RE-309

Report on the Evaluability of Bank Country Strategies

Office of Evaluation and Oversight, OVE

Inter-American Development Bank
Washington, D.C.
September, 2005

TABLE OF CONTENTS

ANNEXES

HTTP://IDBDOCS.IADB.ORG/WSDOCS/GETDOCUMENT.ASPX?DOCNUM=614305

I. INTRODUCTION

1.1 This is a report of OVE's main findings regarding the office's two years process of review of Bank Country Strategies. The report is organized in four chapters. The first chapter describes the proposed architecture of country strategies, as they are laid out in Management's *Country Paper Guidelines*[1]. The second chapter outlines the methods used to analyze the country strategies reviewed. The third chapter presents and discusses results. The last chapter summarizes findings and presents recommendations.

1.2 The Bank's country strategies are the primary instrument through which the Bank formulates its development program with each of its borrowing countries. The strategies and the activities surrounding them are required according to the Bank's Eighth Replenishment[2]. Due to their synchronization with a country's electoral cycle, the development of the strategies represents an opportunity for the Bank to share its medium-term development objectives with the new administration, thus placing the Bank's strategy in the context of the new government priorities and development objectives. The country strategies have also taken on the role as one of the main opportunities for the Bank to reevaluate the country's development challenges, assess the effectiveness of its past engagement with the country, and propose new medium term objectives, strategies to achieve these objectives, as well as a lending and non-lending proposal with which to execute these strategies.

1.3 The Bank's current normative for its country strategies can be found in the *Country Paper Guidelines,* heretofore referred to as the *'Guidelines'*. These Guidelines were developed in 2002, and serve as a working framework within which to develop the Bank's country strategies. At the time of the formulation of this framework the Board decided that Management and OVE would reevaluate the instrument in 2005 in order to determine the effectiveness of the *Guidelines*. To this end OVE reviewed eleven strategies over the past two years, as they were presented to the Programming Committee of Management[3].

A. The Architecture of the Country Strategy Papers

1.4 The architecture and purpose of the country strategies are defined by the *Guidelines*[4]. Accordingly, the Bank's strategies should accomplish four distinct

---

[1] See GN-2020-6

[2] According to AB-1704 "*Each year, for each borrowing member country, Management will present to the Board of Executive Directors a new or revised country paper (CP), providing a basis for Board-Management discussion of the status of the country's portfolio and of planned areas for Bank Group cooperation.*"

[3] The strategies reviewed were Argentina, Bahamas, Belize, Bolivia, Brazil, Colombia, Costa Rica, Ecuador, Guatemala, Paraguay and Trinidad and Tobago. The strategy for Haiti was not reviewed, given that it was a transitional strategy.

[4] The *Guidelines* present a somewhat disjointed mix of (i) varied descriptions regarding the purported objectives of strategies, (ii) prescriptions regarding specific content requirements for strategies (iii) and recommendations regarding the form, style, and length of strategies, where the connection among these

goals. They must (i) "*serve as an efficient vehicle to focus Bank-country dialogue*", (ii) "*design an agreed upon strategy", (iii) "anchor the programming process*", (iv) and "*provide a basis for program assessment*"[5]. In order to accomplish these goals, the *Guidelines* identify a mixture of six different properties and activities for strategies: (i) the strategies have to be consistent with the institutional strategy, (ii) the strategies should reflect a consensus between country and Bank, (iii) the strategies must be grounded in country and sector analytic work, (iv) the strategies must involve stakeholders and be disclosed to the public, (v) the strategies must be synchronized with the country's political cycle and (vi) the strategies must facilitate an assessment of the Bank's contribution to the development process.

1.5 The *Guidelines* also established a three-chapter structure for strategies, where each chapter has a specific purpose, and should contain specific elements. The first chapter should contain the country diagnostic. This diagnostic should lead to the identification of development challenges. The purpose of the diagnostic is to "*provide a macroeconomic and developmental context for the design of an appropriate Bank strategy towards the country, based on existing relevant literature and on the results of the Bank's analytical work on the country.*" It should be noted that although the document is not a macro-analytical document, it is required to be grounded in sector work. In fact, this requirement applies more broadly to all facets of the strategy, and not just the diagnostic, as is established by the *Guidelines*: "*[sector work] should serve three main purposes: provide a sound basis for dialogue with the country on development objectives and policy, inform the Bank's country strategy formulation, and facilitate project identification…*".

1.6 The second chapter should present an assessment of the past strategy. According to the Guidelines, the purpose is to "assess the Bank's strategic focus in the preceding programming cycle and evaluate the efficiency of strategy implementation." This includes an assessment of the effectiveness of the Bank, and identification of accomplishments and obstacles encountered, so as to uncover lessons at the strategic level. This requires a stock-taking of changes in development indicators from the previous strategy. Such changes would provide a basis for assessment of the Bank's impact. This chapter, along with the diagnostic chapter, should form a basis for the adjustment in the Bank's model for working with the country. This basis, in turn, should inform changes in the Bank's medium-term objectives, based on past experience and on changes in the country context. Chapter two also plays a central part in Bank accountability by establishing the Bank's achievements (and failures) during the past programming cycle, and by comparing these achievements with the Bank's formal intent.

---

items is not always evident. This structure necessitated a certain degree of latitude when reconstructing the logical flow of strategies in the context of the Bank's programming.

[5] See GN-2020-6, paragraph 2.1.

1.7 The last chapter should outline the Bank's strategy. It should identify the Bank's objectives, which should be clear and should be informed by the Bank's sector work. The chapter should then develop the Bank's strategy as well as the operational program. This is the main programmatic chapter, containing Bank objectives as well as a proposed Bank strategy and corresponding lending and non-lending program.

1.8 The *Guidelines* also set out specific requirements for this chapter. These include that indicators be identified for each level of objectives; both those that track implementation and those that "*provide a basis for future program impact*." The *Guidelines* also require the last chapter to address the consistency between the Bank's activities and its existing portfolio, the activities of other agencies, and the country's program. In addition, the *Guidelines* require that strategy risks be assessed and that mitigation measures be identified. Regarding financial risk, the Guidelines require that alternative lending scenarios be analyzed by taking into account "*exposure and creditworthiness indicators*".

