

## Concept Note

# Results and Performance of the World Bank Group 2021

June 2, 2021

## 1. Introduction and Context

1.1 The Results and Performance of the World Bank Group (RAP) report is the annual review of evidence from IEG evaluation and validation work on the development effectiveness of the World Bank Group (WBG) - that is, the World Bank (International Bank for Reconstruction and Development and the International Development Association - IBRD/IDA), the International Finance Corporation (IFC), and the Multilateral Investment Guarantee Agency (MIGA). The RAP 2021 will be the eleventh in a series that began in 2010; it will also be the second report departing from the exclusive traditional focus on ratings to also provide additional evidence on the nature of intended outcomes<sup>1</sup> across the WBG.

1.2 The RAP 2020 (World Bank 2020) introduced an innovative focus on classification of intended outcomes. Part I of RAP 2020 synthesized IEG ratings and other evidence from IEG evaluations and validations to give an aggregated picture of the results and performance of the World Bank, IFC and MIGA. Part II of the report presented a new outcome classification, distinguishing between four outcome levels, ranging from outputs to early outcomes to intermediate outcomes to long-term outcomes.

1.3 RAP 2021 will build and expand on the RAP 2020 innovations by refining the classification framework for intended outcomes and integrating analysis of existing ratings (trends) with the outcome classification analysis. Like past RAP reports, RAP 2021 will provide an analysis of project ratings and factors associated with performance as measured by those ratings. Expanding on the past, RAP 2021 will analyze and interpret these ratings through the lens created by the refined *typology* of intended outcomes. This lens would enable an examination of ratings that takes into account portfolio composition in terms of the type (classification) of intended outcomes, as well as the likelihood of achieving those intended outcomes. In other words, RAP 2021 aims at providing a joint assessment of ratings and the risk-return profile of the portfolio generating those ratings.

---

<sup>1</sup> Intended development outcomes should be here broadly understood as intended goals or objectives of projects or programs (as opposed to an element within a results chain or theory of change or specific achievements as captured by, for example, IEG outcome ratings in ICRRs).

1.4 Considering the risk-return profile as part of the RAP analysis may enable reading the findings in the light of the COVID-19 pandemic. While the impact of COVID-19 on the results and performance of the institution will be only partially detectable in the sample of operations reviewed by RAP 2021, the findings on outcomes and risk is expected to provide useful insights regarding the implications that the pandemic may have for WBG operations.<sup>2</sup>

## **2. Objectives and Audience**

2.1 RAP 2021 aims at reaching out to a broad audience. The Board of Executive Directors, which was renewed in November 2020, is the most immediate audience. The WBG Management at all levels is another key audience.<sup>3</sup> At the same time, the report also caters to a more technical audience; for this reason (as well as for transparency), the report will be complemented by an IEG webpage including an online dashboard that will allow users to produce tabulations of the ratings. The website will also provide complementary materials to learn about definitions and background information of the analysis presented in the main report.

### **Report Scope and Questions**

2.2 The report covers WBG projects, country programs and corporate priorities. It will synthesize ratings and other IEG evidence of the outcomes and performance of World Bank, IFC, and MIGA projects and WBG country programs. It will refine, expand, and use as required the outcome classification of project outcomes developed by RAP 2020 and explore the risk-return profile contextually to the ratings. The RAP will produce a technical annex and other background material that will be placed on the IEG website to explain how data were produced and used for the report, and their limitations.

2.3 Following this scope, the report's overarching question is, "what does the existing evidence show about the WBG's results and performance, taking into account types of intended development outcomes, risk level, and other project and country

---

<sup>2</sup> For IFC projects, the COVID impact will not be detectable in the reviewed sample. Although the reviewed sample includes a few projects of validated evaluations of CY20, the ratings of CY20 evaluations are based on pre-COVID results.

