievements during project, CSP cycle and help in knowledge management. Increasing the number of knowledgeable country and regional experts and building capacities on new processes for Results based management including on safeguards and cross cutting issues would enhance the quality of CSP SESP..<br>Evaluation criteria such as relevance, coherence, efficacy, efficiency, impact and sustainability at CSP-SESP increase the usefulness of the self-assessment exercise and help in introducing an independent validation of CSP achievements. Harmonization of evaluation criteria and ratings would increase the reliability of the assessments and learning from CSP experience. | |
| **CONCLUSIONS**<br>I-CSP-SESP in combination with CPPR-CPIP assessments for Tunisia offered a solid base for enhanced performance management, accountability and learning due to the short period of program implementation and the close scrutiny exercised by i-CSP staff. I-CSP SESP tools were useful to demonstrate the achievement of Bank interim assistance strategy outputs and outcomes but have lacked precision and clarity with the absence of rigorous data reporting systems, data collection mand analysis methodology and M&E systems at project level. This has reduced the effectiveness of the I-CSP SESP in assessing and managing portfolio performance particularly in assessing country performances in managing interim program implementation issues.<br>The dual accountability dimension at country level of TMs, continued team work for I-CSP design, MTR and CR feedback and interaction with stable portfolio manager (CPO) have enhanced the I-CSP SESP quality.<br>The quality of the results-based framework at I-CSP level lacked precision on performance indicators and measurement particularly for the 2014-2015 I-CSP and improved substantially during the extension to 2016. Lessons and recommendations in I-CSP CR are more based on portfolio performance with no assessment of what has worked and what has not worked. An assessment of critical success (or failure) factors of the strategy implementation would have enhanced their quality. This was thoroughly done in designing the new 2017-2021 strategy.<br>The assessments at I-CSP-CR was made without categorization (rating) of the achievements of results of Bank interim assistance strategy. CPPRs ratings (based on IPRs) were generous for projects which have distorted I-CSP results considering delays in achieving country performance and development results. | |

## B. RISPs Case Study

# RISP SESP ASSESSMENT TEMPLATE

### NARRATIVE ASSESSMENT OF RISP SESP DIMENSIONS

## RISP Basic Data

| | |
|---|---|
| *Region:* CENTRAL AFRICA | |
| **Countries Covered:** | **Eleven (11) countries covered**[57]: Angola, Burundi*, Cameroon, Chad, Central African Republic, Congo, Equatorial Guinea, Democratic Republic of Congo, Gabon, Rwanda*, Sao Tome and Principe, and Rwanda. |
| **Countries Status: Fragile State/ADF/ADB** | ADF-only countries: 6 (Burundi**, Central African Republic**, Chad**, Democratic Republic of Congo**, Rwanda and Sao Tome & Principe)<br>ADF/ADB blend countries: 1 (Cameroon)<br>ADB Countries: 4 (Angola, Congo, Equatorial Guinea (REG), and Gabon) |
| **RISP Program 2011-2015 extended to 2017** | **Report Date: 1st April 2011 extension to 2017** |
| **RISP – MTR 2011-2015 and 2015 RPPR** | **Report Date: 31 December 2013** |
| **RISP-CR Report data: 2011-2017 and 2017 RPPR** | **Report Date: 15 May 2018**[58] |

*Covered by the Easter Africa RISP

**Fragile and conflict affected States

## NEW RISP (2019-2025)

| | |
|---|---|
| **Date of report:** | **March 2019** |
| | Mission date? |

## Overall RISP – RISP-MTR – RISP-CR AND RISP-E

| SESP Instruments | Narrative Assessment |
|---|---|
| RISP RESULTS FRAMEWORK MONITORING MATRIX | The results framework monitoring matrix described the ECCAS and CEMAC regional economic programs of high priority needs for physical and economic integration of countries in the region as defined by ECCAS Member States. The two pillars: (i) regional infrastructure development; and (ii) regional institutional and human capacity building – were translated into regional operations identified in the ECCAS/CEMAC and Central African Forest Commission sectoral programs (COMIFAC) – in particular the Central African Consensual Transport Master Plan PDCT-AC and Central African Power Pool (CAPP). |

[57] Within the African Development Bank Group, the operational area of the Regional Development and Business Delivery Office for Central Africa (RDGC) covers seven countries: Cameroon, Chad, Congo, Gabon, Central African Republic, Equatorial Guinea, Democratic Republic of Congo, and Equatorial Guinea. This area does not cover all the member countries of the Economic Community of Central African States (ECCAS). The Regional Economic Community of Central Africa. ECCAS comprises 11 countries: Angola, Burundi, Cameroon, Chad, Central African Republic, Congo, Equatorial Guinea, Democratic Republic of Congo, Gabon, Rwanda, Sao Tome and Principe, and Rwanda. As regards the Bank's operational portfolio, Burundi and Rwanda are covered by the Regional Office for East Africa (RDGE) and are included in RISP 2018-2023 for that region, while Angola and São Tomé and Principe are covered by the Regional Office for Southern Africa (RDGS) and included in its next RISP 2019-2024.

[58] Presentation at OPSCOM on 15 May 2018. The full report not made available to the team.

| | |
|---|---|
| | Constraints that hampered regional integration and opportunities were analyzed and key outputs and outcomes at MTR and at the end of the RISP period (2011-2015) were projected with indicative on-going and new financing program under the 2011-2015 period. It was based on analysis of the economic situation and focused on regional integration leverages (macro-economic management, energy, transport, ICT,.). The matrix included quantitative targets of outputs and outcomes for the main strategic objectives for institutional and human development capacity building. However, the capacity building results lacked measurable indicators. The selection of projects followed the 2-step approach[59] as defined by ADF-12 deputies (February 2011). The matrix was updated and revised at MTR with incorporation of new regional programs appraised during the period 2011-2013. At the end of the period, and extension of 2 years (on June 2017) (?), RISP-CR findings were presented to OPSCOM and highlighted several issues with regard the 2011-2015 RISP results and presented the portfolio performance as well the orientations for the next strategy (2018-2025). In the same vein, an independent evaluation of the Central Africa regional integration strategy (2008-2016) highlighted the relevance of the strategy in addressing infrastructure and capacity building obstacles but was unrealistic and not well known by development partners with no visible influence on the portfolio implementation and effectiveness of Multinational operations in Central Africa. |
| RESULTS MONITORING | The summary outcome monitoring at RISP-CR was presented for the main regional projects (corridors, power interconnections, fiver optic interconnections, trade facilitation and institutional capacity building of RECs), including for environmental management and preservation programs (Congo Basin ecosystem and conservation support programme, the Lake Chad Basin sustainable development project). However, new monitoring results matrix was not prepared based on updates of the CR assessment after the 2-year extension. The main weakness is around delays in approval and implementation of multinational operations mainly for capacity building operations. However, the RISP was flexible enough to accommodate new interventions following the political crisis and vulnerability in the region showing some flexibility and ability to align with regional priority needs. |
| PROGRAMME IMPLEMENTATION EVALUATION | The programme implementation self-assessment was undertaken at MTR and at completion (2017). The implementation of the programme at MTR highlighted several issues which necessitated the extension of the RISP period (2011-2015) to 2017. The program did not include specific actions or activities with regard fragile states that belong to the region. |

59 The five-year Regional Integration Strategy Paper should contain an indicative pipeline of operations recommended for Bank Group support with consideration of (i) Operations‟ alignment with the Bank's corporate priorities and the Regional Integration Strategy; (ii) Operations' alignment with the priorities of the African Union, NEPAD, RECs and Regional Member Countries; (iii) Operations' impact on regional integration; and Ownership of participating countries and entities. The prioritization framework (Scorecard) includes indicators related to the participating countries (CPIA, portfolio performances of regional operations from APPR, countries trade facilitation policy commitments to regional integration) and regional operations expected development outcomes and contribution to regional integration and quality at entry and readiness (existence of MOU agreement between beneficiary countries, evidence of support from relevant RECs, and collaboration/co-financing with development partners). Source: Regional Operations Selection and Prioritization Framework, February 2011.

| | |
|---|---|
| | Institutional capacity building (social and multi-sector) operations' implementation was difficult due to the weak capacity of the RECs. |
| PORTFOLIO ALIGNMENT WITH THE BANK'S HIGH 5S | The portfolio alignment with the high 5s was made under the preparation of the RISP-CR. The alignment with the country CSPs was made to ensure consistency with the RISP pillars at CR. The RISP-CR highlighted also the new strategy pillars and expected outcomes and the alignment to the Bank and continental priorities in line with the new Bank-wide Strategic Framework (RISF) (approved in March 2018), and UN-SDGs |
| IMPLEMENTATION STATUS OF THE COUNTRY PORTFOLIO PERFORMANCE IMPROVEMENT PLAN | The implementation status of the RPIP was undertaken at MTR (2015) and at completion (with a revised RPIP for 2018). The portfolio performance issues, corrective measures to be undertaken, responsible entities and indicative monitorable indicators (qualitative) were provided in the 2018 RPIP. |
| BANK GROUP PORTFOLIO PERFORMANCE REVIEW | RISP portfolio reviews were undertaken at MTR (2014) and at Completion (2018) with RPIP prepared for 2018. Portfolio performance indicators were identified and reviewed during the RISP period with stagnation if not regression in most of regional and multi-country operations with a disbursement rate decreasing from 41% in 2014 to 38% at RISP CR (2018) contradicting with an overall portfolio performance of 3 (out of 4 in a 4-point scale). The Bank Group regional portfolio reflected the fragile economic development challenges in the Central African region, especially the recent efforts of the Bank Group in addressing the political crisis in Burundi, Central African Republic, DRC and other fragile states. Portfolio implementation challenges in Central Africa include: (i) long timeframes in meeting conditions precedent to first disbursement; (ii) delay in the provision of counterpart funds; (iii) delays in the procurement process and in the performance of works and services; (iv) delayed submission of audit reports and poor quality of some reports; (v) non-compliance with the frequency of Bank supervision and launch missions; (vi) level of insecurity in some areas of the region; (vii) absence of key staff at the start of projects; and (viii) cumbersome internal Bank and ECCAS procedures. |
| RISP MTR | The implementation status of the program was assessed showing a limited progress of approvals. Delays in implementing the new operations programmed in the RISP have occurred while out of the 6 operations scheduled in 2011 and 2012, two were approved by the Board, representing a 33% implementation rate. studies. |
| RISP CR | The RISP-CR was made as a presentation of the regional context and assessment of performance in terms of results (outputs and outcomes). The CR presentation highlighted several issues with substantial delays at start-up showing few tangible results to report. The RISP-CR discussed very well the regional context and assessed the results achievements, constraints, challenges and opportunities, and discussed the pillars adjustments to be made under the new RISP (2019-2025). |
| **Overall RISP SESP Instruments Assessments**: The RISP SESP reports were analytical and attempted to assess candidly and proactively the results achieved during the RISP period while highlighting the challenges and bottlenecks to regional integration strategy and programs. These are mostly program oriented although regional integration, vulnerability and fragility issues were at stake for the central African region. The | |

monitoring results matrix was revised at MTR and CR without adjustment of the pillars. The RISP MTR was candid to highlight the poor performance in achieving outputs and outcomes of the program. This was further at RISP CR which offered a candid assessment of results and identified options for the new strategic orientation of the 2019-2025 RISP while raising an important issue related to the ownership, lack of commitments of member states, multiplicity of multinational operations and mechanisms that were not conducive to policy convergence and real economic integration (low trade facilitation and harmonization of customs duties, etc. RISP CR came after the BDEV evaluation of the RISP and drew lessons for the new RISP 2019-2025. Discrepancies between RISP SESP in terms of results and portfolio management, were highlighted in the RISP-CR describing the low achievements and reasons for low performance of the program in terms of disbursement but in contradiction with the overall performance of the portfolio (3 out 4 point-scale). The 2019-2025 followed the new guidelines and offered a better shape of the regional integration strategy underpinned by a regional diagnostic and several studies of the challenges, constraints and bottlenecks. The new RISP-CA sought to "improve economic diversification and structural transformation through the improvement of intra-regional trade in Central Africa" through two pillars, namely: (i) Strengthening regional infrastructure (energy, transport and ICT); and (ii) Supporting reforms for intra-regional trade development and build the institutional capacity of RECs.

| **1. Performance Management** | Narrative Assessment |
|---|---|
| 1. (a) Is the RISP considered a strategy document or more of a programmatic document for regional projects; or both?<br>(b) When evaluating the RISP, what weight do you give to its different components: strategy vs. portfolio?<br><br>2. RISPs can cover several CSPs over more than one cycle, and a portfolio of projects with various durations. What are the main criteria used in assessing the performance of the RISP through its main outputs: MTR, RPPR, CPPR and the CR?<br><br>3. Are the RISP SESP outputs (MTR-RPPR-CR) aligned with:<br>• Main operational policy documents?<br>• High 5s?<br>• Environmental and social safeguards such as: gender, fragility, climate change?<br>• Fiduciary & governance policy?<br><br>4. To what extent have alignment and/or complementarity between RISPs and CSPs been dealt with as some projects faced challenges of ownership and cross-country | The RISP-CA rested on through pillars, namely (i) enhancing connectivity and interconnections through physical infrastructure, (ii) building institutional and human capacity for effective implementation of the regional integration agenda, and (iii) environmental management and ecosystems presentations. This strategy was aligned with the CCAS and CEMAC Vision 2025, and the outcome of consultations with regional stakeholders.<br><br>The RISP SESP provided the basis for the analysis of the regional context but was rather a programmatic document despite the analysis of regional integration and bottlenecks with a high number of fragile states for which the Bank conducted fragility and vulnerability assessments. The RISP-MTR and CR made an attempt to align the regional integration strategic objectives with the country strategies, continental priorities and the UN SDGs. The prioritization framework (Scorecard) did include indicators related to the participating countries (CPIA, portfolio performances of regional operations from APPR, countries trade facilitation policy commitments to regional integration) and quality at entry and readiness but revealed the weak capacity of the RECS as a serious impediment to the smooth implementation of the soft components of the regional programs.<br><br>Security and fragility were the main issues to further regional integration, adding the security in some part of the region.<br><br>With the recent guidelines on regional diagnostic note (RDN) preceding the RISP preparation has, according to regional coordinator, facilitated the understanding of the strategic components of the regional integration strategy and assisted in the adjustment of the main strategy pillars. However, RISP SESP should have reflected on program implementation issues as |

| | |
|---|---|
| coordination since regional operations are typically implemented at national levels?<br><br>5. RPPRs and CPPRs covered by the RISP include a rating on a scale of 1 to 4, which is based on a simple average of reported performance of public sector operations. How well were other non-lending activities considered as part of the RISP self-evaluation?<br><br>6. Is the recent decision to separate the regional diagnostic note (RDN) from the RISP per se, likely to facilitate the alignment with the Bank's key strategic documents and the robustness of the RISP self-evaluation process in the MTR and the CR?<br><br>7. (a) Should SESP outputs of the RISP cover both the RDN and the RISP document or only the latter?<br>(b) How could the RISP-MTR and RISP-CR better reflect the performance at the regional level and its country components?<br><br>8. Are the SESP main outputs (RISP-MTR, CR, Annual Portfolio Reviews) geared towards addressing:<br>(a) strategic issues,<br>(b) policy dialogue,<br>(c) knowledge management?<br><br>9. Was IDEV validation of the project's PCRs, CSPs and RISP-Es factored-in when discussing the performance of the regional portfolio and of the RISP? Does external validation help improve the candour of the self-assessment? Should the RISP-CR be rated? | highlighted by the independent evaluation 2008-2016 which highlighted the need for a proportionate and appropriate approach; more realistic operational indicative program; Improvement of political dialogue and leadership; dissemination and feedback; private sector support; and continued support for REC capacity building.<br><br>The 2019-2025 RPIP included a series of remedial actions to improve portfolio management performance, quality at entry, procurement, disbursement and M&E with monitorable indicators, responsible entities and timeframe including clear output and outcome indicators.<br><br>RISP CR offered a candid assessment of results and identified options for the new strategic orientation of the RISP-CA while raising an important issue related to the lack of ownership, state members' commitment, and multiplicity of mechanisms related to multi-country projects and program.<br><br>The independent evaluation highlighted the satisfactory relevance of the program but also the unsatisfactory performance of effectiveness, efficiency and highly sustainability unsatisfactory performance.<br><br>RISP were not based on clear theory of change (TOC) both on operational and strategic level in the CA countries or region. RISP Results framework did not inform on regional integration progress nor on macroeconomic convergence. This was highlighted in RISP-CR and independent evaluation.<br><br>RECs and implementation agencies and beneficiaries' involved in RISP SESP preparation did happen. However, capacity development of RECs (including in statistics and impacts assessments) was necessary to enhance the quality of design and follow up of strategic development objectives of the region. This was not significantly assessed.<br><br>Portfolio performance indicators were identified and reviewed during the RISP period but with stagnation and even decreasing performance in multi-country operations contradicting with the overall performance of the Portfolio rated satisfactory (3 out of 4-point scale). Disbursement rate has decreased from 41% in 2014 to 38% in 2018.<br><br>De-linking RPPRs and RPIPs from RISP MTR or CR was not obvious as the programming projects with focus on operational issues was predominantly of strategic importance. |