1.9 OVE's contribution to the assessment of the effectiveness of the new guidelines has been centered on the issue of evaluability, which is a measure of how well a proposed intervention (strategy or project) initially sets out criteria and metrics to be used in its subsequent evaluation. Since the impact of strategies must be assessed, strategies must articulate precisely the objectives being pursued, the data that measure progress towards those objectives, and the mechanisms to be put into place to generate data on future performance. This analysis will focus on those aspects of the *Guidelines* that are central to the ultimate evaluability of the strategies. Thus, the study's main evaluative questions are:

1. Do the strategy papers contain the elements required to make them evaluable?

1.10 In the process of answering this first question the evaluation will also provide evidence toward a second question, which relates to compliance:

2. Have the strategy papers complied with the Guideline requirements?

1.11 It should be noted that the nature of this review establishes limits as to what can be evaluated. Aspects dealing with the process of strategy implementation were not assessed in the present evaluation, although they are assessed continually by this office in the context of each country's Country Program Evaluation (CPE)[6].

1.12 Lastly, much of the discussion surrounding the conclusions and recommendations is forward-looking, and ties into the needs of the Bank's borrowing clients, and by

[6] Likewise, the analysis is not a comparison of strategies before and after the Guidelines, as only the new strategies were reviewed; nor does the evaluation presume to evaluate the quality and extent of work done in preparation of strategies.

extension the needs of the Bank itself, as it adapts its programming process to the New Lending Framework.

II. METHODS

2.1 To answer the two questions posed in chapter 1 OVE has, over the past two years, reviewed the Bank's country strategies. The methodology of this review process, as well as the methodology of textual analysis applied as part of this evaluation are discussed below.

2.2 Prior to the presentation of each new strategy to Management's Programming Committee, OVE convened to review the strategies. As part of the review strategies were rated on seven main themes, by applying an evaluability instrument. The themes were: (i) diagnosis, (ii) analysis of previous programming cycle, (iii) objectives, (iv) logical consistency, (v) risks, (vi) indicators, and (vii) monitoring and evaluation. The instrument assessed the evaluability of the reviewed strategies, as well as the extent of compliance with the *Guidelines*; questions are derived from requirements of the *Guidelines*, or are derived from the basic principles of the same. For each strategy the office issued a *Evaluability Note* summarizing the findings of the review process, which were shared with Management.

2.3 The instrument was comprised of a series of questions in the themes outlined above. Each question was scored on a four-point scale. A score of '1' represents "no content" and '4' represents "superior content". Guidelines for each question are specified based on a mixture of content and quality dimensions[7]. Early versions of the instruments were applied to the strategies for Peru, Honduras and Guyana, in order to test the instrument and to refine it. Through this process questions were rephrased (and dropped), and scoring rules were introduced. The results of this analysis are based on the eleven strategies evaluated after the three pilot strategy reviews. Furthermore, the results from the application of the instrument to the eleven strategies were tested for validity prior to this analysis. The process of validation is explained in detail in Annex C.

2.4 The textual analysis complemented the methodology described above. The textual analysis entailed deconstructing the strategy into its different component parts and generating metrics for each of the components and the linkages between components. The standalone characteristics of each of the components included (i) a results framework for the assessment of the past strategy (this is based on chapter 2 of each strategy), (ii) a results framework for the proposed Bank

---

[7] In each scoring exercise five panelists participated, being that the author of the office's Country Program Evaluation for that country was not allowed to either score or be involved in the actual writing of the Note. In some cases more than five members were involved in the discussions, but only the five panelists scored the strategies.

program (this was based on the "Strategy Matrix"), (iii) an assessment of the extent to which strategy components reflected the use of sector work. In addition to the standalone components, the linkages between components were quantified regarding their logical consistency. The objective of this last step was to quantify the degree to which objectives were consistent with diagnostics; the degree to which the strategy was consistent with objectives; and the degree to which each component was consistent with proposed activities.

III. FINDINGS

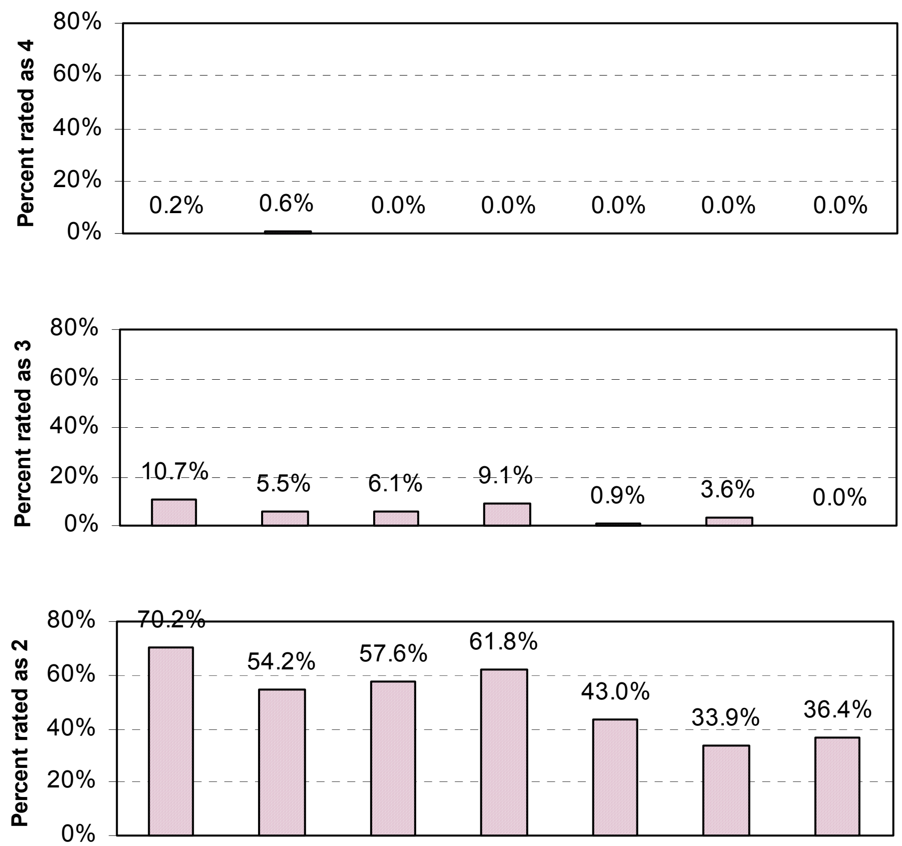
3.1 This chapter presents results from both the scoring exercise and from the textual analysis. Before proceeding to report on findings it is important to put the results in context. Throughout OVE's review it became obvious that much effort was put into the elaboration of the strategies, and likewise that much analytical output was produced. The list of economic and sector work done, along with the description of the process of consultation is indicative of the magnitude of the effort that went into the elaboration of strategy documents. In contrast with this effort, as will be seen below, the evaluation found significant deficits in the actual documents produced. This raises the issue of why the strategies do not reflect the actual effort and output that has gone into their production. This is an issue that we cannot answer, but that nevertheless is important for any future assessment of the process of country strategy formulation that Management may undertake.