<sup>3</sup> The RAP 2018 Concept Note (Appendix G, pp. 25-26) provides a detailed Stakeholder Analysis and describes the specific interests that might prompt the Board and specific managerial groups (at the World Bank, IFC, and MIGA, and at the sector, region, and country level) to peruse the RAP report.

program characteristics?” To answer this question, the RAP will address the following three subquestions:

- How do existing ratings reflect the portfolio composition (varying across sectors, regions, WBG organizations, and over time) in terms of types of intended outcomes and risk (likelihood of achieving those outcomes)?<sup>4</sup>
- What patterns do we observe in outcome types, risk, and ratings across projects with different characteristics?
- What patterns do we observe in outcome types, risk, and ratings across countries with different characteristics?

2.4 The analysis will be carried out at the project- and country-level and for the World Bank, IFC, and MIGA separately. While the questions and sub-questions will be the same, the report will acknowledge that intended development outcomes and outcome types may differ considerably by sector, lending instrument, type of countries, and organization within the WBG – hence the ratings are not comparable across institutions and may have different meanings and interpretations depending on context or grouping mentioned above. This lack of direct comparability will require the team to generate, analyze, and report results separately for different groupings; it may also require use of different methodologies.

### **3. Outcomes, Risk, and Ratings**

3.1 RAP 2021 will build on a key finding of RAP 2020 regarding the relationship between development outcomes and ratings. RAP 2020 found no correlation between the *level* of development outcomes and outcome ratings after controlling for other project characteristics and it concluded that “many projects with higher-level objectives manage to achieve good outcome ratings, in part by having strong results frameworks to measure outcome achievement” (World Bank 2020, p. 46). This lack of correlation can be interpreted as a positive finding – the WBG can focus on and pursue important higher-level outcomes without this higher level of ambition being penalized with lower ratings<sup>5</sup>.

---

<sup>4</sup> When ratings are referred to, this means evaluation ratings, such as Development Outcome.

<sup>5</sup> RAP 2020 offered this interpretation: “The relationship between objectives’ outcome levels and projects’ performance is only modest and becomes insignificant when controlling for other factors. This finding runs counter to a key assumption prior to doing the analysis that one of the reasons for not setting higher level objectives is the risk of a lower rating. Instead, the finding

3.2 RAP 2021 will probe this relationship further and will analyze more explicitly how the risk-return profile is reflected in WBG ratings. The point of departure is the consideration that success in achieving project development objectives – expressed through outcome ratings (at the project level) – does not reflect the type of outcomes pursued or the expected probability (risk) of achieving (or failing to achieve) those outcomes.<sup>6</sup> As a result, it is difficult to interpret what higher or lower ratings mean in terms of substantial WBG achievements. Lower ratings are not necessarily an indication of poorer performance if they are the result of the WBG “pushing the envelope further”, that is, pursuing more transformative, innovative, and – plausibly – more challenging and difficult-to-achieve development objectives. Conversely, higher ratings are not automatically a marker of increased performance if they are achieved in a progressively more risk-averse environment.

3.3 To further explore the correlation of outcomes and ratings, RAP 2021 will deepen the analysis of both returns (types of outcomes) and risk (probability of achieving, or not, those outcomes). It will identify characteristics of intended development outcomes, projects, programs, and contextual elements that are associated with different risk levels and analyze the relationship between types of intended development outcomes, risk, and ratings. To do so, RAP 2021 will need to go beyond the outcome level framework developed by RAP 2020 and define more granular outcome *typologies* pursued by projects. Analytical approaches to do so are proposed in the next section.

## 4. Methods and Data

4.1 RAP 2021 will use statistical and qualitative approaches to address the questions related to results and performance listed in section 3. IEG envisages the analysis to be carried out in two steps. First, RAP 2021 will conduct a risk-return analysis – defining “return” in terms of types of outcomes and “risk” in terms of the probability of achieving (or failing to achieve) those outcomes. Second, RAP 2021 will use the results of this risk-return analysis to interpret outcome ratings (and potentially other project-level

---

shows no systematic trade-off between projects’ outcome level and ratings. This implies that many projects with higher-level objectives manage to achieve good outcome ratings, in part by having strong results frameworks to measure outcome achievement” (p. 46).