| |
|---|
| **Performance Management Overall Assessment**<br><br>Portfolio implementation challenges in Central Africa include: (i) delays in the implementation of loan agreements, (ii) delay in the fulfilment of conditions precedent to first disbursement, (iii) insufficient knowledge of the Bank's rules and procedures, particularly with regard to procurement and disbursements, and (iv) weak implementation capacity. The RISP SESP provided a basis for the analysis of regional program implementation risks related to performance and results. However, the RISP was rather a programmatic document despite the analysis of regional context and challenges with a high number of fragile states.<br><br>New guidelines did not make it clear which from CSP or RISP is the main strategic document or the subsidiary and how they complement each other. The independent evaluation assessed the relevance, effectiveness, efficiency and sustainability and concluded that the strategy was based on an unrealistic theory of change. In spite of an improved physical connectivity, poor enforcement of existing trade agreements as well as weak governance continued to hamper intraregional trade. There was a low ownership of the strategy and its multinational operations with no visible influence of the program on the design, portfolio implementation and effectiveness of Multinational operations in Central Africa which is the least integrated region of the continent. This region received the least resources for its integration and was the only one where the Bank's assistance to regional integration has declined during the implementation period of the strategy. Effectiveness and efficiency were unsatisfactory while sustainability of results was highly unsatisfactory suggesting the Bank's failure to engage in high level policy dialogue, strong donor coordination with a considerable financial engagement in the region, despite the complexity of institutional and organizational architecture with ineffective implementation arrangements up to the complexity of the situation.<br><br>The RISP-MTR and CR did not assess the extent to which the RISP implementation achieved its ultimate goals of advancing the regional integration agenda and facilitating regional solutions. It did not inform on regional integration progress nor on macroeconomic convergence.<br><br>De-linking CPPRs from CSP/RIPS MTR or CR, in the case of the RISP-CA is not obvious due to the focus on the multinational operations rather on the regional policy convergence and macroeconomic management. Program design and quality at entry issues were the major impediment to quality design of the regional integration. RISPs SESP in region with high number of fragile states lack valid and reliable data on implementation and results due to conflicts and unavailability of data. |

| 2. **Accountability** | |
|---|---|
| 1. (a) Are roles and responsibilities sufficiently clear in the preparation, conduct, review, sign-off, follow-up for the various steps of the RISP SESP outputs (country team, regional team, country managers, DG and chief economist)?<br>(b) Do Regional DG and Country managers see the RISP SESP outputs as a relevant accountability tool?<br>2. Do SESP of the RISP outputs provide a relevant perspective on | Regional Directorate plays a fundamental role in adopting an operational approach to RISPs and RISP SESP by engaging ADB staff in adapting country specific strategies to the regional context and enhancing Bank's agility and ability to include urgent security and vulnerability crisis in the region.<br><br>The SESP at regional multinational projects were restricted to operational issues. However, the mid-term review provided an opportunity to reflect on the adequacy of Bank's response to the regional integration issues, challenges and bottlenecks but did not carefully adjust the program to focus more on strategic issues. The extension to 2017 was due to the low progress in program implementation and weak capacity of the implementing agencies despite the increased number of consultations. These |

| | |
|---|---|
| the results achieved and communicate overall performance in a credible way?<br><br>3. Is Regional and HQ Management exerting leadership over the correct implementation of the RISP SESP outputs and lines of accountability?<br><br>4. Are TMs of regional programs being held accountable for timely submission and proper implementation of the RISP SESP outputs?<br><br>5. To what extent do SESP outputs of regional program: (a) rely on strong M&E systems? (b) offer a credible accountability framework?<br><br>6. To what extent do managers discuss with TMs regional program performance and results during staff performance conversations?<br><br>7. Does SESP aggregation of RISP products provide a relevant and cost-effective reporting of results through the RMF, ADER and other reporting tools (dashboard, etc..)? | were considered as insufficient to enhance ownership and states commitments to regional integration. Accountability for results was judged by the independent program evaluation (2008-2016) as unsatisfactory with insufficient performances of the Bank, the RECs, Governments and other actors… In fact, although countries have ratified key international conventions and provided political backing to regional organizations, effective political commitment and financial contributions of countries to regional organizations have been sporadic with weak leadership and low capacity of regional organizations to oversee regional programs.<br><br>RISP SESP are considered for both accountability and for knowledge base of it entails in promoting regional integration. This did not happen in Central Africa RISP and was highlighted at mid-course (2013-2014). The over-optimistic programming was that it did not sufficiently take into account the level of preparation of investment projects or factors that have negatively affected the preparation of these projects such as: (i) the poor quality or absence of feasibility studies within the prescribed timeframe; (ii) differences between the parties as regards project structuring from the institutional standpoint; (iii) weak State commitment to the preparation of certain operations; and lastly, (iv) insufficient or poor quality dialogue which sometimes led to substantial changes in project design that required additional resources.<br><br>Collaboration with RECs proved to be difficult without strong capacity building in terms of policy convergence, harmonization of macro-economic policies and a good understanding of the political economy of how regional integration is built. The independent program evaluation highlighted the low disbursement performance, the high commitment at risks and the low disbursement rate.<br><br>The RISP-MTR was candid in highlighting the main reasons for insufficient performance in regional integration objectives in terms of free movement of people and goods, harmonization of customs standards and the business environment which were very limited due essentially to the weak commitment of States. In contrast, progress was made in RECs' capacity to implement projects although the challenges remain significant. Some progress was made, however, in the area of environmental management and preservation of ecosystems. Reasons of unsatisfactory performance of the Bank were: inefficient supervision, highly centralized decision-making process due to the relocation of the Bank in Tunis, lack of real time technical assistance to RECs and insufficient familiarization of implementing agencies on Bank's rules and procedures.<br><br>The RISP MTR suggested a gradual approach through which reforms and regional integration process would be implemented |

| | |
|---|---|
| | in tandem and in synergy with the consolidation of national economies. From the operational standpoint, States and the RECs needed to be more involved in project monitoring while the Bank needed to improve the quality of operations at start-up with a more proactive approach to problem-solving. Furthermore, more focus should be given to aspects of fragility in the design and implementation of operations.<br><br>The Results monitoring Framework did not cascade to allow meaningful IPR reports on regional programs and results matrices were mostly not adapted to country and region context with M&E reporting systems. Results frameworks lacked smart indicators mainly on the social sector (poverty assessment, vulnerability, gender disparities, etc..).<br><br>The regional directorate (and also the RDVP) were exerting leadership for the implementation of the regional program both on hard and less on soft components due to the relocation of the bank in Tunis, centralization of the decision-making process posed several issues for timely corrective actions.<br><br>Discrepancies between RISP SESP and the independent evaluation conducted by BDEV highlighted the need for more candor and rigorous assessments of effectiveness and efficiency and effective M&E systems within the implementing agencies and RECs and suggested high commitments of member states and high effectiveness and efficiency of RECs in implementing the regional programs. Lessons and recommendations in both RISP SESP and independent evaluation would offer an opportunity for the Bank to enhance the ownership and the level of commitments in leading donor coordination and partnerships and encourage high level consultations with private sector investors, civil society organizations and regional institutions and organizations.<br><br>Countries and RECs need to be involved in SESP in reporting on regional difficulties and bottlenecks or challenges as well as on opportunities for enhanced policy convergence, macroeconomic management, and the political economy of regional integration with security and political vulnerability issues.<br><br>As for the WA-RISP, the RISP-CA's institutional framework has to be reviewed to enhance reporting on results and integration into the ADER and RMF with regional experts' involvement in RISP assessments, policy dialogue, validations and corrective actions mainly of safeguards (example compensation in resettlements) and gender. The implementation support divisions should be well staffed with all subject matter experts such as gender, climate change, E&S experts, etc… |

| | |
|---|---|
| | |

**Accountability Overall Assessment**

Regional Directorate plays a fundamental role in adopting an operational approach to RISPs and RISP SESP by engaging ADB staff in adapting country specific strategies to the regional context and promoted their ability and agility to face the regional political and security crisis. The regional delivery unit would help assist the regional directorate in reviewing what has worked or not in terms of regional strategic implementation, achievements and results including the risk factors around them.

Accountability is based on management validation and accuracy of performance data in RISP SESP. However, integrated convergence of national strategies encapsulated in CSPs did not allow for strong accountability of RISP SESP tools. Collaboration with RECs proved to be difficult without strong capacity building in terms of policy convergence, harmonization of macro-economic policies and a good understanding of the political economy of how regional integration is built.

The specificity of the Central African region is the low ownership of the regional strategy objectives and its multinational operations with a limited visible influence of the operations on the design, portfolio implementation and effectiveness of Multinational operations in Central Africa. Central Africa region is considered as the least integrated region of the continent with the least financial flow for its integration and only one where the Bank's assistance to regional integration has declined during the implementation period of the strategy.

Discrepancies between RISP SESP and independent validation undertaken by BDEV highlighted the need for more candor and rigorous assessments of effectiveness and efficiency and effective M&E systems within the implementing agencies and RECs.

The independent evaluation of the RISP-CA (2008-2016) recommended that the Bank, due to the high number of fragile states and landlocked countries, suffering from political instability and violence, which some of them were severely hit in their public finances by the sharp decline of oil, should tailor its approach and consider the preparation of an indicative operational program for regional integration with a tailored project pipeline to better reflect the specificities of the region. Better understanding of what the Bank wants to achieve and foster an increased ownership of the regional actions needs to be undertaken with improved policy dialogue and leadership by using its leadership position in terms of financing and knowledge sharing of regional activities to foster the coordination of donor interventions at the regional level making countries and donors work better together.

| **3. Learning** | |
|---|---|
| 1. To what extent have the RISP SESP outputs for regional program been used as a source of learning and knowledge management to inform | The involvement of regional economists and regional coordinators in the design and implementation phases of RISPs contributes mainly to enhance the knowledge sharing on what regional integration entails. This was corrected by the new |

| | |
|---|---|
| the next RISP or other regional operations/strategies<br><br>2. To what extent has the RPPR or CPPRs been used for the preparation of the RISP- MTR and the RISP-CR?<br><br>3. (a) Did feedback loop sessions and lessons learned discussions take place after RISP- MTR or RISP CR? (b) Have these feedback sessions led to better procedures for regional programs, restructuring, differentiation according to regional situations (fragile context, MIC, non-lending)<br><br>4. To what extent are recommendations and lessons learned described in RISP-MTR and RISP-CR: (a) actionable and useful? (b) Has any of it been used?<br><br>5. What incentives could change behaviours in terms of promoting critical analysis, best practices of results reporting, follow up, awards for innovation, increase of value of the knowledge created, learn from failure?<br><br>6. To what extent will the new RDN (as a separate document from the RISP) strengthen the Bank's analysis of regional integration issues and their linkages with national issues, leading to better alignment of the Bank's country and regional operational programmes<br><br>7. Are there concerns over ratings and disconnects between RISP-CR and RISP-E provided by IDEV that could distract from learning?<br><br>8. To what extent were leadership signals received that learning and knowledge management are key outcome of the RISP's SESP. | CSP/RISP guidelines.<br><br>Training of TMs and implementing agencies on integrating regional integration issues, M&E, procurement and Bank rules and regulations on E&S safeguards should be continuous due of the implementing staff. There is need to increase learning from RISP case studies on E&S safeguards and CC adaptation and also on SDG and mainstreaming guidelines of S&E should reach out countries and RECs.<br><br>New RDN (as a separate document from the RISP) was judged as a good opportunity to focus more on the challenges and opportunities for a strong regional integration that would strengthen the Bank's analysis of regional integration issues, policy convergence, macroeconomic management and on political economy of regional integration in Central Africa.<br><br>The RISP-CR has taken on board findings and lessons learned drawn from the BDEV regional strategy evaluation (2008-2016) and went beyond to suggest orientations and options for the next RISP. The main lessons identified at RISP-CR (2018) were to: (i) Improve RECs ownership and leadership in the implementation of regional program; (ii)<br>Take into account issues of fragility and resilience in the implementation of regional projects; (iii) Improve project preparation mechanisms and capacities at RECs level; (iv) establish specific criteria and indicators for monitoring regional integration; (v) Establish new regional frameworks for dialogue with other TFPs and RECs (in particular CEMAC and ECCAS), the private sector and civil society on regional integration issues; and (vi) continue to strengthen the links between national CSPs and the RISP. Key challenges and opportunities were defined and taken on board in the 2019-2025 RISP which came after the Bank's approval, in March 2018, of the new continental Regional Integration Strategic Framework (RISF) for the 2019-2025 period with several recommendations for enhancing the regional integration strategy implementation.<br><br>The feedback loop from RISP SESP that provide success-failure stories and results based on assessment of achievements in reaching regional integration should be encouraged with more opportunities for consultation and knowledge sharing and dissemination. Lessons learned should be institutionalized during life cycle and documented for capitalization of lessons learned in RISPs. There is need to increase the number of knowledgeable regional experts to build up strategic cooperation framework at regional level particularly with Private investors in building integrated regional infrastructure and with research centers in the region. |

| | |
|---|---|
| | |

**Learning Overall Assessment**

The involvement of regional economists and regional coordinators in the design and implementation phases of RISP programs contributes mainly to enhance the knowledge sharing on what regional integration entails. This was corrected by the new CSP/RISP guidelines.

There is need to increase learning from RISP case studies on political convergence, macroeconomic management and the political economy of the WA regional integration as well as for E&S safeguards and CC adaptation and also on SDG and mainstreaming guidelines of S&E should reach out countries and RECs.

New RDN (as a separate document from the RISP) was judged as a good opportunity to focus more on the challenges and opportunities for a strong regional integration that would strengthen the Bank's analysis of regional integration issues and their linkages with national issues, leading to better alignment of the Bank's country and regional operational programmes.

The feedback loop from RISP SESP on success-failure stories and results based on assessment of achievements in reaching regional integration was eagerly requested by regional staff. IDEV regional strategy program evaluation helped in capturing lessons and generating knowledge that enabled new strategic orientations notably of cross-cutting issues.

**CONCLUSIONS**

The RISP-CA has provided support for both institutions but due to capacity of institutions and provided the basis for the analysis of regional context, challenges and opportunities. The RISP was rather a programmatic document despite the analysis of regional integration and bottlenecks with a high number of fragile states for which the Bank did not offer, unfortunately a high-level policy dialogue to engage with the member states which have shown a low level of commitment. The RISP-MTR and CR were candid to highlight several issues with regard the performance of the multinational operations. They made, however, an attempt to align the regional integration strategic objectives, the Bank and continental priorities as well as with the country CSPs and the UN SDGs.

Discrepancies between RISP SESP in terms of results and portfolio management, were highlighted in the RISP-CR describing the low achievements and reasons for low performance of the program in terms of disbursement but in contradiction with the overall performance of the portfolio (3 out 4 point-scale). Such discrepancies between RISP SESP and the independent evaluation conducted by BDEV highlighted the need for more candor and rigorous assessments of effectiveness and efficiency and effective M&E systems within the implementing agencies and RECs and suggested high commitments of member states and high effectiveness and efficiency of RECs in implementing the regional programs.

Countries and RECs need to be involved in SESP in reporting on regional difficulties and bottlenecks or challenges as well as on opportunities for enhanced policy convergence, macroeconomic management, and the political economy of regional integration with security and political vulnerability issues.

As for the WA-RISP, the RISP-CA's institutional framework has to be reviewed to enhance reporting on results and integration into the ADER and RMF with regional experts' involvement in RISP assessments, policy dialogue, validations and corrective actions mainly of safeguards (example compensation in resettlements) and gender. The implementation support divisions should be well staffed with all subject matter experts such as gender, climate change, E&S experts, etc…

Lessons and recommendations in both RISP SESP and independent evaluation would offer an opportunity for the Bank to enhance the ownership and the level of commitments in leading donor coordination and partnerships and encourage high level consultations with private sector investors, civil society organizations and regional institutions and organizations.

The 2019-2025 followed the new guidelines and offered a better shape of the regional integration strategy underpinned by a regional diagnostic and several studies of the challenges, constraints and bottlenecks. The new RISP-CA sought to "improve economic diversification and structural transformation through the improvement of intra-regional trade in Central Africa" through two pillars, namely: (i) Strengthening regional infrastructure (energy, transport and ICT); and (ii) Supporting reforms for intra-regional trade development and build the institutional capacity of RECs.