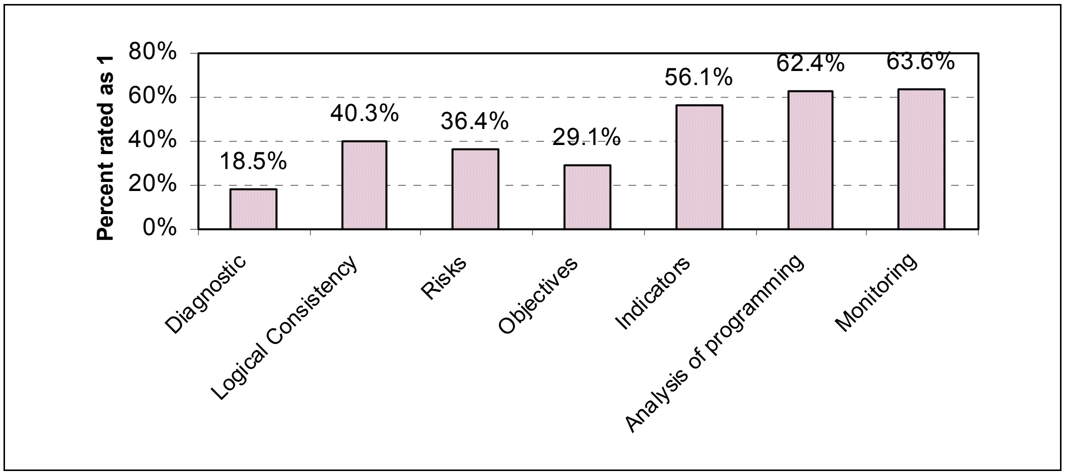
3.2 The evidence reviewed suggests that strategies were characterized by poorly developed and in many instances inconsistent diagnostics, a paucity of outcome indicators and on metrics for these indicators, limitations in the logical relationship between different components of the strategies, and the absence of a risk assessment. The implication of this finding is that the strategies reviewed are not evaluable. On the whole, the strategies did not contain the elements required in order to serve the purposes identified in the *Guidelines.* Details relating to each of the different evaluative categories considered are given below.

3.3 Figure 3.1 presents the summary findings from the application of the evaluability instrument to the 11 strategies reviewed. The results are presented by evaluability theme. The scores represent the proportion of ratings in each of the four categories. As can be seen, ratings were concentrated in the categories of '1' and '2'. There were few ratings of '3' and essentially no ratings of '4'. However, there is also much heterogeneity among results. For example, in the case of diagnostic there are a few ratings in the '1' category, indicating that strategies on the whole did contain diagnostic of development challenges. On the other hand, themes such as monitoring, indicators, and analysis of the previous programming period had over half of the ratings in the '1' category.

3.4 The data represented in Figure 3.1 also show that strategies performed worst in areas that require the identification and analysis of metrics that can be used to monitor and evaluate Bank performance. In the case of indicators, the low marks reflect deficiencies with metrics, baselines, milestones and targets for Bank objectives. In the case of monitoring the absence of specific activities to monitor implementation and outcomes, beyond the usual Bank instruments (PPMR, PCR, etc.), was one of the main problems encountered in the strategies. And in the case of the analysis of previous programming, there is little in the way of metrics regarding the evolution of outcomes associated with Bank-supported activities. In the case of risks the low marks reflect the absence of a risk analysis.

Figure 3.1: Summary Frequency Distributions





A. Diagnostic

3.5 Overall the strategy's diagnostics contained deficiencies related to: (i) overly general and descriptive diagnostics with little analysis, (ii) diagnostics with little or no foundation in evidence, and (iii) poor link between challenges identified and prescriptions given. Furthermore, the evaluation also looked at two sub-themes that are specifically highlighted in the *Guidelines*: public sector institutions and assessments of risk. With respect to the first, the evaluation found that Strategies do not reflect the existence of a model linking different elements of institutional strengthening of the public sector to development objectives. The evaluation also found that key concepts related to the efficiency and effectiveness of the public sector were not developed. With respect to the assessment of risk relating to the Bank's lending the evaluation found little discussion of country risk and the Bank's lending.

Table 3.1: Frequency Distribution and Means, Diagnostic Questions

		Percentage scoring			
	Mean	1	2	3	4
Sector Work (D3)	2.05	7%	82%	12%	0%
Development Challenges (D4)	2.12	8%	72%	20%	0%
Financial Exposure (D5)	1.69	43%	46%	9%	2%
Institutional Capacity (D6)	1.92	10%	88%	2%	0%

1 "no content"; 2 "poor content"; 3 "adequate content"; 4 "superior content"

3.6 As can be seen by the table 3.1 above[8], diagnostic scores were concentrated around a rating of '2'. A breakdown of the diagnostic score by specific questions revealed that strategies did better in providing broad characterization of development challenges and worse on themes that required analysis of specific

[8] Full sets of frequency distributions for all questions are in Annex B.

country characteristics. This was the case of financial exposure, in which 43% of the ratings were of '1' and 46% of '2'. In many cases there was no evidence of analysis or evidence of only a cursory analysis. In the case of institutional analysis, this is reflected by the concentration of ratings in the 'poor' category. The same pattern is repeated in the case of sector work. There was some evidence that the strategies were informed by this sector work, but this work was found to be of "poor analysis, not coherent, or incomplete"[9].