<sup>6</sup> This follows from how the rating system for World Bank projects is designed. In self-assessments and IEG validations, the relevance of PDOs rating, which contributes to the outcome rating, does not necessarily provide an assessment of the type of outcomes or of how challenging they are. In the case of IFC and MIGA, no rating is provided at PDO level. Ratings are provided at overall development outcome level and for a number of other dimensions. The ratings take into account achievement of objectives as well as performance of projects/companies against benchmarks.

ratings). Details on the analytical approach follow below. An important caveat is that the two-step approach proposed here might not yield definitive findings, and therefore adjustments and corrections to the approach may be needed as the analytical work proceeds.

4.2 For the risk-return analysis (first step) three approaches are envisaged. The team plans to use some or all of them in parallel for triangulation purposes and assess the consistency of the results:<sup>7</sup>

- (i) Classify intended development outcomes using a *typology* at the level of each individual project development objective that relates to how challenging (and hence risky) these outcomes are, informed by findings from the international development literature and experts' opinions. The expectation is that some types of outcomes can be more challenging than others and therefore may be associated with higher or lower outcome efficacy ratings (hence may have different average probabilities of success, or a wider variance of ratings).<sup>8</sup>;
- (ii) Use the conceptual frameworks of internal tools (such as ORAF/SORT, IRP/AIMM or IMPACT) to identify development outcome/project characteristics that are associated with high or low risk. Note that this option would not require the team to use the project-level risk ratings data generated by these internal tools. Rather, the team would study the framework(s) underlying these risk tools to gain insights into the elements that the institution associates with risk. For IFC, it is envisaged that overall credit risk ratings, environmental and social risk category and environmental and social risk ratings could be included in the analysis, in addition to several project characteristics (see box 4.1);<sup>9</sup>

---

<sup>7</sup> Moreover, each of these methods better applies to one or another WBG institution — for example method (i) is more appropriate for IBRD projects, while method (ii) for IFC projects.

<sup>8</sup> For example, building roads is generally considered/may be expected to be less challenging than creating government capacity to manage transport systems.

<sup>9</sup> Both AIMM and SORT have been introduced only recently, so the two data series are very short. Moreover, SORT data suffer from inconsistencies that make them (partially) unreliable for the RAP purposes. According to a recent OPCS analysis, TTLs implement SORT guidelines incorrectly (especially when it comes to update risk ratings, during project implementation), which prompted OPCS to revise the SORT implementation guidelines. RAP will therefore use the SORT guidelines and, potentially, the appraisal risk ratings to understand the association between different outcome, project, and context characteristics and risk ratings.

- (iii) Use supervised machine learning, to the extent feasible, to analyze project appraisal documents (PADs)/project Board proposals to explore the association between intended development outcomes and level of risk expressed within the narrative.

---

#### Box 4.1. World Bank Group Risk Management Tools

---

**SORT (Systematic Operations Risk-rating Tool)** was introduced in 2014 to rate risks of World Bank operations and country engagements. It replaces ORAF (the Operational Risk Assessment Framework) introduced in 2010 and never successfully operationalized. Previously, the main risks of investment lending operations and recommended mitigation measures were identified for all projects, but not in a structured and homogenous way, and no system existed to record risk ratings. SORT aims at measuring and monitoring the risk to the client's ability to achieve the development outcomes - effectively, efficiently, and sustainably - and the risk of unintended consequences. The assessment of risks takes into account both the likelihood of the risk materializing, as well as the severity of its impact on the achievement of the intended results, should the risk materialize. The SORT identifies nine risk categories, including, for example, technical design or political and governance risks, plus an overall risk rating. Some shortcomings in how SORT is implemented have recently emerged – teams have been inconsistent in updating risk ratings during implementation. OPCS is addressing these issues through updating SORT implementation guidelines and increased training.