New guidelines did not make it clear which from CSP or RISP is the main strategic document or the subsidiary and how they complement each other, according to regional coordinators. Reporting on supported actions to promote regional integration in RISP SESP was considered as of utmost importance to reflect on strategy implementation.

De-linking RPPRs from RIPS MTR or CR is not obvious with a high concentration of operations with little evidence on regional integration issues.

The specificity of the Central Africa region is the low commitment of its member states, low ownership and the low capacity of its regional institutions with a variety of complex mechanisms of regional integration.

The new RDN (as a separate document from the RISP) was judged as a good opportunity to focus more on the challenges and opportunities for a strong regional integration leading to better alignment of the Bank’s country and regional operational programmes. The feedback loop from RISP SESP on success-failure stories and results based on assessment of achievements in reaching regional integration was eagerly requested by regional staff. BDEV independent evaluation of the RISP period captured useful lessons and recommendations using a gradual approach to regional integration.

# RISP SESP ASSESSMENT TEMPLATE

## NARRATIVE ASSESSMENT OF RISP SESP DIMENSIONS

### RISP Basic Data

| | |
|---|---|
| ***Region:*** **EAST AFRICA** | |
| **Countries Covered:** | **13 countries**: Burundi, Comoros, Djibouti, Eritrea, Ethiopia, Kenya, Rwanda, Seychelles, Somalia, Sudan, Tanzania, Uganda and, as of 9 July 2011, the newly-independent Republic of South-Sudan. |
| **Countries Status: Fragile State/ADF/ADB** | Fragile States: 6 (Burundi*, Comoros*, Eritrea*, Somalia*, Sudan*, South Sudan*)<br>ADF: 5 (Djibouti**, Ethiopia, Rwanda, Tanzania, Uganda)<br>Blend ADB/ADF: 1 (Kenya)<br>ADB:1 (Seychelles) |
| **RISP Program 2011-2015 extended to 2016** | **Report Date: September 2011** |
| **RISP – MTR 2011-2015 and RPPR** | **Report Date: 16 January 2014** |
| **RISP-CR Report data: 2011-2016 and 2016 RPPR** | **Report Date: 9 January 2017** |

*Fragile and Conflict-affected States
**ADF-Gap

### NEW RISP 2018-2022[60]

| | |
|---|---|
| **Date of report:** | **September 2018** |
| | Mission date? |

## Overall RISP – RISP-MTR – RISP-CR AND RISP-E

| | Narrative Assessment |
|---|---|
| RISP RESULTS FRAMEWORK MONITORING MATRIX | The results framework monitoring matrix described the region's development objectives and constraints that hampered regional integration and projected key outputs and outcomes at MTR and at the end of the RISP period (2011-2015) with indicative on-going and new financing program under each pillar. Although based on sector notes and Bank strategic sector interventions, the matrix lacked quantitative targets and precision of output and outcome indicators of regional integration objectives[61] which were mostly related to the regional infrastructure operations (energy, transportation and ICT). Capacity building activities were soft components to build regional institutions *(RECs, Continental Organizations and National Implementation Unit),* with a support to transport and trade facilitation, customs modernization and reforms and aid for Trade. These also lacked quantifiable monitoring indicators despite a clear 2-step approach[62] defined by ADF-12 deputies (February 2011). The matrix was |

[60] The 2018-2022 RISP followed the new guidelines and template offering a better shape with regional diagnostic note and several sector notes and regional economic and sector work

[61] Non measurable expected results and confusion between outcome and output indicators and lack of sex dis-aggregated or gender related indicators were reported by the combined RISP-MTR report (January 2014 report). This has posed serious challenges to the assessment of regional integration and -specific results generated by the EA-RISP at mid-term.

[62] The five-year Regional Integration Strategy Paper should contain an indicative pipeline of operations recommended for Bank Group support with consideration of (i) Operations�??? alignment with the Bank's corporate priorities and the Regional Integration Strategy; (ii) Operations' alignment with

| | |
|---|---|
| | updated and revised at MTR and reviewed with comparison to actual results at completion highlighting envisaged objectives and achievements compared to the revised monitoring matrix. |
| RESULTS MONITORING | There was no specific summary results monitoring. However, new monitoring results matrix was prepared based on updates of the MTR assessment. |
| PROGRAMME IMPLEMENTATION EVALUATION | The programme implementation evaluation was undertaken at completion (2016) and stated the independent evaluation findings submitted in February 2017. The independent evaluation exercise cannot be considered as a validation exercise but rather an evaluation of a series of RISP programs. The new RISP 2018-2022 submitted in September 2018 benefitted from the lessons drawn at completion and by the independent evaluation. The implementation of the programme at completion highlighted several issues which necessitated the extension of the RISP period (2011-2015) to 2016. The program did not specific actions or activities with regard fragile states that belong to the region. Institutional capacity building (social and multi-sector) operations were suffering from the weak capacity of the RECs although the capacity issue of RECs including COMESA was recognized at the outset with identified risk factors. Yet, Bank's operations did not contribute significantly to reinforcing RECs capacities nor regional integration harmonized strategy. |
| PORTFOLIO ALIGNMENT WITH THE BANK'S HIGH 5S | The portfolio alignment with the high 5s was made under the preparation of the 2018-2022 RISP and the new results tool comprising a strategic alignment matrix (compliance with AfDB corporate policies and thematic strategies of continental and regional programs). |
| IMPLEMENTATION STATUS OF THE COUNTRY PORTFOLIO PERFORMANCE IMPROVEMENT PLAN | The implementation status of the RPIP was undertaken at MTR (2015) and at completion (with a revised RPIP for 2016). A new RPIP was prepared under the new 2018-2022 RISP specifying the portfolio performance issues, corrective measures to be undertaken, responsible entities and indicative monitorable indicators (qualitative). |
| BANK GROUP PORTFOL | RISP portfolio reviews were undertaken at MTR (2015) and Completion (2017) and RPIPS were prepared for the same period including for the 1st year and then2018-2022 RISP. Portfolio performance indicators were identified and reviewed during the RISP period with slight improvements in most of regional and multi-country operations (reduction of potentially problematic projects and projects as well as commitment at risks). Disbursement rate has improved between 2012 and 2016 from 33.6% to 41.2% (?). Challenges that hampered portfolio performance were identified and measures were proposed to be on a continuous basis and their implementation was closely monitored by the Bank, the RECs and the partner governments. Intensified dialogue with Governments and Executing Agencies facilitated by the EARC and the Bank's country offices in the region to improve the design and implementation of new operations. |
| RISP MTR | The implementation status of the program was assessed showing a progress of approvals but stagnation of disbursements at mid-term. The RISP MTR highlighted |

the priorities of the African Union, NEPAD, RECs and Regional Member Countries; (iii) Operations' impact on regional integration; and Ownership of participating countries and entities. The prioritization framework (Scorecard) includes indicators related to the participating countries (CPIA, portfolio performances of regional operations from APPR, countries trade facilitation policy commitments to regional integration) and regional operations expected development outcomes and contribution to regional integration and quality at entry and readiness (existence of MOU agreement between beneficiary countries, evidence of support from relevant RECs, and collaboration/co-financing with development partners). Source: Regional Operations Selection and Prioritization Framework, February 2011.

| | |
|---|---|
| | the inadequacy of the RISP original Results-Based Framework (RBF) and proposed a revised RBF with more monitorable indicators. Bank's performance was judged to be satisfactory with areas for improvement mainly in field supervision field offices, task management and application of more rigorous readiness filter at regional projects appraisal and approval stages. Governments and RECs performance was judged moderately satisfactory. Weaknesses exist in the adherence to project implementation schedules, due to weak oversight, inadequate coordination and weak contract management skills of implementing agencies. Lessons learned were identified for the Bank as well as Governments, RECS and development partners with strategic options to enhance regional integration strategy implementation, M&E of results and RECs capacity building initiatives. |
| RISP CR | The RISP-CR was prepared based on the RISP-MTR with assessment of performance in terms of results (outputs and outcomes) but did not match progress made in terms of regional project/program approvals. The CR report highlighted procurement issues with substantial delays at start-up showing few tangible results to report. Only a few projects have been completed at RISP completion. The RISP-CR as well as implementation experience from the Bank's RISPs and portfolio reviews raised a large number of lessons at the strategic and portfolio management levels, as well as at RECs'/RMCs' and Development Partners' levels, which have informed the Bank's new RISP 2018-2022. It has also highlighted much greater attention that needed to be given to 'soft' issues hindering deeper regional integration and trade, and to promote the structural transformation of the region through industrialization and value chain development. However, there was no specific action to accompany fragile states to be covered by the transformative change proposed in the RISP-MTR. |
| Overall RISP SESP Instruments Assessments: The RISP SESP reports were analytical and attempted to assess candidly and proactively the results achieved during the RISP period while highlighting the challenges and bottlenecks to regional integration strategy and programs. These are mostly program oriented although regional integration issues were analysed. The monitoring results matrix was revised at MTR due to lack of quantifiable monitoring results and precise measurement methods. This was clarified at RISP CR which offered a candid assessment of results and identified options for the new strategic orientation of the RISP while raising an important issue related to the close coordination and 'division of labor' between the RISP and country CSPs of the region to maximize complementarity and synergies. While CSPs focus on supporting the development of sources of transformative growth (growth-poles), through targeted support at micro-level, RISP focuses rather on removing bottlenecks to free movement/evacuation of goods (and also of services and capital) generated by those growth poles, by applying a spatial approach through corridor development. However, there was no strategic action to remove vulnerability and fragility of the affected countries. Discrepancies between RISP SESP and independent regional strategy evaluation in terms of results and portfolio management highlighted the need for timely validations of RISP-MTR and CR and RPPRs in order to enhance the candor and rigor in assessing performance at the strategic, portfolio management, capacity building and partnership levels and in informing the new 2018-2022 RISP. Lessons and recommendations in both RISP SESP and independent evaluation offered an opportunity to enhance the evaluability of the new RISP and reflect on new measurement and monitoring framework for results achievement. | |

| 1- **PERFORMANCE MANAGEMENT** | Narrative Assessment |
|---|---|
| 1. (a) Is the RISP considered a strategy document or more of a programmatic document for | The RISP was underpinned by the Strategic Frameworks of both EAC and COMESA and the COMESA-EAC-SADC (CES) Tripartite Arrangement. Support for the Tripartite Arrangement was the key |

| | |
|---|---|
| regional projects; or both?<br>(b) When evaluating the RISP, what weight do you give to its different components: strategy vs. portfolio?<br><br>2. RISPs can cover several CSPs over more than one cycle, and a portfolio of projects with various durations. What are the main criteria used in assessing the performance of the RISP through its main outputs: MTR, RPPR, CPPR and the CR?<br><br>3. Are the RISP SESP outputs (MTR-RPPR-CR) aligned with:<br>• Main operational policy documents?<br>• High 5s?<br>• Environmental and social safeguards such as: gender, fragility, climate change?<br>• Fiduciary & governance policy? | feature of the RISP.<br>The RISP SESP provided the basis for the analysis of regional integration strategies' implementation risks related to performance and results. The RISP was rather a programmatic document despite the analysis of regional integration and bottlenecks with a high number of fragile states. The RISP was considered a stand-alone document with its financing program that may or may not be part of the individual CSPs such as the multi-country operations. As specified in the ADF-12 paper, the five-year Regional Integration Strategy Paper should have contained an indicative pipeline of operations recommended for Bank Group support with consideration of (i) Operations‟ alignment with the Bank's corporate priorities and the Regional Integration Strategy; (ii) Operations' alignment with the priorities of the African Union, NEPAD, RECs and Regional Member Countries; (iii) Operations' impact on regional integration; and Ownership of participating countries and entities. The prioritization framework (Scorecard) includes indicators related to the participating countries (CPIA, portfolio performances of regional operations from APPR, countries trade facilitation policy commitments to regional integration) and regional operations expected development outcomes and contribution to regional integration and quality at entry and readiness (existence of MOU agreement between beneficiary countries, evidence of support from relevant RECs, and collaboration/co-financing with development partners). |
| 4. To what extent have alignment and/or complementarity between RISPs and CSPs been dealt with as some projects faced challenges of ownership and cross-country coordination since regional operations are typically implemented at national levels? | New guidelines did not make it clear which from CSP or RISP is the main strategic document or the subsidiary and how they complement each other.<br>Regional integration coordinators were not conversant with M&E and impact tools of regional projects and able to lead the process in conducting regional diagnostic and understanding the political economy of the regional integration (actually regional economists dominate the design of RISPs). |
| 5. RPPRs and CPPRs covered by the RISP include a rating on a scale of 1 to 4, which is based on a simple average of reported performance of public sector operations. How well were other non-lending activities considered as part of the RISP self-evaluation?<br><br>6. Is the recent decision to separate the regional diagnostic note (RDN) from the RISP per se, likely to facilitate the alignment with the Bank's key strategic documents and the robustness of the RISP self-evaluation process | With the recent guidelines on regional diagnostic note (RDN) preceding the RISP preparation would, according to regional coordinator, facilitate the understanding of the strategic components of the regional integration strategy (policy convergence, spatial integration, capacity strengthening of the RECS and regional institutions to support policy development and harmonized macro-economic management, etc..). Also, reporting on supported actions to promote regional integration in RISP SESP was considered as of utmost importance to reflect on strategy implementation while regional programs and multi-country projects results assessment can provide only a trend analysis of the regional integration strategy by building regional infrastructure to support transport and trade facilitation, customs modernization and reforms and aid for Trade. These have lacked quantifiable monitoring indicators despite a clear 2-step approach described in the ADF-12 paper. |

| | |
|---|---|
| in the MTR and the CR?<br><br>7. (a) Should SESP outputs of the RISP cover both the RDN and the RISP document or only the latter? (b) How could the RISP-MTR and RISP-CR better reflect the performance at the regional level and its country components?<br><br>8. Are the SESP main outputs (RISP-MTR, CR, Annual Portfolio Reviews) geared towards addressing:<br>(a) strategic issues,<br>(b) policy dialogue,<br>(c) knowledge management?<br><br>9. Was IDEV validation of the projects PCRs, CSPs and RISP-Es factored-in when discussing the performance of the regional portfolio and of the RISP? Does external validation help improve the candour of the self-assessment? Should the RISP-CR be rated? | RISP-MTR and CR offered a candid assessment of results and identified options for the new strategic orientation of the RISP while raising an important issue related to the close coordination and 'division of labor' between the RISP and country CSPs in the region to maximize complementarity and synergies. While CSPs focused on supporting the development of sources of transformative growth (growth-poles), through targeted support at micro-level, RISP focused rather on removing bottlenecks to free movement/evacuation of goods (and also of services and capital) generated by those growth poles, by applying a spatial approach through corridor development. However, there was no strategic action to remove vulnerability and fragility of the affected fragile countries in the region.<br><br>RISP SESP were aligned with main operational policies and sector strategies but areas such E&S safeguards were mostly covered at national levels. Gender issues were not sufficiently covered in regional and multi-country projects, RISPs from design to completion were more output oriented due to weak methodology in anticipating and assessing results (lack of outcome indicators) as gender experts are not systematically involved during project cycle and RISP mid-term reviews or RISP-CRs.<br><br>RISP were not based on clear theory of change (TOC) both on operational and strategic level in countries or region. RISP Results framework did not inform on regional integration progress nor on macroeconomic convergence.<br><br>RECs and implementation agencies and beneficiaries' involved in RISP SESP preparation increased their validity based on agreed KPIs and deliverables. However, capacity development of RECs (including in statistics and impacts assessments) was necessary to enhance the quality of design and follow up of strategic development objectives of the region but were not rated. Independent validation of RPPRs was considered as a good way to ensure candor and objectivity and reliability of RISP SESP.<br>Disregarding numerical ratings in assessing implementation performances as well as Bank and Governments or RECs performance may encourage better dialogue focusing on strategic objectives, according to regional economists and regional coordinators. Consultations/coordination with implementing agencies in the region did happen but their involvement should be enhanced to increase the RISP SESP validity.<br><br>Overall portfolio performance rating improved from 3.4 in 2012 to 3.8 in 2016 (on a 4-point scale of 1 to 4). The implementation progress (IP) rating improved from 3.2 to 3.7, while the development objective (DO) rating also improved from 3.6 to 3.9. Discrepancies with independent evaluation based on evaluation criteria such as relevance, effectiveness, efficiency and sustainability |

| | |
|---|---|
| | showed an overall rating less than satisfactory (moderately satisfactory for relevance and effectiveness and moderately unsatisfactory for efficiency and sustainability). The discrepancy was also around the disbursement ratio (26% according to IDEV independent evaluation) after taking into consideration all approved regional and multi-country operations as at February 2016). Time efficiency as it relates to project start-ups, large-scale regional infrastructure projects in the transport and power sectors have encountered severe delays (15 to 23 months) with no distinct difference between operations under Pillars (hard and soft). Sustainability of results was judged as weak by IDEV evaluation with variations across sectors. Regional programs and projects in the financial and transport sectors fared better on sustainability compared to the agriculture sector and institution building operations, due to the lack of convergence.<br><br>RPPRs were not synchronized with RISP timeline and PCREN were fed to staff and considered for official ratings in RPPRs. RMF did not cascade to allow meaningful IPR reports on regional programs and results matrices were mostly not adapted to country and region context with M&E reporting systems.  Results frameworks lack smart indicators mainly on the social sector (poverty assessment, vulnerability, gender disparities, etc..).<br>RISP SESP should also report on processes to deliver sector results on social and environmental safeguards. These processes should be described and documented at inception (early engagement of safeguards during post approval) with adequate risk assessment.<br><br>De-linking RPPRs from RIPS MTR or CR may allow to focus more on strategic issues and few on operational issues of strategic importance. The flagship highlights are insufficient to deal with both implementation and development issues at the regional level. RISPs SESP in the region with a high number of fragile states lacked valid and reliable data on implementation and results due to conflicts and unavailability of data.<br>The implementation support division may improve the situation but there is need to share findings among these divisions for problem anticipation and timely corrections.<br><br>Disregarding numerical ratings in assessing implementation performances as well as Bank and Governments or RECs performance may encourage better dialogue focusing on strategic objectives, according to regional economists and regional coordinators. Consultations/coordination with implementing agencies in the region did happen but their involvement should be enhanced to increase the RISP SESP validity |

**Performance Management Overall Assessment**

The RISP SESP provided the basis for the analysis of regional integration strategies' implementation risks related to performance and results. However, the RISP was rather a programmatic document despite the analysis of regional integration and bottlenecks with a high number of fragile states. It was considered as a stand-alone document with its financing program that may or may not be part of the individual CSPs such as the multi-country operations as it was prepared before the ADF-12 directive on RISP preparation.