3.7 Diagnostics fell into one of four typologies, based on a 2x2 combination of diagnostic and policy prescriptions. The first is a strategy *with diagnostic* and *with prescription* for a particular theme. The second is a strategy *without diagnostic* and *without prescription*. The third is a strategy *without diagnostic* yet *with prescriptions*. The last type is a strategy *with diagnostic* yet *without prescription*. With respect to the match between diagnostic and prescription, the two can be logically related (leading to a consistent diagnostic) or not related (leading to an inconsistent diagnostic). Here the term inconsistent is not used as a value judgment, but rather to identify a diagnostic that does not logically match its corresponding policy prescription. Understanding which strategies fell into each of these categories will help explain the poor overall performance of strategies vis-à-vis the rating results above. In order to do this OVE textually analyzed the strategies with respect to the theme of institutional capacity of the public sector. This theme was chosen both for its perceived importance by the part of stakeholders and because it is one of two themes that all strategies must specifically analyze in their diagnostics, according to the *Guidelines.*

3.8 Annex D presents the summary of the diagnostics of all eleven strategies regarding aspects of institutional capacity of the public sector. The breadth of this topic was substantial: some strategies focused on efficiency of social spending, others on human resource and infrastructure deficits, while still others highlighted the role of political institutions in government. The central finding is that diagnostics reviewed were not analytical, and as such they were not able to link problem statements to strategy interventions through a model. The diagnostics rarely went beyond description. They reported on current levels of very general macro, fiscal and social variables, but added little information that would explain why these variables have behaved in the way documented, and what their impact on development may have been. For example, many strategies identified low levels of public spending on social programs as a major development challenge regarding public sector institutions. However, there was no model linking spending on social sectors to development outcomes, that would motivate the diagnostic statements regarding low levels of spending.

3.9 The second finding, which is related to the one above, is that many of the diagnostics were not sustained by evidence, or that the evidence employed was not appropriate for the problem statements given. The textual review revealed

[9] See rating guidelines in Annex for specific standards for each rating question.

that there were 53 distinct identified diagnostics statements over the 11 strategies[10]. Of these only 16 (30 percent) were based on corroborating evidence, either directly or as reflected in background work. This does not necessarily imply that additional work was not done, but if this additional work was done, it was not reflected in the diagnostics.

3.10 Regarding the sub-theme of public sector institutions and governance, the evaluation found that strategy diagnostic evidence was not adjusted to their problem statements. For example, many strategies relied heavily on opinion indicators, such as those found in *Transparency International*, *Kaufmann et all*, or *Latinobarometro* as a basis for sustaining their diagnostics[11]. The logical connection between high levels of perceived corruption, as measured by *Kaufman et all*, or low levels of public confidence in government, as measured by the questions in *Latinobarometro*, and specific judicial reform, civil service reform, and strengthening of planning and budgeting activities was not demonstrated. Although these indicators can serve to present a general overview of broad issues of governance within a country—particularly if they are tracked over time—it is not clear how they can serve as an input that would inform specific policy.

3.11 The last finding is that in many cases there was no evidence of a model that would connect the specific prescriptions to the development objectives. This is due in part to the fact that many key concepts regarding public spending and the fundamental role of the public sector were usually not developed. For example, from a reading of the 11 strategies, the Bank frequently identified member country's public spending as inefficient and ineffective. However, the meaning of what efficiency means was never clearly articulated. Does this refer to obtaining the maximum level of public services with a given level of inputs, perhaps quality adjusted?[12] Or does it refer to the ability to select spending that has the largest development impact per dollar spent and prioritize that spending? Is efficiency perhaps related to maximizing benefits for a particular target population while minimizing leakage? Given the fundamentally ambiguous nature of such words, the absence of an articulated definition of these key concepts rendered their usage uninformative for identifying a Bank-Country program.

3.12 Institutional capacity is one of the two *specific* diagnostic requirements called for in the *Guidelines*. A second is the analysis of the Bank's financial exposure and discussion of fiduciary issues in general. According to the Guidelines this analysis should be incorporated into the Bank's lending scenarios, *"[lending scenarios] should be discussed within the context of the country's overall debt*

[10] See the first column of Annex D.

[11] This was the case, for example, of strategy number 9, in which the only evidence presented in support of statements regarding corruption, political gridlock, judicial system inadequacies, and lack of transparency was in the form of these indicators. This strategy also contained 11 distinct policy prescriptions regarding institutional capacity and reform, including implementing e-government, civil service reform, and judicial reform.

[12] This is usually known in the literature as technical efficiency.

situation, showing the Bank's contribution to bridging the foreign savings gap, and dealing concisely with exposure and creditworthiness indicators." As can be seen in Table 3.1 above, only 11 percent of ratings are adequate or superior. In most cases this analysis was in the form of projections from third parties; there were no identified instances in which these projections explicitly modeled the role of the IDB, as a debt-generating agent, in preserving fiscal stability. This could have been done, for example, in the form of a mapping between different lending scenarios and fiscal outcomes, such as expected primary surplus, debt ratios, and expected spread over LIBOR, so as to motivate an optimal lending scenario for the Bank and the Country. If such an analysis was done, it was not reflected in the strategies.

Table 3.2: Public sector finance challenge and exposure score

Country Number	Public Finance Challenge	Bank financial exposure score
1	MEDIUM	1.20
2	MEDIUM	1.20
3	HIGH	1.00
4	LOW	2.00
5	HIGH	1.33
6	MEDIUM	2.00
7	MEDIUM	1.20
8	MEDIUM	1.50
9	MEDIUM	2.60
10	LOW	2.00
11	LOW	2.60

3.13 Table 3.2 presents the diagnostic score with respect to financial exposure, alongside with OVE's assessment of the degree of importance attributed to public finance challenge in each of the country strategies reviewed. This assessment was based on public sector primary surplus/deficit and on debt/GDP ratios (net of reserves), as well as on the preeminence attributed to fiscal stability in the strategy itself. The financial exposure analysis should be most prominent in countries in which the strategy identifies public finance as a top-level development challenge. That is not what was found. The two strategies that identified public deficits or debt as fundamental development challenges did little to address the issue of financial exposure analytically. In contrast, in cases where fiscal stability was less of a challenge, there was more analysis. From the above it is unclear what criteria Management is employing in prioritizing among sector work in country strategies, at least in the instances related to financial exposure.

3.14 It is generally maintained that decision-making that is informed by evidence will have a higher likelihood of being effective. This is particularly relevant to country strategies, given that these instruments represent a periodic opportunity for the Bank to formally realign its priorities in the context of the country's objectives, and to revisit the composition and magnitude of the types of interventions at the Bank's disposal. However, the conclusions that fall out from the analysis above is that the poor quality of the diagnostic, be it because of the lack of relevant sector work or the lack of a logical connection between diagnostic and specific Bank activities, deprives the strategy of one of the fundamental inputs for evidence-based decision making, thereby reducing the potential value added of the Bank for its borrowing members.