Maximizing development impact while maintaining financial sustainability is one of the key risk management principles of the **Enterprise Risk Management (ERM)** framework of IFC. The ERM framework categorizes IFC's risk profile along five risk dimensions (credit, market, operational, liquidity, and business). In the past, credit risk and environmental and social risks, which are key elements for development outcome risks of projects, have been assessed and tracked systematically. Credit risks have been measured by **IRP (Investment Risk Platform)**, which replaced the previous CRR system in 2017. Environmental and social risk were measured by **ESRR (Environmental and Social Risk Rating)**. Since 2017, the concept of risks for projects to achieve development impacts has been incorporated in the **AIMM** framework, the ex-ante and monitoring framework for development impacts of IFC projects. In AIMM, the development outcome potential is discounted by "likelihood adjustments", which are based on operational, sector, political/regulatory policy, and country macro factors and are applied at approval and during monitoring. The overall AIMM scores reflect therefore the balance between outcome potential and risk (likelihood adjustments).

MIGA provides guarantees to ensure against political risk and non-honoring of financial obligations of sovereign/sub-sovereign and SOE. MIGA has been systematically assessing and tracking country risks, as MIGA is taking political risks through its guarantees. Like for IFC, the concept of risk for projects to achieve development impact has been recently incorporated in the **IMPACT** framework, the ex-ante and monitoring framework for development impacts of MIGA projects, which mirrors IFC's AIMM framework. Using a similar approach, "likelihood adjustments" are used to discount the development outcome potential.

*Source: IEG*

---

4.3 The risk-return analysis will involve sharpening the classification of intended development outcomes. The classification proposed in RAP 2020 is a four-level typology that approximates level of ambition (and risk) – the lowest level (level one) corresponds to “outputs” (understood as project deliverables: activities, products, and participation generated through the investment of resources, or goods and services delivered), while the highest level (level four) identifies outcomes generally associated with sustained change and aspirational goals such as the SCDs (World Bank 2020, pp. 37-41). RAP 2021 will move beyond this four-level typology by developing a more granular outcome typology.

4.4 The outcome-type framework to be developed for RAP 2021 will address some limitations of the outcome-level framework developed for RAP 2020. First, the outcome-type framework will use categories that classify the intended outcomes that the project is trying to achieve, based on the theory of change inherent to the project. The analysis will take place at the level of these outcome types – rather than collapsing the outcome types into levels – to preserve explicit descriptions of the intended changes, which will enable more granular classification and hence greater variation.<sup>10</sup> Second, each individual objective present within the PDO will be coded (rather than coding only the “highest” individual objective), to avoid discretionality (that is, to avoid the complexity and potential bias introduced when analysts must decide which individual objective to consider “highest” and then decide which category to apply to that objective). Third, objectives will be coded both at entry (start of the project) and at exit (project closing) to capture changes in objectives that may have occurred at restructuring.<sup>11</sup> Doing so will ensure that the pairing of outcome types with ratings is time consistent (i.e. done using the outcome types valid at exit) and will also allow for an analysis of how outcome types change during the life of the project.

4.5 The second step of the analysis consists of measuring the relationship between development outcomes, associated risk, and ratings. This analysis will be carried out using regression analysis. Development outcome characteristics associated with risk

---

<sup>10</sup> The classification of outcome types will include about 15 categories (such as, for example, “improved access to services”, “human capital built”, “strengthened institutions”, “increased equity” etc.), as opposed to 4 levels. These outcome types will not be ranked, in recognition of the fact that their role and placement in the project’s theory of change may vary depending on the context (type of project, country). Using outcome types, rather than outcome levels, will also allow for a more explicit interpretation of what the project is set up to achieve.

<sup>11</sup> In the case of IFC, objectives are assessments generally made at the project approval, given the fact that development objectives are not formally changed by the time of self-evaluation (at early operating maturity.)

identified at step 1 could be combined in an index (using principal component analysis or other techniques) or kept as is. The goal of the regression analysis is to explain ratings “adjusting” for development outcome risk, taking into account other project- and context-related characteristics. For MIGA, given the smaller number of projects, mixed qualitative and quantitative method could be used as an alternative.<sup>12</sup>

4.6 Assembling relevant information to explore this relationship will involve selecting and sharpening the definitions of relevant project- and context-characteristics. Project and country/context typologies will be constructed for regression analysis, similarly to what has been traditionally done for the identification of “drivers” of (or correlates with) ratings. As was done for outcome typologies, some project and context typologies will be defined based on their associated risk – for example, projects triggering specific safeguards, or projects or programs implemented in fragile environments.