New guidelines did not make it clear which from CSP or RISP is the main strategic document or the subsidiary and how they complement each other. Reporting on supported actions to promote regional integration in RISP SESP was considered as of utmost importance to reflect on strategy implementation while regional programs and multi-country projects results assessment can only provide a trend analysis of the implementation of the regional integration strategy by building regional infrastructure to support transport and trade facilitation, customs modernization and reforms and aid for Trade, etc…. De-linking RPPRs from RIPS MTR or CR may allow to focus more on strategic issues and few on operational issues of strategic importance. The flagship highlights are insufficient to deal with both implementation and development issues at the regional level. RISPs SESP in the region with a high number of fragile states lacked valid and reliable data on implementation and results due to conflicts and unavailability of data.

**ACCOUNTABILITY**

| | |
|---|---|
| 1. (a) Are roles and responsibilities sufficiently clear in the preparation, conduct, review, sign-off, follow-up for the various steps of the RISP SESP outputs (country team, regional team, country managers, DG and chief economist)?<br>(b) Do Regional DG and Country managers see the RISP SESP outputs as a relevant accountability tool?<br>2. Do SESP of the RISP outputs provide a relevant perspective on the results achieved and communicate overall performance in a credible way?<br>3. Is Regional and HQ Management exerting leadership over the correct implementation of the RISP SESP outputs and lines of accountability?<br>4. Are TMs of regional programs being held accountable for timely submission and proper implementation of the RISP SESP outputs?<br>5. To what extent do SESP outputs | Regional Directorate plays a fundamental role in adopting an operational approach to RISPs and RISP SESP by engaging ADB staff in adapting country specific strategies to the regional context. The regional delivery unit helps assist the regional directorate in reviewing what has worked or not in terms of regional strategic implementation, achievements and results including the risk factors around them.<br><br>The lack of staff in sector divisions did not allow for continued strategic and operational dialogue and for increased supervision and just-in time actions to address regional project or multi-country projects' implementation issues. The SESP at regional integration project level is restricted to operational issues and only mid-term reviews provide an opportunity to reflect on the adequacy of Bank's response to the regional integration issues, challenges and bottlenecks. RISP SESP are thus considered as both for accountability and for knowledge base of it entails in promoting regional integration. This did not happen in East Africa RISP which did not widen the scope of the regional strategy encapsulated in the tripartite approach: EAC and COMESA and the COMESA-EAC-SADC (CES) Tripartite Arrangement which was the key feature of the EA-RISP.<br><br>The Bank recognizes that the RISP SESP are only for addressing implementation issues and allowing mistakes to be corrected for the good of the region. However, involving all actors, beneficiaries, RECs, Governments and other regional institutions increases their objectivity and validity.<br><br>Accountability is based on management validation and accuracy of performance data in RISP SESP. The regional team concept is not yet |

| | |
|---|---|
| of regional program: (a) rely on strong M&E systems? (b) offer a credible accountability framework?<br><br>6. To what extent do managers discuss with TMs regional program performance and results during staff performance conversations?<br><br>7. Does SESP aggregation of RISP products provide a relevant and cost-effective reporting of results through the RMF, ADER and other reporting tools (dashboard, etc..)? | anchored with no integrated convergence of national strategies encapsulated in CSPs. Collaboration with RECs proved to be difficult without strong capacity building in terms of policy convergence, harmonization of macro-economic policies and a good understanding of the political economy of how regional integration is built.<br><br>RISP SESP outputs did provide a relevant perspective on results but were more focused on the regional and multi-country operations based on assessment of the IP and DOs from the RPPRs with no independent validation. With regard the results monitoring indicators, these were reviewed at MTR and revised at RISP-CR to enhance their credibility. TMs of regional programs were held accountable for their timely submissions RISP SESP but did not provide a discussion of task management and application of more rigorous readiness filter at regional projects appraisal and approval stages. Governments and RECs performance was judged moderately satisfactory and weaknesses were highlighted at RISP-MTR on the adherence to regional project implementation schedules, due to weak oversight, inadequate coordination and weak contract management skills of implementing agencies and RECs. Lessons learned were identified for the Bank as well as for Governments, RECS and development partners with strategic options to enhance regional integration strategy implementation, M&E of results and RECs capacity building initiatives.<br><br>The regional directorate (and also the RDVP) are exerting leadership for the implementation of the regional program both hard and sift components but more focused on the operational side without a sound strategy to policy convergence. The specificity of the Eastern African region with its variety of regional institutions limited the scope of the 2011-2015 RISP to the tripartite arrangement cited above, and increased the complexity of how regional integration should entail.<br><br>Discrepancies between RISP SESP and independent regional strategy evaluation in terms of results and portfolio management highlighted the need for effective M&E systems within the implementing agencies and RECS which the Bank strived to introduce as part of the soft components of the regional integration strategy. However, this did not happen. Furthermore, timely validations of RISP-MTR and CR and RPPRs in order to enhance the candor and rigor in assessing performance at the strategic, portfolio management, capacity building and partnership levels and in informing the new 2018-2022 RISP was judged as of utmost importance. Lessons and recommendations in both RISP SESP and independent evaluation offered an opportunity to enhance the **evaluability** of the new RISP with M&E approach to apply measurement and monitoring framework for results achievement.<br><br>Countries and RECs should be involved in SESP to report on E&S |

| | |
|---|---|
| | safeguards with adequate reporting guidelines on E&S management/mitigation plan. Status of E&S safeguards in portfolio flagship reports is missing.<br><br>Regional coordinators for E&S safeguards and climate change are few and cannot cover all countries and regions. They are not systematically involved in field missions or meetings of RISP meetings although they are custodian of the Bank's climate change action plan (2011-2015). Climate change, social and environmental safeguards should be used as support team to project and RISP preparation and implementation and involved in early stage of project/RISP design (pre-board stages) for better recognition of the subject matter.<br><br>Few gender specialists are located in regions. Gender policy guidelines have to be reviewed as no clear plan of action is in place to report on results in SESP. No dedicated budget is in place to cover specifically gender issues during regional programs' preparation, negotiation down to completion.<br><br>The institutional framework should be reviewed to enhance reporting on results and integration into the ADER and RMF with regional experts' involvement in RISP assessments, policy dialogue, validations and corrective actions mainly of safeguards (example compensation in resettlements) and gender. The implementation support divisions should be well staffed with all subject matter experts such as gender, climate change, E&S experts, etc… |

**Accountability Overall Assessment**

Regional Directorate plays a fundamental role in adopting an operational approach to RISPs and RISP SESP by engaging ADB staff in adapting country specific strategies to the regional context. The regional delivery unit helps assist the regional directorate in reviewing what has worked or not in terms of regional strategic implementation, achievements and results including the risk factors around them.

The lack of staff in sector divisions did not allow for continued strategic and operational dialogue and for increased supervision and just-in time actions to address regional project or multi-country projects' implementation issues.

involving all actors, beneficiaries, RECs, Governments and other regional institutions increases their objectivity and validity.

Accountability is based on management validation and accuracy of performance data in RISP SESP. However, integrated convergence of national strategies encapsulated in CSPs did allow for strong accountability of RISP SESP tools. Collaboration with RECs proved to be difficult without strong capacity building in terms of policy convergence, harmonization of macro-economic policies and a good understanding of the political economy of how regional integration is built.

The specificity of the Eastern African region with its variety of regional institutions limited the scope of the 2011-2015 RISP to the tripartite arrangement cited above, and increased the complexity of how regional integration should entail. The lack of effective M&E systems within the implementing agencies and RECS limited the assessment of results, and their attribution.

| | |
|---|---|
| Lessons and recommendations in both RISP SESP and independent evaluation offered an opportunity to enhance the evaluability of the new RISP with M&E approach to apply measurement and monitoring framework for results achievement. | |
| **3. Learning** | |
| 1. To what extent have the RISP SESP outputs for regional program been used as a source of learning and knowledge management to inform the next RISP or other regional operations/strategies<br><br>2. To what extent has the RPPR or CPPRs been used for the preparation of the RISP- MTR and the RISP-CR?<br><br>3. (a) Did feedback loop sessions and lessons learned discussions take place after RISP- MTR or RISP CR? (b) Have these feedback sessions led to better procedures for regional programs, restructuring, differentiation according to regional situations (fragile context, MIC, non-lending)<br><br>4. To what extent are recommendations and lessons learned described in RISP-MTR and RISP-CR: (a) actionable and useful? (b) Has any of it been used?<br><br>5. What incentives could change behaviours in terms of promoting critical analysis, best practices of results reporting, follow up, awards for innovation, increase of value of the knowledge created, learn from failure?<br><br>6. To what extent will the new RDN (as a separate document from the RISP) strengthen the Bank's analysis of regional integration issues and their linkages with national issues, leading to better | Peer reviews and independent review process encouraged the learning from RISPs SESP and improved their quality together with the Economic and Sector Work that RISPs recommended and strived for their implementation in order to enhance their evaluability and results.<br><br>The involvement of regional economists and regional coordinators in the design and implementation phases of RISPs contributes mainly to enhance the knowledge sharing on what regional integration entails. This was corrected by the new CSP/RISP guidelines.<br><br>Undertaking special supervision solely on these issues is too late, not productive or influencing. Training of TMs on integrating these issues should be continuous because of the staff turnover and procedures simplified for better mainstreaming. There is need to increase learning from RISP case studies on E&S safeguards and CC adaptation and also on SDG and mainstreaming guidelines of S&E should reach out countries and RECs.<br><br>New RDN (as a separate document from the RISP) was judged as a good opportunity to focus more on the challenges and opportunities for a strong regional integration that would strengthen the Bank's analysis of regional integration issues and their linkages with national issues, leading to better alignment of the Bank's country and regional operational programmes.<br><br>There is no feedback loop from RISP SESP that provide success-failure stories and results based on assessment of achievements in reaching regional integration. Lessons learned should be institutionalized during life cycle and documented for capitalization of lessons learned. IDEV country/regional strategy program evaluations helped in capturing lessons and generating knowledge that enabled new strategic orientations notably of cross-cutting issues such as climate change, gender and regional disparities. There is need to increase the number of knowledgeable regional experts to build up strategic cooperation framework at regional level particularly with Private investors in building integrated regional infrastructure.<br>Lessons learnt are well captured at RISP levels with regard FM due to development partners consultations. Lessons from implementation of PFM programs and support of PFM agenda are well incorporated in CSPs. But not in the regional integration project and program and in RISP SESP. |

| | |
|---|---|
| alignment of the Bank's country and regional operational programmes<br><br>7. Are there concerns over ratings and disconnects between RISP-CR and RISP-E provided by IDEV that could distract from learning?<br><br>8. To what extent were leadership signals received that learning and knowledge management are key outcome of the RISP's SESP. | |

**Learning Overall Assessment**

The involvement of regional economists and regional coordinators in the design and implementation phases of RISPs contributes mainly to enhance the knowledge sharing on what regional integration entails. This was corrected by the new CSP/RISP guidelines.

There is need to increase learning from RISP case studies on E&S safeguards and CC adaptation and also on SDG and mainstreaming guidelines of S&E should reach out countries and RECs.

New RDN (as a separate document from the RISP) was judged as a good opportunity to focus more on the challenges and opportunities for a strong regional integration that would strengthen the Bank's analysis of regional integration issues and their linkages with national issues, leading to better alignment of the Bank's country and regional operational programmes.

The feedback loop from RISP SESP on success-failure stories and results based on assessment of achievements in reaching regional integration was eagerly requested by regional staff. IDEV regional strategy program evaluation helped in capturing lessons and generating knowledge that enabled new strategic orientations notably of cross-cutting issues.

**CONCLUSIONS**

The RISP was underpinned by the Strategic Frameworks of both EAC and COMESA and the COMESA-EAC-SADC (CES) Tripartite Arrangement. <it provided support for the Tripartite Arrangement as the key feature of the RISP prepared on the basis of analysis of regional integration strategies' implementation risks related to performance and results. The RISP was rather a programmatic document despite the analysis of regional integration and bottlenecks with a high number of fragile states. The RISP was considered a stand-alone document with its financing program that may or may not be part of the individual CSPs such as the multi-country operations.

New guidelines did not make it clear which from CSP or RISP is the main strategic document or the subsidiary and how they complement each other, according to regional coordinators. Reporting on supported actions to promote regional integration in RISP SESP was considered as of utmost importance to reflect on strategy implementation while regional programs and multi-country projects results assessment can only provide a trend analysis of the implementation of the regional integration strategy with support to transport and trade facilitation, customs modernization and reforms and aid for Trade, etc….