B. The Bank's Self-Assessment

3.15 The Strategy's analysis of the Bank's previous programming cycle's results is a chapter in and of itself. This reflects the importance of this self-assessment, both for the purposes of accountability and as an input into the Bank's program. The chapter's purpose, according to the *Guidelines*, is to *"…assess the appropriateness of the objectives and strategy selected, highlight the main areas of Bank activity, identify principal accomplishments and obstacles encountered at the country level and assess the effectiveness of the Bank instruments employed"*.

3.16 Assessment of past performance is complex, given that it requires addressing the problem of identification of the effect of the project or policy that the Bank had supported. Bank activities are conducted in an environment influenced by a multitude of other factors with direct incidence on the development objectives of interest, so sorting things out requires taking these factors into account and placing the magnitude of the Bank's activities in the wider context of other development activities. These considerations notwithstanding, a minimum requirement to address the Bank's contribution to development is to define objectives, indictors, and to measure these indicators before, during and after Bank interventions. If this trinity of information is not present, parsing the Bank's contribution is not possible given that there will be no information to work with.

3.17 This information is a necessary condition for the Bank to assess the *"effectiveness of its instruments employed"*. Therefore it constitutes the backbone of OVE's analysis in this section. In order to do this exercise, results frameworks were constructed and matched with the findings from the application of OVE's evaluability instrument. Results frameworks identify the presence or absence of indicators, baseline information, as well as information tracked over time on these indicators, for each stated objective of the prior strategy.

Table 3.3: Evaluability of the Bank's Previous Programming					
		Percentage scoring			
	Mean	1	2	3	4
Changes in Indicator values (P1)	1.29	71%	29%	0%	0%
Development objectives achieved (P2)	1.40	64%	33%	3%	0%
Mitigating factors (P5)	1.83	28%	62%	11%	0%
Percentage of objectives with indicators	23				
Percentage of objectives with baselines	2				
Percentage of objectives tracked (targets)	4				

1 "no content"; 2 "poor content"; 3 "adequate content"; 4 "superior content"

3.18 Table 3.3 presents the results of the scoring exercise for questions related to the previous programming exercise (top panel) and of OVE's results framework exercise regarding the results framework of the previous programming cycle (bottom panel). The scores show that no strategy reviewed adequately tracked indicators of prior programming cycle objectives. In 71 percent of the cases the rating is of '1', indicating that there was no reporting of outcome indicators. The second question measures the extent to which discussed progress toward development objectives. The results show that most strategies did not identify or discuss progress toward development objectives, with 97 percent of responses scoring a '1' or '2'. Strategies performed better at identifying mitigating factors that impacted on the Bank's activities. However, in this case the results indicate that factors were identified, but that, according to the rating guide, they were not relevant or not specific enough, with 62% of ratings in the category of '2'.

3.19 The bottom panel of Table 3.3 tracks the extent to which the strategies report on the indicators, with baselines and currently observed values, of Bank objectives established for the previous strategy. Ninety-eight objectives were identified for past strategies, with some strategies showing a large number of objectives and others only displaying select overarching objectives. The Table shows that in only 23% of cases were indicators identified for the Bank's development objectives. This reflects in part the poor identification of indicators from previous strategies; however, in cases where indicators are not identified *ex ante* there was usually no evidence of an attempt to reconstruct them *ex post*. When indicators are defined, in only 2% of cases do they also report baselines, and in only 4% of cases is the current status of the indicator reported[13]. In only two cases (strategy numbers 5 and 7) is there an explicit methodology to follow up on strategy development impact. In other cases, rather than report own indicators, strategies rely on third party assessments of progress. Although this is a legitimate

[13] Indicators that refer to approval of laws or other measures of institutional change often included as part of sector loans are not considered in the calculation. This is done for two reasons. First, these are actions that constitute means toward an end; second, in these cases baselines are nonsensical, since they will always imply the absence of a particular activity and the target the presence of such activity.

approach, in most cases where the approach was used the relationship between these indicators and the Bank's program was not demonstrated[14].

3.20 Strategies did contain statements regarding the Bank's development effectiveness. These statements were, nonetheless, either statements of what the Bank did—rather than what its development impact was—or they were statements of development impact, only lacking the required supporting data. Strategy number 4 is illustrative of the former. The strategy recognizes the importance of looking over a longer term in order to better identify contributions of the Bank to development by reviewing the previous two programming cycles. However, the review is limited to statements such as "*En el área de reducción de la desigualdad social y la pobreza, se dio prioridad a los sectores de reformas en educación, salud y desarrollo municipal y urbano*" or "*En el área de competitividad y modernización, el Banco continuó apoyando los esfuerzos del país destinados a reforzar la infraestructura (transporte y energía), fortalecer el sistema financiero, apoyar a medianas, pequeñas y micro empresas, asistir al desarrollo tecnológico, desarrollar el turismo, apoyar la competitividad del sector agrícola, y mejorar los marcos regulatorios."* The strategy contained no indicators nor were data presented regarding what improved over the period. In other cases the document came closer to identifying specific impacts, such as in the case of municipal administration: "*dichas operaciones fueron exitosas, reconociéndoles un impacto directo y significativo en la recaudación tributaria*". However, even here there was no specific mention of the improvements in revenue creation. In all, the document made 16 similar statements regarding the Bank's development effectiveness; in none of the cases was data given to substantiate these claims.

3.21 The most common finding was either no comment on development objectives or comments restricted solely to discussions of execution issues. In the case of strategy number 10, the strategy contained a detailed discussion of execution, and identified many of the issues related to disbursement delays. It also discussed the fall in demand for Bank lending by part of the country and recognized that, essentially, the country demonstrated a revealed preference for financing its development through other means (most of the originally programmed Bank operations were dropped). However, the document did not link these problems to development objectives, nor did it report on the evolution of these objectives over time.

3.22 In two cases efforts were made to present tables reporting on previous development objectives, with indications corresponding to positive or negative progress. This resulted in above average scores for the strategies. However, the potential value added from having this form of results tracking was mitigated by the lack of supporting data. For most objectives the tables and the ensuing

[14] The most common approach is to mention millennium development goals (MDG). Other approaches included relying on World Bank assessments.

discussion in the text did not identify indicators, nor did they specify baselines, targets, and corresponding current values for indicators, nor the magnitude of change. The result is that the table's representation of the Bank's impact cannot be confirmed.