4.7 Some of the steps described above may require a deep dive in specific GPs. This is because outcome types (and associated risk) are likely sector-specific and require a close analysis of theories of change and project types. As a result, some portion of the analysis, which will allow generating more granular findings, may be limited to a few sectors. Qualitative methods, including analysis of theories of change and experts’ interviews, may be used for the deep dive.

4.8 Across RAP 2021, the analysis will be carried out separately for World Bank, IFC, and MIGA, acknowledging the differences across risk frameworks, development outcomes, and ratings systems. For instance, the analysis for IFC project could take advantage of AIMM backfilled scores/ratings to provide additional insights on the developmental challenge that the projects intended to address.

4.9 The RAP 2021 analysis will be carried out at the project and country level. At the project level, the focus will be on the association between individual project ratings, outcomes, and risk. At the country level, the focus will be on the association between aggregate project ratings across a country, patterns in outcome types across projects in a country, and country characteristics associated with greater challenge (for example, FCV status; country income grouping; relevant macroeconomic indicators).

4.10 Different aspects of the analysis will use different time periods as appropriate. The most recent year for overall ratings trends will be FY20 for World Bank projects closed in that year (CY19 for IFC). Because the self-evaluation system of the WBG

---

<sup>12</sup> The number of MIGA projects that close and are evaluated each year is about 10, which makes regression analysis challenging.



involves a time lag (projects are evaluated after closing, and IEG validation comes 6 to 18 months later depending on when the self-evaluation is completed), analyses requiring more complete coverage would need to use earlier periods. For analyses requiring close-to-complete coverage across a three-year period, for example, the most recent 3-year period used would be World Bank projects closed in FY17-19 (CY17-19 for IFC), and an earlier period for comparison could be World Bank projects closed in, for example, FY12-14 (CY12-CY14 for IFC).<sup>13</sup> For IFC, these time periods may need to be adjusted based on the number of projects with backfilled AIMM.

4.11 RAP 2021 will also explore whether the WBG's corporate priorities could be assessed in the proposed framework. One possible avenue could be to analyze how aligned the outcomes (as captured by the new typologies) are with select SDGs. Another approach could consist of selecting one specific corporate priority and analyze how well reflected it is in the WBG portfolio. The corporate priority could be selected simultaneously to the sectors for the deep-dive, in order to maximize the efficiency of the analysis. The final decision will be undertaken at the conclusion of the piloting phase. For IFC, the changes of project portfolio objectives over time to better address IFC's corporate priorities such as FCV, climate change, and gender, will be assessed.

4.12 RAP 2021 will still report on aggregate ratings and trends over time, for continuity with past RAPs. The analysis of development outcomes and risk proposed here is meant to deepen the analysis of ratings and complement the traditional RAP reporting. IEG will use, like in previous reports, a wide variety of ratings, evaluative evidence, project document reviews, and other data from IEG's ratings databases (for World Bank, IFC, and MIGA projects and WBG country programs) to produce a variety of tabulations. These more traditional cross-tabulations of ratings and trends over time will be provided through a complementary web page (see next section). The main report, however, will focus on the specific relationship between ratings and the risk-return profile.

4.13 The report findings can have implications for understanding the impacts of the COVID-19 pandemic on results and performance of projects. An in-depth analysis of COVID-19 impacts will not be feasible, as many projects were delayed by the pandemic and only a limited number closed/assessed since its start<sup>14</sup> (still, it will be important to identify projects that closed during the pandemic to detect potential impacts in ratings and achievements.) However, the RAP analysis and findings – especially around risk –

---

<sup>13</sup> For IFC, the year indicates XPSR program year, not the year when the project was closed.

<sup>14</sup> Very few or none of these would have had completion reporting finalized, including IEG validation, by the time RAP 2021 analysis takes place.

might help identify sectors, types of projects, and countries where outcomes and risk may be affected the most because of COVID-19.