De-linking RPPRs from RIPS MTR or CR may allow to focus more on strategic issues and few on operational issues of strategic importance.

| |
|---|
| The specificity of the Eastern African region with its variety of regional institutions limited the scope of the 2011-2015 RISP to the tripartite arrangement cited above, and increased the complexity of how regional integration should entail. The lack of effective M&E systems within the implementing agencies and RECs limited the assessment of results, and their attribution.<br><br>Lessons and recommendations in both RISP SESP and independent evaluation offered an opportunity to enhance the evaluability of the new RISP with M&E approach to apply measurement and monitoring framework for results achievement.<br><br>New RDN (as a separate document from the RISP) was judged as a good opportunity to focus more on the challenges and opportunities for a strong regional integration leading to better alignment of the Bank's country and regional operational programmes. The feedback loop from RISP SESP on success-failure stories and results based on assessment of achievements in reaching regional integration was eagerly requested by regional staff. IDEV regional strategy program evaluation helped in capturing lessons and generating knowledge that enabled new strategic orientations notably of cross-cutting issues. |

# RISP SESP ASSESSMENT TEMPLATE

## NARRATIVE ASSESSMENT OF RISP SESP DIMENSIONS

### RISP Basic Data

| | |
|---|---|
| ***Region:*** **WEST AFRICA** | |
| **Countries Covered:** | **15 countries**: ECOWAS region: Benin, Burkina Faso, Cape Verde, Côte d'Ivoire, Gambia, Ghana, Guinea, Guinea Bissau, Liberia, Mali, Niger, Nigeria, Senegal, Sierra Leone, and Togo. Eight of the countries are francophone, 5 are Anglophone and, two are lusophone. |
| **Countries Status: Fragile State/ADF/ADB** | Fragile States: 6[63] (Gambia*, Guinea*, Guinea Bissau*, Liberia*, Niger*, and Sierra Leone*)<br>ADF-only: 4 (Benin, Burkina Faso, Mali, Togo)<br>ADF Gap: 2 (Côte d'Ivoire, Ghana)<br>Blend: 1 (Senegal)<br>Asîring ADB: 1 (Nigeria)<br>ADB:1 (Cape Verde) |
| **RISP Program 2011-2015 extended to 2017** | **Report Date: March 2011 Extension: April 2016** |
| **RISP – MTR 2011-2015 and 2015 RPPR** | **Report Date: May 2014** |
| **RISP-CR Report data: 2011-2017 and 2017 RPPR** | **Report Date: 22 February 2019[64]** |

### NEW RISP (Under preparation)

| | |
|---|---|
| **Date of report:** | **?** |
| | Mission date? |

## Overall RISP – RISP-MTR – RISP-CR AND RISP-E

| SESP Instruments | Narrative Assessment |
|---|---|
| RISP RESULTS FRAMEWORK MONITORING MATRIX | The results framework monitoring matrix described the region's development objectives and constraints that hampered regional integration and projected key outputs and outcomes at MTR and at the end of the RISP period (2011-2015) with indicative on-going and new financing program under each pillar. It was based on economic and sector work which focused on regional integration leverages (macro-economic management, transport, ports, energy, trade, ICT, monetary policy convergence and monetary cooperation program convergence criteria etc..). The matrix included quantitative targets of outputs and outcomes for the main strategic objectives (linking markets through road corridors, transport and trade facilitation; regional energy production and market integration; support to integrated financial sector, and capacity building for enhanced implementation of the regional integration agenda, effective policy and multi-country projects' implementation, support to centres of excellence, and statistical support to ECOWAS institutions). The |

[63] The Bank opted to engage through a short-term programming document (CB) until conditions on the ground improve to a point where a CSP could be prepared. Source: WA-RISP-CR (2017)

[64] A validation of the WA RISP-CR was conducted by BDEV on a pilot basis in March 2019 (Draft Validation Note 22 March 2019)

| | |
|---|---|
| | capacity building results lacked measurable indicators but projects selection was made as an attempt to follow the 2-step approach[65] defined by ADF-12 deputies (February 2011). The matrix was updated and revised at MTR. The RISP period was extended by 2 years (up to 2017). The reasons for the extension were: i) to synchronize the Bank's strategy with the preparation and launching of the new ECOWAS strategic plan 2016-2020, ii) to allow the Bank to finalise the preparation of the new format, guidelines and templates of future RISPs and iii) to provide sufficient time for the preparation of the WA RISP 2018-2022. The results matrix was slightly reviewed and updated based on a comparison to actual results at completion (2017). A pilot independent validation of the WA-RISP-CR highlighted the limited information on outputs and a failure to provide details on the progress made in achieving intermediate outcomes and outcomes. However, the pilot validation exercise did not use the results monitoring framework as a basis for its validation nor assessed its validity and reliability as a main M&E tool. |
| RESULTS MONITORING | There was no specific summary results monitoring. However, new monitoring results matrix was prepared based on updates of the MTR assessment and on CR after the 2-year extension. The main weakness is around delays in approval and implementation of multi-year operations mainly for capacity building operations. However, the RISP was flexible enough to accommodate new interventions following the Ebola crisis in the region showing its flexibility and ability to align with regional priority needs. |
| PROGRAMME IMPLEMENTATION EVALUATION | The programme implementation self-assessment was undertaken at MTR and completion (2017). The independent validation exercise of the RISP-CR provided critical assessment on the CR report which lacked precise reporting on the program outputs and outcomes. The new RISP (2020-2024)(?) is underway but will benefit from the lessons drawn at completion and by the independent validation exercise. The implementation of the programme at MTR highlighted several issues which necessitated the extension of the RISP period (2011-2015) to 2017. The program did not include specific actions or activities with regard fragile states that belong to the region. Institutional capacity building (social and multi-sector) operations' implementation was difficult due to the weak capacity of the RECs. |
| PORTFOLIO ALIGNMENT WITH THE BANK'S HIGH 5S | The portfolio alignment with the high 5s was made under the preparation of the RISP-CR. An alignment with the country CSPs was made to ensure consistency with the RISP pillars in RISP-MTR and CR. |
| IMPLEMENTATION STATUS OF THE COUNTRY PORTFOLIO PERFORMANCE IMPROVEMENT PLAN | The implementation status of the RPIP was undertaken at MTR (2015) and at completion (with a revised RPIP for 2018). The portfolio performance issues, corrective measures to be undertaken, responsible entities and indicative monitorable indicators (qualitative) were provided in the 2018 RPIP. |
| BANK GROUP PORTF( | RISP portfolio reviews were undertaken at MTR (2014) and at completion (2017) with |

[65] The five-year Regional Integration Strategy Paper should contain an indicative pipeline of operations recommended for Bank Group support with consideration of (i) Operations�much alignment with the Bank's corporate priorities and the Regional Integration Strategy; (ii) Operations' alignment with the priorities of the African Union, NEPAD, RECs and Regional Member Countries; (iii) Operations' impact on regional integration; and Ownership of participating countries and entities. The prioritization framework (Scorecard) includes indicators related to the participating countries (CPIA, portfolio performances of regional operations from APPR, countries trade facilitation policy commitments to regional integration) and regional operations expected development outcomes and contribution to regional integration and quality at entry and readiness (existence of MOU agreement between beneficiary countries, evidence of support from relevant RECs, and collaboration/co-financing with development partners). Source: Regional Operations Selection and Prioritization Framework, February 2011.

| | |
|---|---|
| | RPIP prepared for 2018. Portfolio performance indicators were identified and reviewed during the RISP period with slight improvements in most of regional and multi-country operations (reduction of potentially problematic projects and projects as well as commitment at risks).<br>Portfolio rating improved from 2.07 in 2013 (on a scale 1-3, equivalent to 2.76 on a scale 1-4) to 3.26 in 2018. As at August 1, 2018 only one project was classified as problematic, down from two identified during the mid-term in 2014, while the commitment at risk reduced drastically from 16 percent to about 1 percent.<br>Disbursement rate has improved between 2014, 2016 and 2018 from 41%, 33.6% to 47.2% during the period. The validation exercise stated that the actual disbursement rate for the public sector (representing 78 out of the 86 operations) is only 23% suggesting that disbursement performance for public sector regional projects is less than satisfactory.<br>The Bank Group regional portfolio reflected the economic development challenges in the WA region, especially the recent efforts of the Bank Group in addressing the Ebola Virus crisis as captured by the social sector operations, as well as the transportation and energy sector. Private sector operations have also been important.<br>Portfolio implementation challenges in West Africa include: (i) delay in the implementation of loan agreements, (ii) delay in the fulfilment of conditions precedent to first disbursement, (iii) insufficient knowledge of the Bank's rules and procedures, particularly with regard to procurement and disbursements, and (iv) weak implementation capacity. |
| RISP MTR | The implementation status of the program was assessed showing a limited progress of approvals and low disbursement ratio (31%). The RISP MTR revised slightly the pillars and attempted to include more monitorable indicators. The main reasons for late start-up of project implementation and disbursement are the lack of follow up system of conditions precedent the loan agreement signature or loan effectiveness and the lack of coordination between the bank and the RECs. The private sector regional operations showed better performance and helped create the necessary conditions for private sector development in the region. Regional operations directly managed by the RECs (ECOWAS and WAEMU), respectively 4 and 3 operations were facing implementation difficulties due to their quality at entry and lack of supervision from the Bank and the weak capacity of the concerned RECs.<br><br>The RISP-MTR was candid in highlighting the main reasons for insufficient performance of the Bank: inefficient supervision, highly centralized decision-making process due to the relocation of the Bank in Tunis, lack of real time technical assistance to RECs and insufficient familiarization of implementing agencies on Bank's rules and procedures. Governments and RECs performance were judged less that satisfactory due to the weak participation of RECs in the design of the operations, lack of monitoring and evaluation systems, delays in selecting the implementation team and the lack of delegation of authorities. Governments performance is hampered by the lack of coordination during the multi-country and regional integration operations and insufficient consultations before multi-country project start-ups and the lack of dissemination and sharing of information on the regional projects. |

| | |
|---|---|
| | Risk factors' assessment related to the dual accountability and multiplicity mechanisms of the integration architecture (Bank, RECs and Government levels) and the political fragility in the WA region has led to the identification of lessons learned with strategic options to enhance regional integration strategy implementation, M&E of results and RECs capacity building mainly by increasing the familiarization of implementing agencies on Bank's rules and procedures, sharing knowledge products on the WA region and dissemination of the regional research centers on policy convergence, regional macroeconomic policies and the political economy management.<br>The extension of the 2011-2015 RISP to 2017 highlighted several issues such as the political and security vulnerability and fragility of the region. Bank Group's fragility assessments provided critical inputs necessary to incorporate fragility-mitigating actions into the design and implementation of regional integration and country strategies (for Mali, Niger, Togo, Guinea Bissau, Sierra Leone and political economy analysis for Côte d'Ivoire and Burkina-Faso) and implemented resilience strengthening program in the Sahel. |
| RISP CR | The RISP-CR was prepared based on the RISP-MTR with assessment of performance in terms of results (outputs and outcomes) but did not match progress made in terms of regional project/program approvals. The CR report highlighted procurement issues with substantial delays at start-up showing few tangible results to report. Only a few projects have been completed at RISP completion (2017). The RISP-CR discussed very well the regional context and assessed the results achievements, constraints, challenges and opportunities, and discussed the pillars adjustments (Pillar1 was adjusted to increased Bank's support to infrastructure development and to addressing fragility, food security and resilience issues in the Sahel; Pillar 2 was adjusted and expanded to cover capacity building for financial sector integration, for effective policy and regional projects implementation, support to Regional Centres of Excellence, and statistical support to ECOWAS institutions).<br>The pilot independent validation exercise, based on a desk review, assessed the quality of evidence provided to support the self-assessment and contribution to the new Regional Integration Strategy. The validation aimed at contributing to improving the quality of RISP completion reports and providing an opportunity for learning. **T**he validation covered three dimensions of the OECD/DAC evaluation criteria: relevance, effectiveness and efficiency. The validation also assessed the Bank's performance in managing risks, policy dialogue, partnership and aid coordination.<br>The validation concluded that the CR lacked sufficient evidence to support its main conclusion and did not report in details the achievement of development objectives for the 2011-2017 RISP period. The CR should have reported on results up to 2018. In addition, the results were reported in some cases at the pillar level (i.e. infrastructure projects) which were clearly linked to performance through supported projects. The CR results framework was not clearly linked to specific projects. Although the RISP aimed to mainstream cross cutting issues in the implementation, this was not taken into consideration in the results framework with only scant details provided in the CR regarding results achieved. The assessment of the efficiency of the RISP CR implementation was not rigorous, although delays were noted. The RISP-CR did not explicitly assess the likely contribution or attribution to the delays in achieving regional and multi-country project results. Finally, the CR did not assess the extent to which the RISP implementation achieved its ultimate goals of advancing the regional integration agenda and facilitating regional solutions. |

| |
|---|
| **Overall RISP SESP Instruments Assessments**: The RISP SESP reports were analytical and attempted to assess candidly and proactively the results achieved during the RISP period while highlighting the challenges and bottlenecks to regional integration strategy and programs. These are mostly program oriented although regional integration, vulnerability and fragility issues were analysed. The monitoring results matrix was revised at MTR and CR with adjustment of the pillars to the new regional challenges and increased political and security issues besides the Ebola crisis which the Bank took quickly on board and responded to the crisis. This was clarified at RISP CR which offered a candid assessment of results and identified options for the new strategic orientation of the RISP while raising an important issue related to the dual accountability and multiplicity of mechanisms related to the regional integration architecture. Strategic actions to remove vulnerability and fragility of the affected countries were included at RISP-MTR (extended by 2 years) and RISP CR which was independently validated by IDEV. Although it was a pilot exercise, the validation revealed several deficiencies and discrepancies between RISP SESP in terms of results and portfolio management, focusing on three main evaluation criteria such as: relevance, effectiveness and efficiency. This is seen as a worth exercise that would enhance the candor and rigor in assessing performance at the strategic, portfolio management, capacity building and partnership levels and in informing the new RISP. Lessons and recommendations in both RISP SESP and independent validation offered an opportunity to enhance the evaluability of the new RISP and reflect on new measurement and monitoring framework for results achievement. |

| 1- **PERFORMANCE MANAGEMENT** | Narrative Assessment |
|---|---|
| 1. (a) Is the RISP considered a strategy document or more of a programmatic document for regional projects; or both?<br>(b) When evaluating the RISP, what weight do you give to its different components: strategy vs. portfolio?<br><br>2. RISPs can cover several CSPs over more than one cycle, and a portfolio of projects with various durations. What are the main criteria used in assessing the performance of the RISP through its main outputs: MTR, RPPR, CPPR and the CR?<br><br>3. Are the RISP SESP outputs (MTR-RPPR-CR) aligned with:<br>• Main operational policy documents?<br>• High 5s?<br>• Environmental and | The WA-RISP rested on two pillars, namely (i) linking regional markets and, (ii) building capacity for effective implementation of the regional integration agenda. This strategy was aligned with the ECOWAS Vision 2020, the Regional Strategic Plan, and the outcome of consultations with regional stakeholders. In coordination with the Africa Action Plan (AAP) of the African Union (AU), ECOWAS led and coordinated implementation of the New Partnership for Africa's Development (NEPAD) programs in West Africa, including the Comprehensive African Agriculture Development Program (CAADP) and the Program for Infrastructure Development in Africa (PIDA). PIDA, which covered cross-border infrastructure investment needs and policy and other regulatory measures to accompany these investments up to 2040, approved by the AU Heads of State Assembly in January 2012.<br><br>The RISP SESP provided the basis for the analysis of regional integration strategies' implementation, political and security risks related to performance and results. The RISP was rather a programmatic document despite the analysis of regional integration and bottlenecks with a high number of fragile states for which the Bank conducted a fragility and vulnerability assessment mainly for the Sahel region. The RISP-MTR and CR made an attempt to align the regional integration strategic objectives with the country.<br>As specified in the ADF-12 paper, the five-year Regional Integration Strategy Paper should have contained an indicative pipeline of operations recommended for Bank Group support with consideration of (i) Operations‟ alignment with the Bank's corporate priorities and the Regional Integration Strategy; (ii) Operations' alignment with the priorities of the African Union, NEPAD, RECs and Regional Member Countries; (iii) Operations' impact on regional integration; and ownership of participating countries and entities. The prioritization framework (Scorecard) did not include indicators related to the participating countries (CPIA, portfolio performances of regional operations from APPR, countries trade facilitation policy commitments to regional integration) and quality at entry and readiness (existence of MOU agreement between beneficiary countries, evidence of support from relevant |

| | |
|---|---|
| social safeguards such as: gender, fragility, climate change?<br>• Fiduciary & governance policy?<br><br>4. To what extent have alignment and/or complementarity between RISPs and CSPs been dealt with as some projects faced challenges of ownership and cross-country coordination since regional operations are typically implemented at national levels?<br><br>5. RPPRs and CPPRs covered by the RISP include a rating on a scale of 1 to 4, which is based on a simple average of reported performance of public sector operations. How well were other non-lending activities considered as part of the RISP SESP?<br><br>6. Is the recent decision to separate the regional diagnostic note (RDN) from the RISP per se, likely to facilitate the alignment with the Bank's key strategic documents and the robustness of the RISP self-evaluation process in the MTR and the CR?<br><br>7. (a) Should SESP outputs of the RISP cover both the RDN and the RISP document or only the latter?<br>(b) How could the RISP- | RECs, and collaboration/co-financing with development partners) which was revealed as a serious impediment to the smooth implementation of the soft components of the regional integration strategy.<br><br>New guidelines did not make it clear which from CSP or RISP is the main strategic document or the subsidiary and how they complement each other. Both regional integration coordinators and RECs implementing agencies were not conversant with M&E and impact tools of regional projects and were able to lead regional diagnostic and understanding of the political economy of the regional integration. However, as security and fragility were the main issues to further regional integration, adding the Ebola virus crisis, priorities in the regional integration strategy have changed leading to adjustment of the 2 main pillars).<br><br>With the recent guidelines on regional diagnostic note (RDN) preceding the RISP preparation would, according to regional coordinator, facilitate the understanding of the strategic components of the regional integration strategy (policy convergence, spatial integration, capacity strengthening of the RECS and regional institutions to support policy development and harmonized macro-economic management, etc..). However, RISP SESP should have reflected on strategy implementation issues related to quality at entry and RECs capacity to manage regional projects. For this reason, the bank took several actions to relieve WAEMU of the acquisitions and implementation of the construction and equipment sub-components of facilities and entrust these activities to States that have demonstrated ability to better implement road components of regional projects, for example. The other option was for WAEMU to continue to implement the regional components but to develop a pool of procurement experts to support acquisitions. With regard to the delays at the Bank's level, project coordinators and WAEMU were requested to be more proactive, through regular communication and monitoring of their files with project managers. The 2018 RPIP included a series of remedial actions to improve portfolio performance, quality at entry, procurement, disbursement and M&E with monitorable indicators, responsible entities and timeframe. Regional programs and multi-country project results' assessment can provide only a trend analysis of the regional integration strategy by building regional infrastructure to support transport and trade facilitation with weak RECs implementation capacity.<br><br>RISP CR offered a candid assessment of results and identified options for the new strategic orientation of the WA-RISP while raising an important issue related to the dual accountability and multiplicity of mechanisms related to the regional integration architecture. Strategic actions to remove vulnerability and fragility of the affected countries were included at RISP-MTR (extended by 2 years) and RISP CR, independently validated by IDEV. Although it was a pilot exercise, the validation revealed several deficiencies and discrepancies between RISP SESP in terms of results and portfolio management, focusing on three main evaluation criteria such as: relevance, effectiveness and efficiency. This is seen as a worth exercise that would enhance the candor and rigor in assessing performance at the strategic, portfolio management, capacity building and partnership levels and in informing the new RISP. Recommendations in both RISP SESP and independent validation offered an |