3.23 A credible assessment of the Bank's effectiveness is no simple task. It requires collecting data on program development objectives and developing relevant counterfactuals regarding what would have occurred had the reforms or investments that the Bank supported not taken place. Or, an equally relevant question, what would have occurred had the Bank and country financed a different development program. However, a necessary condition for this analysis to be attempted is that information on development objectives be presented. The results framework findings, along with the evaluability scoring results, and the textual analysis of the strategies all indicate that none of the required elements for this analysis were developed in the Bank's country strategies.

C. Indicators and Monitoring of the Bank's Proposed Program

3.24 The previous section discussed the results of the Bank's self-assessment. In this section the paper looks at Indicators and at the Monitoring of the Bank's proposed program. The establishment of indicators serves two main purposes. The first is to ensure that the minimum required information to track changes in development objectives over time is identified. Data on baselines and targets provide a necessary condition for the Bank to identify its development impact. The second is the purpose of accountability. By establishing indicators that can track development objectives the Bank can hold itself accountable for both the successes and the failures of its program with borrowing countries.

3.25 The Guidelines call for the identification of separate Bank and Country indicators: *"An attempt should be made to include, in the strategy matrix, performance indicators that can be used to monitor the implementation of the Bank's strategy for the country, as well as those that might be employed—either directly or by proxy—to assess the Bank's contribution to the attainment of the country's development objectives."*[15] However, the evaluation results revealed a number of findings regarding evaluability: (i) gaps in the identification of indicators and their metric values, (ii) indicators that are not specific to the interventions proposed, (iii) indicators that do not measure the purported benefits of the Bank's interventions. There were also problems detected in the identification of baseline, milestone and target data that would allow tracking of indicators. Lastly, there were many instances in which activities, objectives and indicators are used inappropriately (e.g. activities identified when an objective is called for).

---

[15] The Guidelines also call for the identification of both implementation and outcome indicators: "*two types of indicators must be identified for each development challenge that the Bank strategy addresses: those that track strategy implementation and those that would provide a basis for future impact assessment.*"

3.26 OVE's assessment was based on two tiers of indicators: those related to the overall development objectives, and those related to each of the strategy component objectives. Scores for indicators, baselines and targets are presented for both the development objective level and the strategy component level in Table 3.4 below. The first finding is that overall there was a paucity of indicators and their corresponding metrics. All of the ratings at the strategy level are '1' or '2'; a similar result was found in the development objective level. However, some categories performed better than others. Most objectives identified some form of indicators (55% in the case of Development Objectives and 72% in the case of strategy objectives). On the other hand, the concentration of scores in the lowest category for targets (74% and 79%) reflects the absence of identification of targets at both levels of indicators[16]. The results for baselines was somewhat better, although a majority of the scores were still in the lowest category (52% and 54%).

Table 3.4: Frequency distribution of basic evaluability requirement scores

		Percentage scoring			
	Mean	1	2	3	4
Development objectives					
Indicators (I1)	1.56	45%	53%	2%	0%
Baselines (I2)	1.51	52%	44%	4%	0%
Targets (I3)	1.26	74%	26%	0%	0%
Startegy component objecties					
Indicators (I1)	1.72	28%	72%	0%	0%
Baselines (I2)	1.46	54%	46%	0%	0%
Targets (I3)	1.21	79%	21%	0%	0%

1 "no content"; 2 "poor content"; 3 "adequate content"; 4 "superior content"

3.27 The lack of evaluability data shown above is corroborated by the analysis of each strategy's "Strategy Matrix". This analysis is presented in Table 3.5. The table shows that the Matrices had few indicators, fewer baselines and targets and almost no milestones. It also shows the variability across strategies mentioned above. For example, in both cases strategy 8 does relatively well across the

[16] In some cases the lack of evaluability information is a result of the strategy's particular approach. Fro example, in strategy number 7 instead of using targets, the strategy establishes "tendencias", in which the "target" in each of the sectors is to show any improvement in indicators, independent of the magnitude of this improvement. The strategy justifies this conceptualization of targest based on "*(i) que la definición de logros específicos corresponde a cada proyecto; (ii) el problema de atribución del Banco; (iii) que en una estrategia de tres años se lograrán avances parciales en las metas, requiriéndose acciones continuas de apoyo en los años venideros.*" However, the strategy does not specify (i) how having targets at the project level precludes having them at the macro level, nor (ii) how not setting targets in the first place will attenuate—rather than aggravate—the problem of identification of the Bank's impact. And lastly, the strategy also does not substantiate the claim in item (iii) that the Bank cannot be expected to produce a measurable impact in three years, particularly given that at any given point in time the Bank's most important asset is its existing portfolio rather than any new approvals coming online.

board. It is the only strategy that has evaluability statistics in all categories, including a serious effort to identify milestones. Other strategies, such as strategies 5, 6, and 7 also do better in identifying indicators and in some cases baselines. However, overall, no strategy achieved an evaluability index of over 50 percent. The results also shows that, as in the case with the application of the evaluability instrument, the results at both the development objective level and at the strategy level are similar[17].

Table 3.5: Evaluability of the Bank's Previous Programming

Strategy Number	Indicators	Baselines	Milestones	Targets	Evaluability
1	12.5%	0.0%	0.0%	6.3%	4.7%
2	16.7%	16.7%	0.0%	0.0%	8.3%
3	30.8%	0.0%	0.0%	0.0%	7.7%
4	4.5%	0.0%	0.0%	0.0%	1.1%
5	55.0%	55.0%	0.0%	55.0%	41.3%
6	60.0%	45.0%	5.0%	55.0%	41.3%
7	46.2%	30.8%	0.0%	7.7%	21.2%
8	46.4%	28.6%	17.9%	32.1%	31.3%
9	18.8%	6.3%	0.0%	18.8%	10.9%
10	14.3%	4.8%	0.0%	14.3%	8.3%
11	54.0%	0.0%	0.0%	18.2%	18.2%
Objective level					
Development Obj.	31.0%	23.0%	2.8%	22.8%	20.0%
Startegy	33.7%	17.8%	3.3%	21.8%	19.2%

3.28 A second finding is that the quality of performance outcome indicators was poor. In the strategies reviewed indicators focus on tracking outputs rather than outcomes. There was usually no separate attempt to identify one type of indicator from another. For example, in strategy number 1 only 2 of the 14 statements made under the column of strategy indicators actually monitored outcomes. This characteristic was present to different levels in each of the strategies. It was explicitly identified as a major problem in six of the ten Evaluability Notes[18].