4.14 Table 4.1 lists activities, methodologies, and sources of information that are planned at this stage to develop evidence to answer the RAP 2021 questions.

**Table 4.1. Planned Data Collection and Analysis Methods**

Question/Line of Inquiry	Activities and Analysis Method	Sources of Information and Sampling
<p>How do existing <b>ratings</b> reflect the portfolio composition (varying across sectors, regions, WBG organizations, and over time) in terms of types of intended <b>outcomes</b> and <b>risk</b> (likelihood of achieving those outcomes)?</p>	<p>For World Bank/IFC/MIGA</p> <ul style="list-style-type: none"> <li>(i) Develop outcome typologies: identify and code outcome characteristics;</li> <li>(ii) extract and classify information on risk using findings from the literature and experts' opinions;</li> <li>(iii) extract and classify information on risk using typologies/examples proposed in ORAF/SORT (WB), CRR/IRP, E&amp;S risk category/ESRR, AIMM (IFC), Country Risk ratings, E&amp;S risk category, IMPACT (MIGA);</li> <li>(iv) extract and classify information on risk using supervised machine learning, to the extent feasible;</li> <li>(v) analyze changes in risk ratings in the ISRs (WB) and CSRs (IFC) during implementation – using statistical methods (tabulations, cross-tabulations);</li> <li>(vi) use regression analysis to identify characteristics of intended development outcomes of projects and country programs associated with different levels of risk.</li> </ul> <p>For IFC, also consider as an alternative:</p> <ul style="list-style-type: none"> <li>(i) Review backfilled AIMM claims and outcome potential scores/ratings in relation to project characteristics and risk.</li> <li>(ii) For risks, use credit risk ratings provided by CRR and IRP system as well as environment and social risk category and environmental and social risk rating.</li> </ul>	<p>For World Bank, IFC, and MIGA: Information on outcomes derived from project documents and IEG validations.</p> <p>Information on risk derived from SORT, ORAF, IFC and MIGA risk assessment tools (IRP, CRR, E&amp;S Category, ESRR, country risk ratings); project documents (including PADs, project proposals) restructuring papers; aide memoires, as appropriate); analysis of results frameworks.</p> <p>Part of the analysis can be limited to 2-3 GPs (deep dive).</p> <p>For IFC and MIGA, AIMM system (IFC) and IMPACT system (MIGA) can be also referred.</p>
<p>What patterns do we observe in outcome levels, risk, and ratings across <b>projects</b> with different characteristics?</p>	<ul style="list-style-type: none"> <li>(i) Identify and code meaningful <b>characteristics of projects</b>. Categorization needs to be adapted to the type of instrument and World Bank, IFC, MIGA. In general, these could be: sector/sub-sector; size (original commitment, total disbursement); client types, client capacity, use of TF/use of ASA/AS; co-financing</li> </ul>	<p>Project characteristics such as size and sector can be found in IEG Datamart, the Enterprise Data Catalog (EDC), and IFC's Management Information System (MIS). Other characteristics such as details of intended outcomes or project</p>

Question/Line of Inquiry	Activities and Analysis Method	Sources of Information and Sampling
	<p>(and use of blended finance); number of TTLs; project team/TTL during implementation located in country. Other potential dimensions could be: consideration of political economy by project team; level of trust between the client and the WB; good knowledge of the country by the project team; TTL experience; team experience;</p> <p>(ii) Correlate outcomes, efficacy ratings/project ratings, and risk level for different types of project.</p>	<p>team composition (or deliberations or decisions) would involve coding of narrative information in project documents and/or IEG validation documents.</p>
<p>What patterns do we observe in outcome levels, risk, and ratings across <b>countries</b> with different characteristics?</p>	<p>(i) Identify and code meaningful <b>country characteristics</b>, such as FCV vs. non-FCV; IDA vs. non-IDA; and other categories based on country income level; population size; share of population in extreme poverty; CPIA ratings; aid dependency; WBG's share of ODA; size of WBG portfolio in the country.</p> <p>(ii) Correlate outcomes types (using RAP 2021 refinements on RAP 2020 definitions), efficacy ratings, and other project ratings for different types of projects in different types of countries.</p>	<p>Data sources on country characteristics:</p> <ul style="list-style-type: none"> <li>• Income levels and lending groups</li> <li>• FCV status</li> <li>• ODA and WBG share – OECD database</li> <li>• CPIA: IDA; IBRD – making request to OPCS</li> <li>• Worldwide Governance Indicators</li> <li>• WBG Country level statistical capacity indicator</li> <li>• Global Open Data Index</li> </ul>

Source: IEG.