| | |
|---|---|
| MTR and RISP-CR better reflect the performance at the regional level and country components?<br><br>8. Are the SESP main outputs (RISP-MTR, CR, Annual Portfolio Reviews) geared towards addressing:<br>(a) strategic issues,<br>(b) policy dialogue,<br>(c) knowledge management?<br><br>9. Was IDEV validation of the projects PCRs, CSPs and RISP-Es factored-in when discussing the performance of the regional portfolio and of the RISP? Does external validation help improve the candour of the SESP? Should the RISP-CR be rated? | opportunity to enhance the evaluability of the new RISP and reflect on new monitoring framework for results achievement.<br><br>RISP were not based on clear theory of change (TOC) both on operational and strategic level in countries or region. RISP Results framework did not inform on regional integration progress nor on macroeconomic convergence. This was highlighted in RISP-CR and the pilot validation exercise.<br><br>RECs and implementation agencies and beneficiaries' involved in RISP SESP preparation did not happen which was seen as a major impediment to WA-regional integration. However, capacity development of RECs (including in statistics and impacts assessments) was necessary to enhance the quality of design and follow up of strategic development objectives of the region but were not significantly assessed. Independent validation of RISP-CR, RPPR and RPIPs was considered as a good way to ensure candor, objectivity and validity of self-assessments in RISP-MTR and CR.<br><br>Portfolio performance indicators were identified and reviewed during the RISP period with slight improvements in most of regional and multi-country operations (reduction of potentially problematic projects and projects as well as commitment at risks). Portfolio rating improved from 2.07 in 2013 (on a scale 1-3, equivalent to 2.76 on a scale 1-4) to 3.26 in 2018. As at August 1, 2018 only one project was classified as problematic, down from two identified during the mid-term in 2014, while the commitment at risk reduced drastically from 16 percent to about 1 percent. Disbursement rate has improved between 2014, 2016 and 2018 from 41%, 33.6% to 47.2% during the period.<br><br>RISP SESP should also report on processes to deliver sector results on social and environmental safeguards. These processes should be described and documented at inception (early engagement of safeguards during post approval) with adequate risk assessment.<br><br>De-linking RPPRs and RPIPs from RISP MTR or CR may allow to focus more on strategic issues and few on operational issues of strategic importance. The flagship highlights are insufficient to deal with both implementation and development issues at the regional level. RISPs SESP in the region with a high number of fragile states with security and political vulnerability issues lacked valid and reliable data on implementation and results due to conflicts and unavailability of data.<br><br>The implementation support division may improve the situation but there is need to share findings for problem anticipation and timely corrections.<br><br>Disregarding numerical ratings in assessing implementation performances as well as Bank and Governments or RECs performance may encourage better dialogue focusing on strategic objectives, according to regional economists and regional coordinators. Consultations/coordination with implementing agencies in the region did happen but their involvement should be enhanced to increase the WA-RISP SESP validity. |

**Performance Management Overall Assessment**

Portfolio implementation challenges in West Africa include: (i) delay in the implementation of loan agreements, (ii) delay in the fulfilment of conditions precedent to first disbursement, (iii) insufficient knowledge of the Bank's rules and procedures, particularly with regard to procurement and disbursements, and (iv) weak implementation capacity. The RISP SESP provided the basis for the analysis of regional integration strategies' implementation risks related to performance and results. However, the RISP was rather a programmatic document despite the analysis of regional integration and bottlenecks with a high number of fragile states.
New guidelines did not make it clear which from CSP or RISP is the main strategic document or the subsidiary and how they complement each other. The pilot independent validation exercise, based on a desk review, assessed the quality of evidence provided to support the self-assessment and contribution to the new Regional Integration Strategy. The validation aimed at contributing to improving the quality of RISP completion report and providing an opportunity for learning. The validation covered three dimensions of the OECD/DAC evaluation criteria: relevance, effectiveness and efficiency. The validation also assessed the Bank's performance in managing risks, policy dialogue, partnership and aid coordination.
The validation concluded that the RISP-CR lacked sufficient evidence to support its main conclusion and did not report in details on the achievement of development objectives for the 2011-2017 RISP period.

The RISP-CR should have reported on results up to 2018. It was not clearly linked to specific regional operations. Although the RISP aimed to mainstream cross cutting issues in the implementation, this was not taken into consideration in the results framework with only scant details provided in the RISP-CR with regard results achievements. The assessment of the efficiency of the RISP-CR implementation was not rigorous, although too much delays were noted. The RISP-CR did not explicitly assess the likely contribution or attribution of these delays to the Bank or RECs or governments. Finally, the RISP-MTR and CR did not assess the extent to which the RISP implementation achieved its ultimate goals of advancing the regional integration agenda and facilitating regional solutions, according to the independent validation.

| **2- Accountability** | Narrative assessment |
|---|---|
| 1- (a) are roles and responsibilities sufficiently clear in the preparation, conduct, review, sign-off, follow-up for the various steps of the RISP SESP outputs (country team, regional team, country managers, dg and chief economist)?<br>2- do regional dg and country managers see the RISP SESP outputs as a relevant accountability tool?<br>3- do SESP of the RISP outputs provide a relevant perspective on the results achieved and communicate overall performance in a credible way? | Regional Directorate plays a fundamental role in adopting an operational approach to RISPs and RISP SESP by engaging ADB staff in adapting country specific strategies to the regional context and enhancing Bank's agility and ability to include urgent security and vulnerability crisis such as the Ebola virus crisis which was expanding rapidly in the region.<br><br>The lack of staff in sector divisions did not allow for continued strategic and operational dialogue and for increased supervision and just-in time actions to address regional project or multi-country projects' implementation issues. The SESP at regional integration project level is restricted to operational issues and only mid-term reviews provide an opportunity to reflect on the adequacy of Bank's response to the regional integration issues, challenges and bottlenecks. RISP SESP are thus considered as both for accountability and for knowledge base of it entails in promoting regional integration. This did not happen in West Africa RISP and was highlighted during the 2014 and after consultation workshops with ECOWAS and WAEMU and other country officials.<br><br>Collaboration with RECs proved to be difficult without strong capacity building in terms of policy convergence, harmonization of macro-economic |

| | |
|---|---|
| 4- is regional and HQ management exerting leadership over the correct implementation of the RISP SESP outputs and lines of accountability?<br>5- are TMs of regional programs being held accountable for timely submission and proper implementation of the RISP SESP outputs?<br>6- to what extent do SESP outputs of regional program: (a) rely on strong M&E systems? (b) offer a credible accountability framework?<br>7- to what extent do managers discuss with TMs regional program performance and results during staff performance conversations?<br>8- does SESP aggregation of RISP products provide a relevant and cost-effective reporting of results through the RMF, ADER and other reporting tools (dashboard, etc..)? | policies and a good understanding of the political economy of how regional integration is built. The pilot validation exercise of the RISP-CR conducted in much 2019 stated that the actual disbursement rate for the public sector (representing 78 out of the 86 operations) is only 23% suggesting that disbursement performance for public sector regional projects is less than satisfactory. The Bank Group regional portfolio reflected the economic development challenges in the WA region, especially the recent efforts of the Bank Group in addressing the Ebola Virus crisis as captured by the social sector operations, as well as the transportation and energy sector. Private sector operations have also been important. PCREN were fed to staff but not considered for official ratings in RPPRs. The results monitoring framework did not cascade to allow meaningful IPR reports on regional programs and results matrices were mostly not adapted to country and region context. Results frameworks lack smart indicators mainly on the social sector (poverty assessment, vulnerability, gender disparities, etc..).<br><br>The RISP-MTR was candid in highlighting the main reasons for insufficient performance of the Bank: inefficient supervision, highly centralized decision-making process due to the relocation of the Bank in Tunis, lack of real time technical assistance to RECs and insufficient familiarization of implementing agencies on Bank's rules and procedures. Governments and RECs performance were judged less that satisfactory due to the weak participation of RECs in the design of the operations, lack of monitoring and evaluation systems, delays in selecting the implementation team and the lack of delegation of authorities. Governments performance is hampered by the lack of coordination during the multi-country and regional integration operations and insufficient consultations before multi-country project start-ups and the lack of dissemination and sharing of information on the regional projects.<br><br>Risk factors' assessment in RISP-MTR and CR was related to the dual accountability and multiplicity of mechanisms of the integration architecture (Bank, RECs and Government levels) and the political fragility in the WA region which have led to the identification of lessons learned and strategic options to enhance regional integration strategy implementation, M&E of results and RECs capacity building mainly by increasing their familiarization of implementing agencies on Bank's rules and procedures, sharing knowledge products on the WA region and dissemination of the regional research centers on policy convergence, regional macroeconomic policies and the political economy management.<br><br>The extension of the 2011-2015 RISP to 2017 highlighted several issues such as the political and security vulnerability and fragility of the region. Bank Group's fragility assessments provided critical inputs necessary to incorporate fragility-mitigating actions into the design and implementation of regional integration and country strategies (for Mali, Niger, Togo, Guinea Bissau, Sierra Leone and political economy analysis for Côte d'Ivoire and Burkina-Faso) and implemented resilience strengthening program in the Sahel<br>The regional directorate (and also the RDVP) were exerting leadership for the implementation of the regional program both on hard and sift components but due to the relocation of the bank in Tunis, centralization of the decision making process posed several issues for timely corrective actions. |

| | |
|---|---|
| | Discrepancies between RISP SESP and independent validation undertaken by BDEV highlighted the need for more candor and rigorous assessments of effectiveness and efficiency and effective M&E systems within the implementing agencies and RECS. Lessons and recommendations in both RISP SESP and independent pilot validation exercise offered an opportunity to enhance the evaluability of the new RISP with M&E approach to apply measurement and monitoring framework for results achievement. It has suggested that capacity building is of utmost importance and should continue to be a priority as illustrated by the delays in implementing regional operations.<br><br>Countries and RECs should be involved in SESP in reporting on regional difficulties and bottlenecks or challenges as well as on opportunities for enhanced policy convergence, macroeconomic management, and the political economy of regional integration in West Africa with security and political vulnerability issues.<br><br>Regional coordinators for E&S safeguards and climate change are few and cannot cover all countries and regions. They are not systematically involved in field missions or meetings of RISP meetings although they are custodian of the Bank's climate change action plan (2011-2015). Climate change, social and environmental safeguards should be used as support team to project and RISP preparation and implementation and involved in early stage of project/RISP design (pre-board stages) for better recognition of the subject matter.<br><br>Few gender specialists are located in regions. Gender policy guidelines have to be reviewed as no clear plan of action is in place to report on results in SESP. No dedicated budget is in place to cover specifically gender issues during regional programs' preparation, negotiation down to completion.<br><br>As the WA-RISP, the institutional framework should be reviewed to enhance reporting on results and integration into the ADER and RMF with regional experts' involvement in RISP assessments, policy dialogue, validations and corrective actions mainly of safeguards (example compensation in resettlements) and gender. The implementation support divisions should be well staffed with all subject matter experts such as gender, climate change, E&S experts, etc…<br><br>Regional Directorate plays a fundamental role in adopting an operational approach to RISPs and RISP SESP by engaging ADB staff in adapting country specific strategies to the regional context and enhancing Bank's agility and ability to include urgent security and vulnerability crisis such as the Ebola virus crisis which was expanding rapidly in the region.<br><br>The lack of staff in sector divisions did not allow for continued strategic and operational dialogue and for increased supervision and just-in time actions to address regional project or multi-country projects' implementation issues. The SESP at regional integration project level is restricted to operational issues and only mid-term reviews provide an opportunity to reflect on the |

| | |
|---|---|
| | adequacy of Bank's response to the regional integration issues, challenges and bottlenecks. RISP SESP are thus considered as both for accountability and for knowledge base of it entails in promoting regional integration. This did not happen in West Africa RISP and was highlighted during the 2014 and after consultation workshops with ECOWAS and WAEMU and other country officials.<br><br>Collaboration with RECs proved to be difficult without strong capacity building in terms of policy convergence, harmonization of macro-economic policies and a good understanding of the political economy of how regional integration is built. The pilot validation exercise of the RISP-CR conducted in much 2019 stated that the actual disbursement rate for the public sector (representing 78 out of the 86 operations) is only 23% suggesting that disbursement performance for public sector regional projects is less than satisfactory. The Bank Group regional portfolio reflected the economic development challenges in the WA region, especially the recent efforts of the Bank Group in addressing the Ebola Virus crisis as captured by the social sector operations, as well as the transportation and energy sector. Private sector operations have also been important. PCREN were fed to staff but not considered for official ratings in RPPRs. The results monitoring framework did not cascade to allow meaningful IPR reports on regional programs and results matrices were mostly not adapted to country and region context. Results frameworks lack smart indicators mainly on the social sector (poverty assessment, vulnerability, gender disparities, etc..).<br><br>The RISP-MTR was candid in highlighting the main reasons for insufficient performance of the Bank: inefficient supervision, highly centralized decision-making process due to the relocation of the Bank in Tunis, lack of real time technical assistance to RECs and insufficient familiarization of implementing agencies on Bank's rules and procedures. Governments and RECs performance were judged less that satisfactory due to the weak participation of RECs in the design of the operations, lack of monitoring and evaluation systems, delays in selecting the implementation team and the lack of delegation of authorities. Governments performance is hampered by the lack of coordination during the multi-country and regional integration operations and insufficient consultations before multi-country project start-ups and the lack of dissemination and sharing of information on the regional projects.<br><br>Risk factors' assessment in RISP-MTR and CR was related to the dual accountability and multiplicity of mechanisms of the integration architecture (Bank, RECs and Government levels) and the political fragility in the WA region which have led to the identification of lessons learned and strategic options to enhance regional integration strategy implementation, M&E of results and RECs capacity building mainly by increasing their familiarization of implementing agencies on Bank's rules and procedures, sharing knowledge products on the WA region and dissemination of the regional research centers on policy convergence, regional macroeconomic policies and the political economy management.<br><br>The extension of the 2011-2015 RISP to 2017 highlighted several issues such as the political and security vulnerability and fragility of the region. Bank |

| | |
|---|---|
| | Group's fragility assessments provided critical inputs necessary to incorporate fragility-mitigating actions into the design and implementation of regional integration and country strategies (for Mali, Niger, Togo, Guinea Bissau, Sierra Leone and political economy analysis for Côte d'Ivoire and Burkina-Faso) and implemented resilience strengthening program in the Sahel<br>The regional directorate (and also the RDVP) were exerting leadership for the implementation of the regional program both on hard and sift components but due to the relocation of the bank in Tunis, centralization of the decision making process posed several issues for timely corrective actions.<br><br>Discrepancies between RISP SESP and independent validation undertaken by BDEV highlighted the need for more candor and rigorous assessments of effectiveness and efficiency and effective M&E systems within the implementing agencies and RECS. Lessons and recommendations in both RISP SESP and independent pilot validation exercise offered an opportunity to enhance the evaluability of the new RISP with M&E approach to apply measurement and monitoring framework for results achievement. It has suggested that capacity building is of utmost importance and should continue to be a priority as illustrated by the delays in implementing regional operations.<br><br>Countries and RECs should be involved in SESP in reporting on regional difficulties and bottlenecks or challenges as well as on opportunities for enhanced policy convergence, macroeconomic management, and the political economy of regional integration in West Africa with security and political vulnerability issues.<br><br>Regional coordinators for E&S safeguards and climate change are few and cannot cover all countries and regions. They are not systematically involved in field missions or meetings of RISP meetings although they are custodian of the Bank's climate change action plan (2011-2015). Climate change, social and environmental safeguards should be used as support team to project and RISP preparation and implementation and involved in early stage of project/RISP design (pre-board stages) for better recognition of the subject matter.<br><br>Few gender specialists are located in regions. Gender policy guidelines have to be reviewed as no clear plan of action is in place to report on results in SESP. No dedicated budget is in place to cover specifically gender issues during regional programs' preparation, negotiation down to completion.<br><br>As the WA-RISP, the institutional framework should be reviewed to enhance reporting on results and integration into the ADER and RMF with regional experts' involvement in RISP assessments, policy dialogue, validations and corrective actions mainly of safeguards (example compensation in resettlements) and gender. The implementation support divisions should be well staffed with all subject matter experts such as gender, climate change, E&S experts, etc… |
| **Accountability Overall** | Regional Directorate plays a fundamental role in adopting an operational |