[17] Although better evaluability would be expected with respect to indicators at the top-tier level, given that at this level indicators are more readily available from national statistical databases, this is not strictly the case. In the case of OVE's evaluability instrument the scores are evenly split between '1' and '2' at both the strategy and development objective level. In some cases strategies score lower on development objectives because of the approach taken to identify only the government's development objective indicators, limiting the Bank's indicators to those at the strategy level. In the case of the review of the "strategy matrices" the metrics are also evenly divided in most categories.

[18] Notes for strategies 2, 4, 5, 8, 9 and 10. In some cases, such as that of strategy number 10, this may have been due to a deliberate structuring of the Strategy Matrix so that strategy indicators are limited to the output level; by design the strategy does not identify outcome measures for strategy components: "*Output performance indicators were selected for each initiative, and country performance outcomes were selected for each pillar. This presentation allows for future monitoring of the implementation of the strategy by*

3.29 In many instances the deficiencies outlined above were compounded by the fact that stated indicators were in fact objectives and not indicators. For example, in strategy number 2 one of the indicators was defined as "*Improve design and marketing of handicraft industry, in order to increase its quality and availability to the growing tourist flow*." In this example the stated indicator was not an indicator rather, it is a combination of different objectives. In strategy number 1, 4 of the 14 statements listed under indicators were actually objectives and not indicators. Conversely, identified Bank objectives were often either not objectives or combination of objectives, activities, and indicators. In the case of strategy number 5, one of the strategic objectives is to "*apoyar la reestructuración de sector agropecuario y manejo de recursos naturales*". This again does not impart what the Bank hopes to achieve, but rather an activity (apoyar la reestructuración). Presumably having a restructured sector is not the objective, but rather the means to an objective.

3.30 The poor quality of indicators is also manifest by the poor linkage between the indicators and their corresponding objectives. This limitation was mentioned in five of the 10 *Evaluability Notes* issued by OVE[19], but was present in different degrees, in all of the strategies reviewed. In some cases this was due to the reliance on indicators that are generated by third parties, and which may not be adequate to measure the impacts expected of specific Bank interventions.

3.31 The last finding has to do with the disparate results by sector. In themes relating to social development and poverty reduction the evaluability index was nearly 30%, while in themes of modernization of state and competitiveness it averaged 16%. This was a result of both better conceptual models in the social sector and of better information content in terms of indicator baselines and targets. In the case of sectors such as education and health, both the model of intervention and the metrics by which outcomes can be measured were better defined, and somewhat consistently defined across strategies. In contrast, in themes such as state modernization, there was no evidence of a systematic use of objectives, indicators, and metrics across strategies. Each strategy used different terminology and identified different indicators. This reflects both the lack of a coherent model in the sector—a finding already echoed by the analysis of the governance and institution sub-theme above—as well as a difficulty in identifying information sources in the country.

---

measuring change in output indicators directly related to actions supported by the Bank. Moreover, country performance outcomes are also appropriate indicators because those point to the final outcome for the country, where Bank actions, added to those of many other actors, determine the final outcome." The implication is that the Bank's responsibility is restricted to monitoring outputs, leaving all accountability regarding development effectiveness to the country.

[19] Notes for strategies 1, 5, 6, 8, and 10. It was also present in note 11, which was not issued. A case in pint was the note for strategy number 5, "*Por su lado los indicadores seleccionados son en muchos casos inhábiles para medir un efectivo progreso hacia el logro de los objetivos. El recurrir a indicadores generados por otras instituciones que no siempre encajan con el objetivo específico definido genera una ruptura entre objetivo e indicador de muy difícil solución.*"

3.32 In addition to identification of metrics, baselines, and metrics to be tracked over time, the *Guidelines* also call for the identification of mechanisms to track and monitor implementation of both the development indicators and the implementation progress. The *Guidelines* state that "*brief statement as to how the Bank's program would be monitored and assessed*" should be included.

3.33 The implementation theme is measured with a set of three questions in OVE's evaluability instrument: measures to monitor implementation, measures to monitor development outcomes and measures to monitor risks. Table 3.6 presents the summary statistics of each of these scores.

Table 3.6: Frequency Distribution of Monitoring Scores

		Percentage scoring			
	Mean	1	2	3	4
Actions to monitor development objectives (M	1.38	62%	38%	0%	0%
Actions to monitor implementation (M2)	1.63	29%	63%	0%	0%
Actions to monitor risks (M3)	1.11	89%	11%	0%	0%

1 "no content"; 2 "poor content"; 3 "adequate content"; 4 "superior content"

3.34 The first finding is that most strategies did not identify measures for monitoring either risks or development objective implementation, above and beyond the Bank's normal monitoring mechanisms. This is reflected by the absence of any scores in the '3' or '4' categories, and the preponderance of scores in the '1' category (62% for development objectives and 89% for risks). In the cases where activities are identified, they are usually identified by mention to data sources that will be used to monitor and not through the explicit assignment of responsibilities for these monitoring activities. In other cases annual revisions of the strategy were mentioned, but usually in the absence of the identification of specific activities. In cases where alternative lending scenarios were specified, there were no specific measures identified to monitor the triggers. That is, the institutional arrangements that would place a country in a high or low lending scenario were not made explicit. In any case, it should be pointed out that the gaps in development objective indicators found in most strategies would make monitoring effectiveness a particularly onerous task. In the case of risks none of the strategies reviewed advanced specific metrics by which these could be tracked, making monitoring an even more difficult task.

3.35 The implications stemming from the absence of evaluability information are twofold. First, future assessment of the Bank's impact will be made more difficult, to the extent that the groundwork required for collecting baseline information was not done. Second, the lack of information will also make it difficult for the Bank to hold itself accountable.

D. The Diagnostic-Objectives-Program Nexus

3.36 The assessment of logical relationship of strategy components is perhaps the most complex and difficult, given that the topic deals with the relationship between different strategy components rather than attributes contained within components. On the other hand, conceptually the questions to be answered here are quite simple. (i) Does the diagnostic logically lead to the identified development challenges? (ii) Are the Bank strategy components logically related to the Bank objectives? (iii) Are the proposed activities logically related to the strategy components and to the Bank objectives?

3.37 The three questions above are analogous to those asked in the evaluability exercise. However, as with the other themes reviewed above, they were supplemented with information from the textual quantification and analysis of the strategy documents. In the case of logical consistency the logical decomposition relies to a greater degree on subjective evaluations regarding the relationship between the components mentioned above.