## 5. Accompanying pieces

5.1 The RAP 2021 will consist of a report and complemented by a website including an online dashboard and reference materials. As part of the RAP 2021 analysis, the dashboard that was originally produced in 2017 will be revamped. Currently, the data in this online dashboard for World Bank project ratings is regularly updated and can be used interactively to tabulate ratings by GPs, regions, lending instruments, etc. As part of RAP 2021, most tables and graphs will be moved to the website, as well as annexes, technical explanations, and additional relevant material, including an updated summary of the evidence about factors that correlate to World Bank project ratings from past RAPs, World Bank working papers, DEC reports, and external academic papers.<sup>15</sup>

<sup>15</sup> Starting point for this could be this blog post [Uncovering Factors of Project Success: A Literature Review](#) (based on work done for RAP 2017), see especially the dashboard there.

## **6. Engagement, Communication, and Dissemination**

6.1 The report will be presented to the Board in October 2021. During the preparation of the report, the team will engage with selected board members and advisors; members of Bank, IFC, and MIGA management; with Bank, IFC, and MIGA staff; with evaluators of other IFIs engaged in analyzing outcomes and outcome orientation of their organization, and academics and experts in organizational effectiveness. The team will develop an outreach and dissemination plan that will feature a robust online outreach.

6.2 The engagement with counterparts during the RAP analysis will be essential for several reasons. First, it will provide elements to sharpen and improve the analysis. Second, it will ensure clarity and consistency of the terminology. Finally, it will generate interest in and buy-in of the proposed framework and prepare the dissemination phase.

## **7. Team, Quality Assurance, Budget, and Timeline**

7.1 The core team members for the evaluation are Elena Bardasi (TTL), Jean Jacques Ahouansou, Joy Maria Behrens, Sylvie Bishweka, Kwabena Antwi Boasiako, Mariana Branco, Elisabeth Goller, Xiaoxiao Peng, Junko Sekine, Ichiro Toda, and Yi Yao. Other IEG staff and consultants will also contribute. Maximillian Ashwill will be the lead editor.

7.2 The report will be produced under the overall supervision of Alison Evans (Director General Evaluation) and the direct supervision of Oscar Calvo-Gonzalez (Director, IEGHE) and Galina Sotirova (Manager, IEGHC). During the production of the report the team will also regularly consult with IEG ICRR coordinators, CLRR coordinators, the IEGFP unit and staff and consultants involved in the validation of self-evaluations.

7.3 The report will also benefit from the advice of an external advisory panel comprised of academics and evaluation professionals. This panel will advise the team on methods and interpretation of findings. External peer reviewers are not envisaged, but this concept note and the draft final report will be subject to internal IEG review and the standard process of management comments.

7.4 The budget for the task is \$530,000 including \$40,000 for dissemination. Staff costs are estimated at 70 percent of the total task budget and variable costs, mostly for consultants, at 30 percent.

**Table 7.1. Timeline**

Timeline	
January 11, 2021	One-Stop Review Meeting for concept note
June 2, 2021	E-submission of concept note to CODE (AOB)
August 19, 2021	One-Stop Review Meeting for the draft report
August 31, 2021	E-submission of report to the WBG Board of Directors
October 2021	Board Meeting

*Source: IEG.*

## References

World Bank. 2018. IEG Review of Results and Performance of the World Bank Group 2018 With a Special Focus on Inclusive Growth. Concept Note. Independent Evaluation Group. Washington, DC: World Bank

World Bank. 2020. Results and Performance of the World Bank Group 2020. Independent Evaluation Group. Washington, DC: World Bank