| | |
|---|---|
| **Assessment** | approach to RISPs and RISP SESP by engaging ADB staff in adapting country specific strategies to the regional context and promoted their ability and agility to face the regional crisis of the Ebola virus. The regional delivery unit would help assist the regional directorate in reviewing what has worked or not in terms of regional strategic implementation, achievements and results including the risk factors around them.<br><br>Involving all actors, including research and excellence centres on policy convergence, macroeconomic management and political economy would increase their validity and reliability.<br><br>Collaboration with RECs proved to be difficult without strong capacity building in terms of policy convergence, harmonization of macro-economic policies and a good understanding of the political economy of how regional integration is built.<br><br>The specificity of the Western African region with its dual accountability between ECOWAS and WAEMU, the complexity of mechanisms around the WA-regional integration agenda and the variety of regional institutions in the Sahel and West Africa limited the scope of the 2011-2017 RISP to the weak capacity of these institutions and the complexity of how WA-regional integration should entail. The lack of effective M&E systems within the implementing agencies and RECS limited the assessment of results, and their attribution.<br><br>Discrepancies between RISP SESP and independent validation undertaken by BDEV highlighted the need for more candor and rigorous assessments of effectiveness and efficiency and effective M&E systems within the implementing agencies and RECS. Lessons and recommendations in both RISP SESP and independent pilot validation exercise offered an opportunity to enhance the evaluability of the new RISP with M&E approach to apply measurement and monitoring framework for results achievement. It has suggested that capacity building is of utmost importance and should continue to be a priority as illustrated by the delays in implementing regional operations. |
| **3. Learning** | Narrative assessment |
| 1. To what extent have the RISP SESP outputs for regional program been used as a source of learning and knowledge management to inform the next RISP or other regional operations/strategies<br><br>2. To what extent has the RPPR or CPPRs been used for the preparation of the RISP- MTR and | The involvement of regional economists and regional coordinators in the design and implementation phases of RISPs contributes mainly to enhance the knowledge sharing on what regional integration entails. This was corrected by the new CSP/RISP guidelines.<br><br>Training of TMs and WA-implementing agencies on integrating regional integration issues, M&E, procurement and Bank rules and regulations on E&S safeguards should be continuous due of the implementing staff. There is need to increase learning from RISP case studies on E&S safeguards and CC adaptation and also on SDG and mainstreaming guidelines of S&E should reach out countries and RECs.<br><br>New RDN (as a separate document from the RISP) was judged as a good opportunity to focus more on the challenges and opportunities for a strong |

| | |
|---|---|
| the RISP-CR?<br><br>3. (a) Did feedback loop sessions and lessons learned discussions take place after RISP- MTR or RISP CR? (b) Have these feedback sessions led to better procedures for regional programs, restructuring, differentiation according to regional situations (fragile context, MIC, non-lending)<br><br>4. To what extent are recommendations and lessons learned described in RISP-MTR and RISP-CR: (a) actionable and useful? (b) Has any of it been used?<br><br>5. What incentives could change behaviours in terms of promoting critical analysis, best practices of results reporting, follow up, awards for innovation, increase of value of the knowledge created, learn from failure?<br><br>6. To what extent will the new RDN (as a separate document from the RISP) strengthen the Bank's analysis of regional integration issues and their linkages with national issues, leading to better alignment of the Bank's country and regional operational programmes | regional integration that would strengthen the Bank's analysis of regional integration issues, policy convergence, macroeconomic management and on political economy of regional integration in West Africa. Relying on regional excellence centres provides an opportunity to learn and share knowledge on regional integration, leading to better alignment of the Bank's country and regional operational programmes.<br><br>The feedback loop from RISP SESP that provide success-failure stories and results based on assessment of achievements in reaching regional integration should be encouraged with more opportunities for consultation and knowledge sharing and dissemination. Lessons learned should be institutionalized during life cycle and documented for capitalization of lessons learned in RISPs. BDEV pilot validation proved to be a valuable exercise as it has provided an independent assessment of the RISP-CR findings and contributed to their validity. However, as a pilot exercise, with its shortcomings should be encouraged and enlarged to include an independent assessment of the theory of change, transformation and increased convergence and coherence with country CSPs. Validations should also include lessons and generating knowledge that will enable new strategic orientations to the WA-Africa region with regard the political fragility and vulnerability including cross-cutting issues such as climate change, gender and regional disparities. There is need to increase the number of knowledgeable regional experts to build up strategic cooperation framework at regional level particularly with Private investors in building integrated regional infrastructure and with research centers in the region. |

| | |
|---|---|
| 7. Are there concerns over ratings and disconnects between RISP-CR and RISP-E provided by IDEV that could distract from learning?<br><br>8. To what extent were leadership signals received that learning and knowledge management are key outcome of the RISP's SESP. | |

**Learning Overall Assessment**

The involvement of regional economists and regional coordinators in the design and implementation phases of RISPs contributes mainly to enhance the knowledge sharing on what regional integration entails. This was corrected by the new CSP/RISP guidelines.

There is need to increase learning from RISP case studies on political convergence, macroeconomic management and the political economy of the WA regional integration as well as for E&S safeguards and CC adaptation and also on SDG and mainstreaming guidelines of S&E should reach out countries and RECs.

New RDN (as a separate document from the RISP) was judged as a good opportunity to focus more on the challenges and opportunities for a strong regional integration that would strengthen the Bank's analysis of regional integration issues and their linkages with national issues, leading to better alignment of the Bank's country and regional operational programmes.

The feedback loop from RISP SESP on success-failure stories and results based on assessment of achievements in reaching regional integration was eagerly requested by regional staff. IDEV regional strategy program evaluation helped in capturing lessons and generating knowledge that enabled new strategic orientations notably of cross-cutting issues.

**CONCLUSIONS**

The RISP was underpinned by the Strategic Frameworks of both ECOWAS and WAEMU. It has provided support for both institutions but due to capacity of institutions the Bank has increased its supervision on multi-country operations directly managed by these institutions. The RISP SESP provided the basis for the analysis of regional integration strategies' implementation, political and security risks related to performance and results. The RISP was rather a programmatic document despite the analysis of regional integration and bottlenecks with a high number of fragile states for which the Bank conducted a fragility and vulnerability assessment mainly for the Sahel region. The RISP-MTR and CR made an attempt to align the regional integration strategic objectives with the country CSPs.

The pilot independent validation exercise, based on a desk review, assessed the quality of evidence provided to support to the self-assessment. The validation, covering three dimensions of the OECD/DAC evaluation criteria: relevance, effectiveness and efficiency concluded that the RISP-CR lacked sufficient evidence to support its main conclusion.

The new RISP/CSP guidelines did not make it clear which from CSP or RISP is the main strategic document or the subsidiary and how they complement each other, according to regional coordinators. Reporting on supported actions to promote regional integration in RISP SESP was considered as of utmost importance to reflect on strategy implementation.

The specificity of the Western African region with its dual accountability and variety of mechanisms of regional integration and weak capacity regional institutions have increased the complexity of how WA regional integration should entail.

The lack of effective M&E systems within the implementing agencies and RECs limited the assessment of results, and their attribution.

Lessons and recommendations in both RISP SESP and independent evaluation offered an opportunity to enhance the **evaluability** of the new RISP with M&E approach to apply measurement and monitoring framework for results achievement.

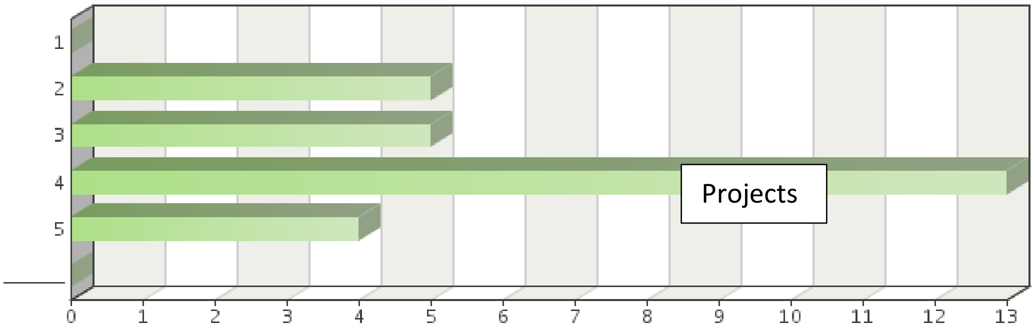
New RDN (as a separate document from the RISP) was judged as a good opportunity to focus more on the challenges and opportunities for a strong regional integration. The feedback loop from RISP SESP on success-failure stories and results based on assessment of achievements in reaching regional integration was eagerly requested by regional staff. BDEV pilot validation may contribute in capturing lessons and generating knowledge that will enable new strategic orientations notably of cross-cutting issues.

De-linking RPPRs from RISP MTR or CR may allow to focus more on strategic issues and few on operational issues of strategic importance.

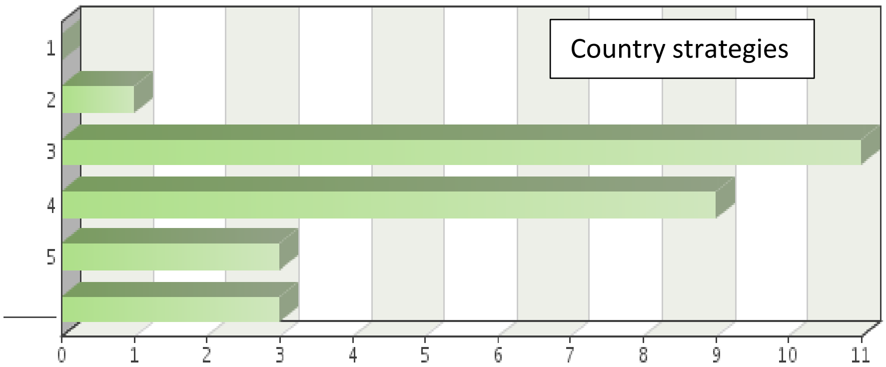
## Annex 8: Staff Survey Results

1. there is evidence that the implementation of the SES contributes to enhancing the design and performance of:

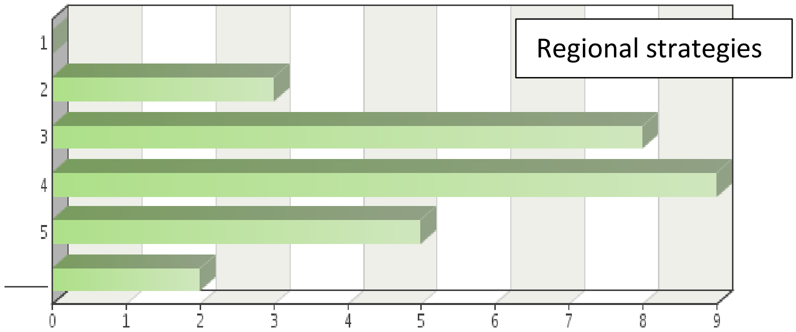
1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly Agree



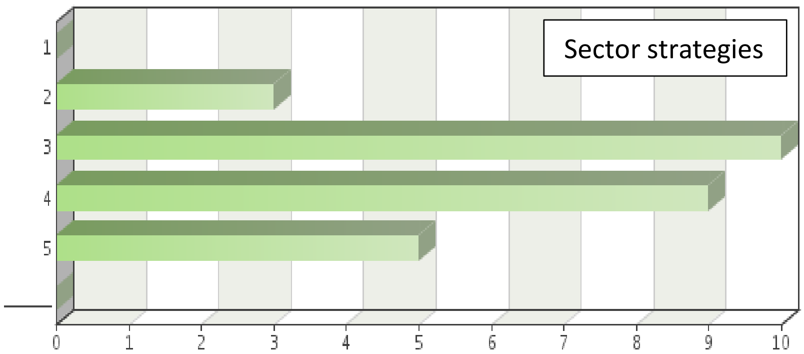

1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly Agree
6. NA




1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly Agree
6. NA

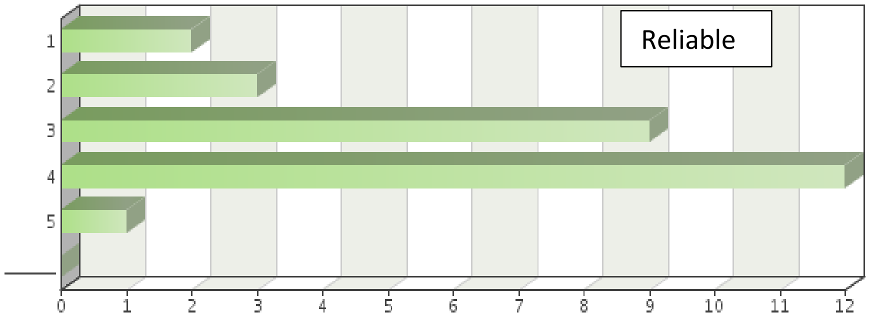



1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly Agree

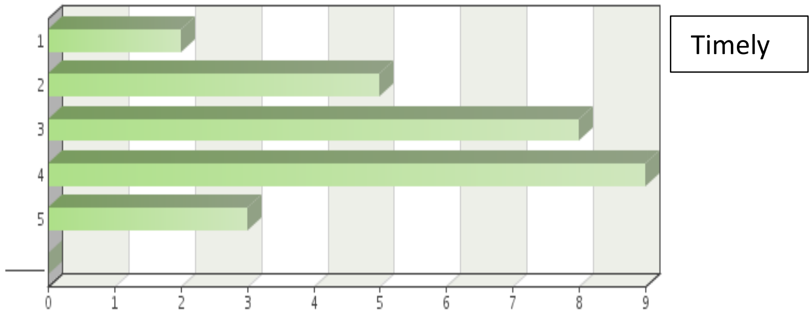


2. the SESP and its outputs provide a framework for portfolio management which is:

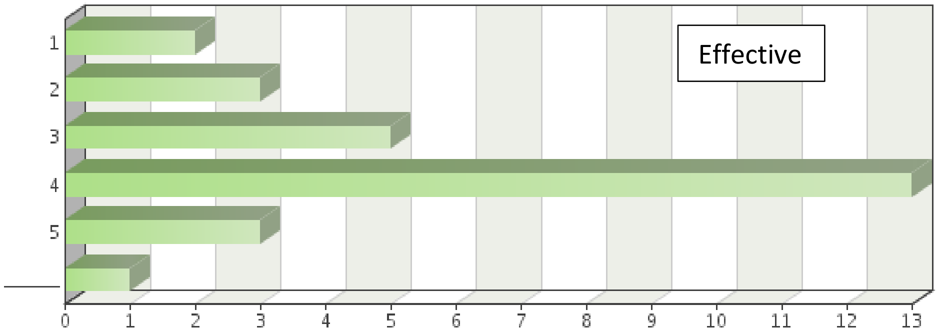
1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly Agree




1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly Agree

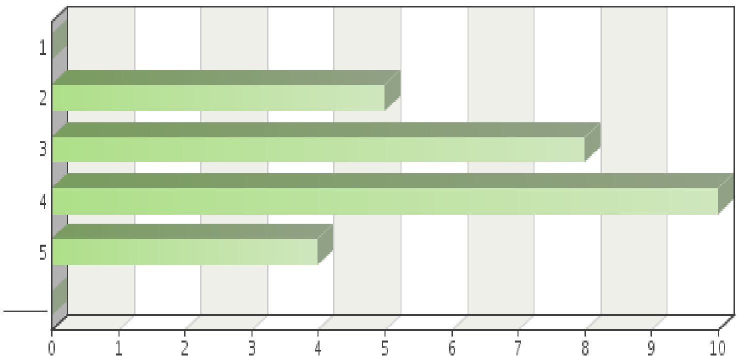



1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly Agree



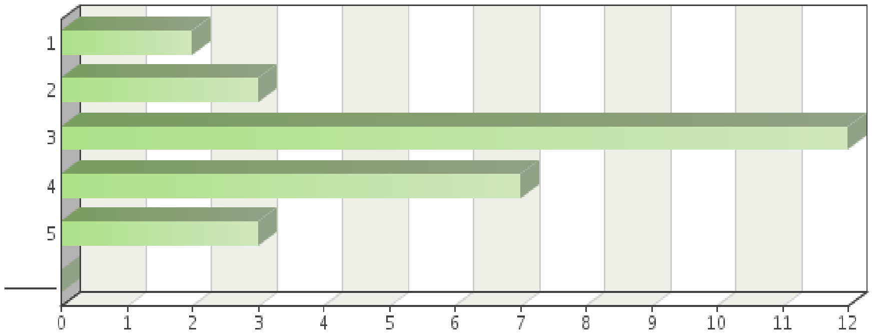

3. the SESP contribute to improved quality at exit (closing) through corrective action during supervision.

1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly Agree

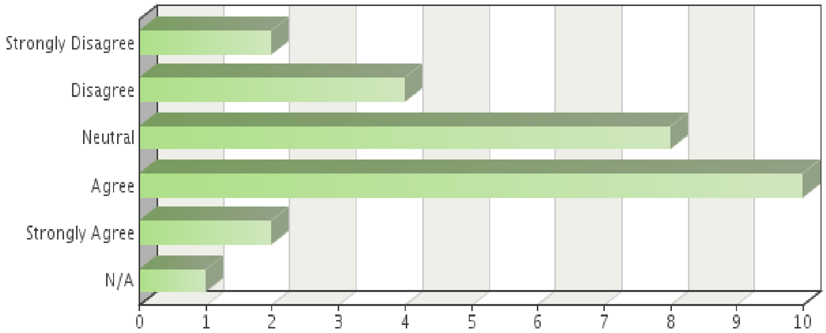


| 4. the SESP contribute to improved quality at exit (closing) through corrective action during supervision. |
|---|

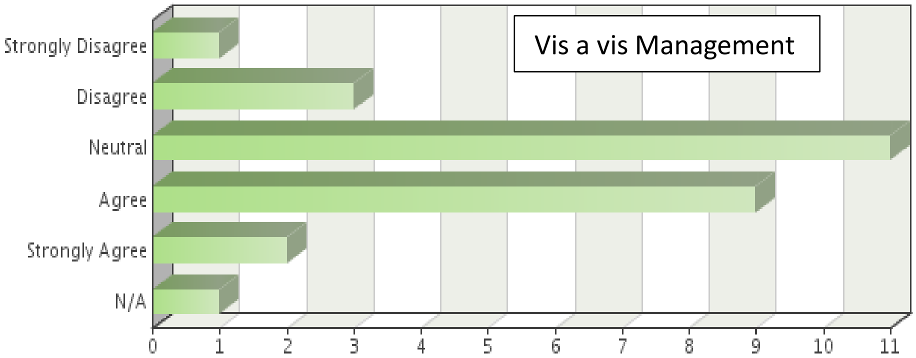
1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly Agree

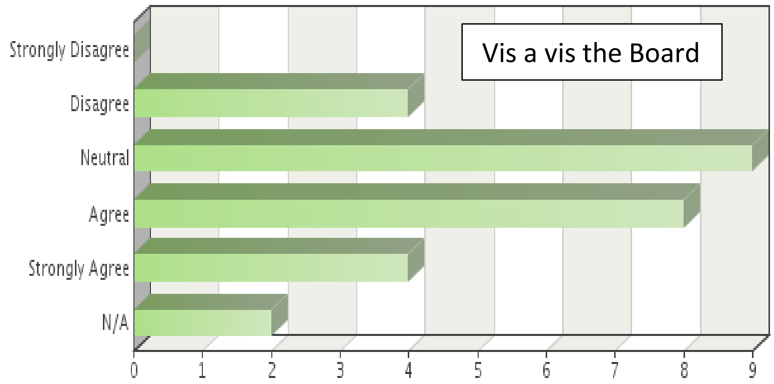



| 5. Management is sending and implementing the right signals as of the importance of the SESP as a tool to achieve better results |
|---|

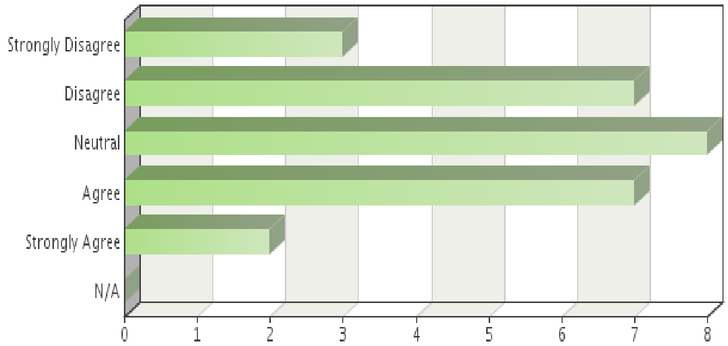



| 6. there is the right degree of accountability on the implementation of the SESP |
|---|

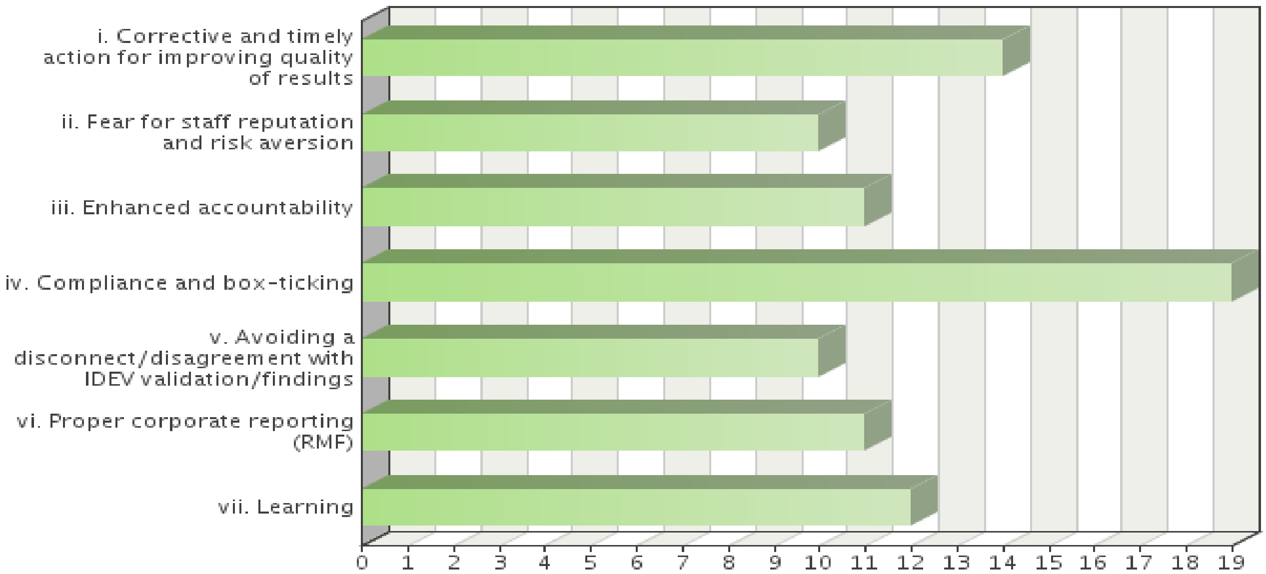





7. the SESP is seen as a tool to achieve results rather than a compliance mechanism

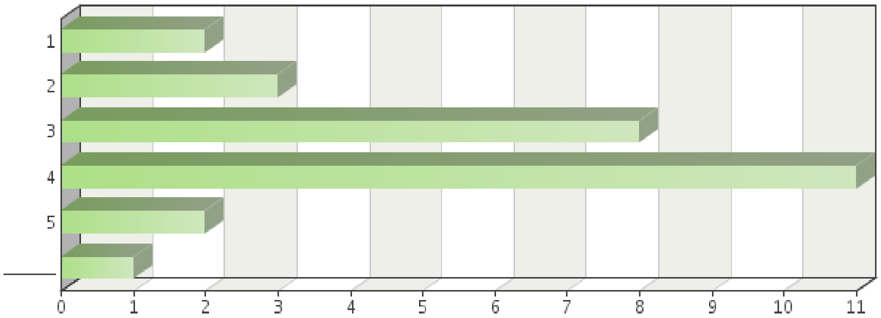



8. In which direction are the incentives around the SESP mostly influencing behavior (can tick more than one):



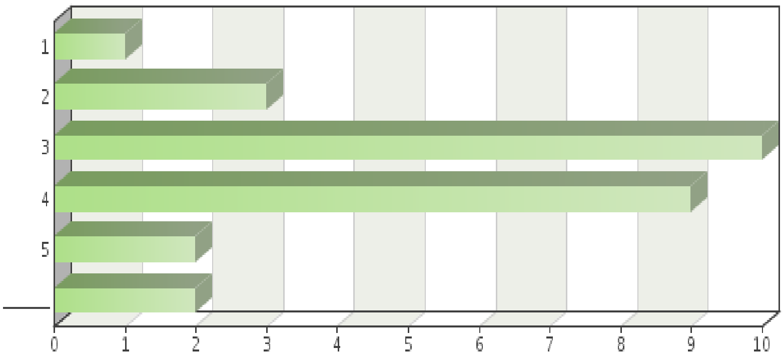

9. The outputs of the SESP are being used as an opportunity for learning and innovation

1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly Agree



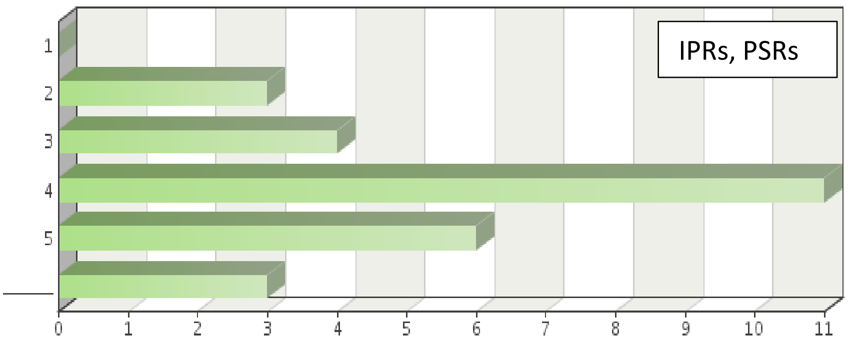

10. The lack of trust in the SESP undermines its learning potential

1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly Agree
6. NA




11. The SESP outputs listed below are the right vehicles for learning and lessons learned (please rate each)

1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly Agree
6. NA

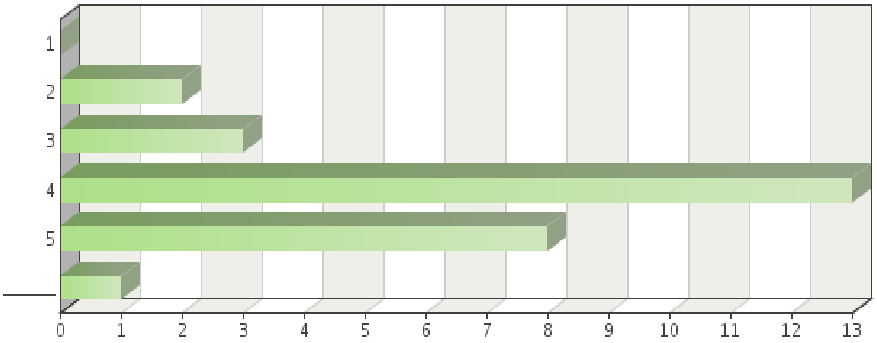


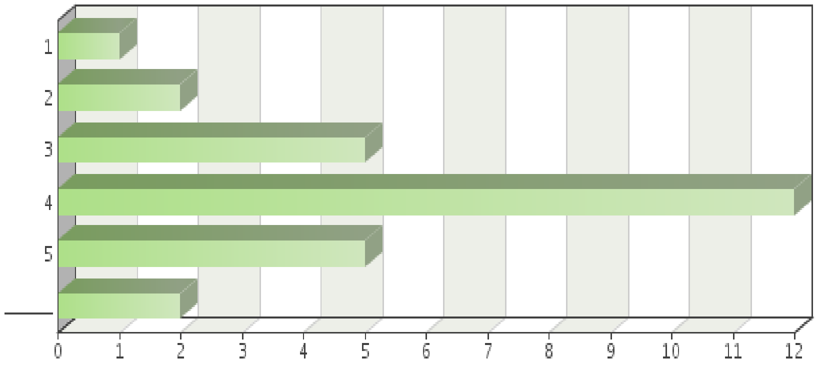

1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly

MTRs

1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly Agree
6. NA

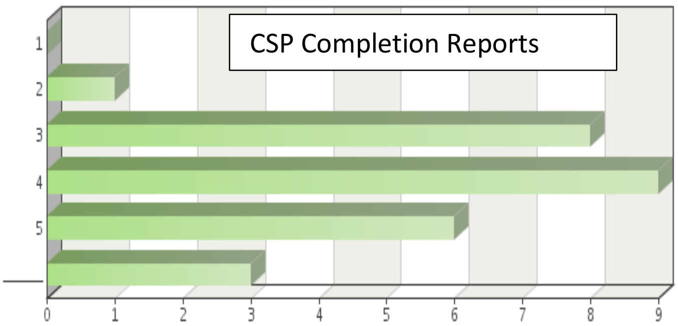
PCRs, XSRs



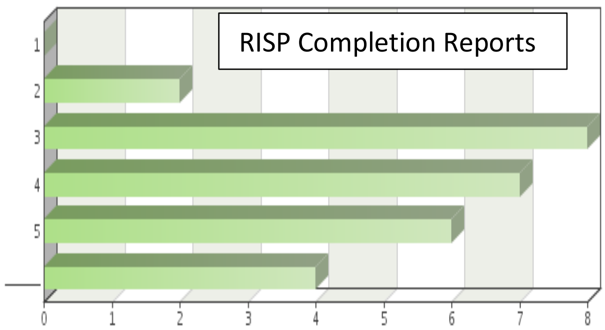

1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly Agree
6. NA

CSP Completion Reports




1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly Agree
6. NA

RISP Completion Reports



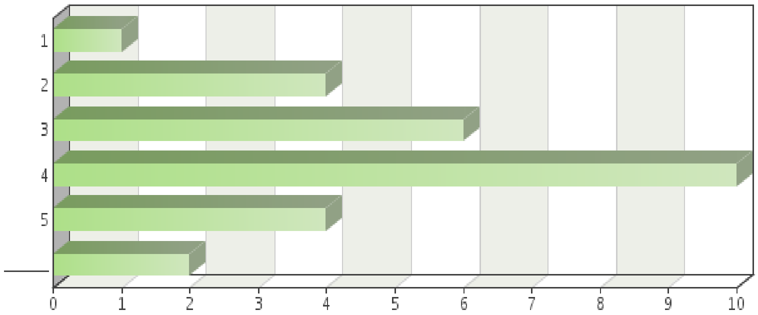

1. Strongly disagree
2. Disagree
3. Neutral
4. Agree

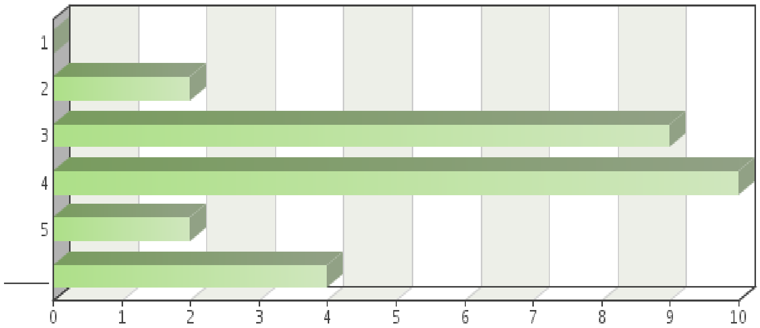
Thematic strategies and reviews

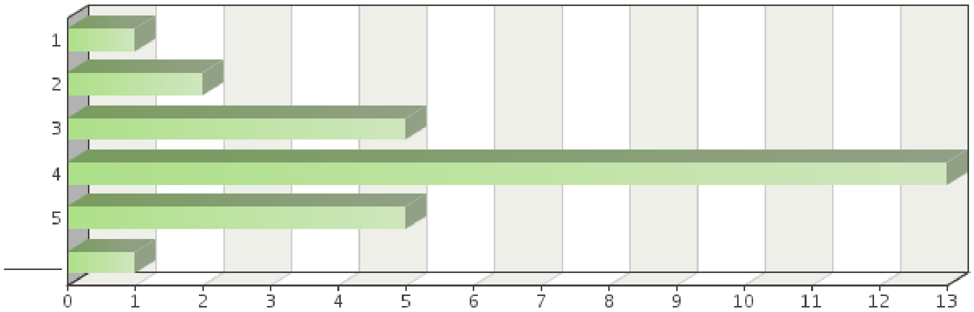
1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly Agree
6. NA



ADOA


1
2
3
4
5
0 1 2 3 4 5 6 7 8 9 10


1
2
3
4
5
0 1 2 3 4 5 6 7 8 9 10


1
2
3
4
5
0 1 2 3 4 5 6 7 8 9 10 11 12 13