3.38 The *Evaluability Notes* identified two major disconnects regarding diagnostic and objectives. The first is the logical relationship between the diagnostic and the objectives selected by the Bank and the second is the logical relationship between diagnostic and the Bank's model for reaching stated objectives. It should be noted that much of the diagnostic regarding the Bank's model is developed in the third chapter of the strategies, rather than in the diagnostic chapter. The *Notes* took all of the diagnostic found in the strategies into consideration and not just that found in the first chapter.

3.39 Problems in the logical connection between the strategy components and the diagnostics were identified in almost all strategies. For example, in the case of strategy number 4 the Note found a weak link between diagnostic and the Bank's model, "*The identification of Bank actions seem not to be drawn from detailed sector analysis that present and discuss potential alternatives to reach the same goal*." A similar finding was echoed in the Note regarding strategy number 8, "*A model linking the deficits identified in the diagnostic with the Bank's objectives is not developed. The diagnostic identifies a number of key development challenges… From this collection of development challenges, the overall task of reducing poverty is selected as the most important, with better governance, a more equal development of human capital, and a more efficient production as specific vehicles by which poverty reduction can be achieved. However, the document does not explicitly specify a model in which these three objectives will lead to poverty reduction, nor does it identify why this approach, rather than other possible approaches, is the most likely to achieve poverty reduction*." In another strategy—strategy number 10—a break between the diagnostic and specific operations was present in the property rights loan. Although a loan for improving property rights was one of the two largest proposed, there was no explanation regarding why land rights were considered a binding development constraint. The only evidence of diagnostic given consisted in the following statement: "*Another area that complements secured transactions is improving property rights, especially in rural areas where until recently land was in public hands.*"

3.40 The findings from the scoring exercise and those reported in the *Notes* are also reflected in the textual analysis of the documents. One exercise conducted was to match the diagnostic elements to the selected bank objectives. In the case of strategy number 5, of the 16 development objective challenges identified in the text as areas for Bank intervention, only three of them had an empirical basis in the diagnostic (either directly or reflected by sector work). The remaining statements of development challenges were given without adequate evidence. This pattern was mentioned in the *Evaluability Note* for that strategy: "*La enumeración de problemas efectuada no refleja la presencia de un diagnóstico en el que sustentar la programación del Banco, ni contiene un trabajo analítico, ni es posible deducirlo del mismo. De hecho, de la lectura del documento no se concluyen las razones que subyacen en los avances y retrocesos sufridos por el país en la década. No se identifican los desafíos específicos para el logro los objetivos de desarrollo, más allá de identificar dos factores o restricciones generales…"*

3.41 In some cases problems in consistency are identified by looking at changing diagnostics over time[20]. This is the case, for example, in strategy number 2. In the previous strategy for that country the Bank indicated that "*Labor legislation has created an array of labor market rigidities, including lack of flexibility in working hours, as well as restrictions to hiring and firing.*" In 2004 the new strategy indicated that "*The expected discounted cost of dismissing a worker is half the [region] average and payroll costs are around 7 percent. In conclusion, even though the wage rate is comparatively high in comparison to other countries [in the region], labor costs and other incentives to the private sector seem to be relatively beneficial to private investment and not a binding constraint.*" In other words, from one strategy to the next there was a reversal in the diagnostic, from labor markets being rigid to labor markets being relatively beneficial. In the first strategy the diagnostic contained no empirical justification, whereas in the second strategy there were indications that the diagnostic finding was informed by evidence. It would, however, be premature to conclude that the change was due to a better informed diagnostic; the strategy did not indicate if the change in findings was due to an incorrect initial diagnostic or to radical changes in labor markets.

3.42 The relationship between the Bank's objectives and its program is perhaps the most important piece of logic in a country strategy, given that it provides the link between a development concept and its implementation. The remainder of this

[20] Although the methodology employed did not include an assessment of past strategies' diagnostics, in reviewing the chapter dealing with the previous cycle it was necessary to briefly review the objectives and program of the Bank in the previous strategy.

chapter looks at this relationship. As in the analysis done previously, the findings rely on evidence from the evaluability exercise, and supplemental evidence from the textual analysis conducted.

3.43 Table 3.7 presents the frequency distribution for four questions related to logic. Each question addresses a different part of the link between the strategy and the actions surrounding it. The first two relate to the Bank's specific objectives. Are the objectives consistent with the strategy and with the activities proposed? The second set of two questions relate to the Bank's institutional objectives. Are the activities related to the institutional objectives, and more specifically to poverty reduction and to sustainable growth?

3.44 Table 3.7 shows that in almost all strategies (88% of ratings) there was a relationship between the objectives and the strategy. This is the area where the strategies scored better. The caveat here is that the consistency was only at the most general level. In other words, the direct pathways (or model) that relate the strategy to the development objectives are usually not developed. This is reflected by the scarceness (13%) of ratings in the '3' or '4' category. In terms of the connection between activities and objectives the strategies fared worse. In this case the logic relating specific interventions and sector objectives was either not given (36%), or was lacking specificity and relevance (58%). With respect to two of the overarching institutional objectives—growth and poverty reduction—the concentration of the ratings evenly in the two bottom categories reflects either the absence of an articulated model or a poorly articulated model. Lastly, the application of the evaluation instrument to the issue of consistency between government and Bank program, revealed both the absence of this consistency in some cases (31% of ratings) or poor consistency (60% of ratings). In most cases this was due to strategies that did not adequately articulate what the country program was.

Table 3.7: Frequency Distribution of Questions Relating to Logic

		Percentage scoring			
	Mean	1	2	3	4
Relationship between Bank objective and strategy (O4)	2.09	4%	84%	13%	0%
Relationship between Bank objective and proposed activities (C4)	1.69	36%	58%	5%	0%
Contribution to poverty reduction demonstrated (C6)	1.55	49%	52%	3%	0%
Contribution to growth demonstrated (C7)	1.64	41%	54%	5%	0%
Consistency between Bank and Country objectives demonstrated (O2)	1.78	31%	60%	9%	0%

1 "no content"; 2 "poor content"; 3 "adequate content"; 4 "superior content"

3.45 A stocktaking of the number of strategic objectives backed by substantial activities by the Bank illustrates why strategies scored low in logic. Overall, (i) many specific Bank objectives were not backed by specific Bank activities, and (ii) many identified Bank activities were not logically related to the objectives they were to support. In the case of strategy number 4, five of the 13 specific operations proposed were not related to the specific objectives that they supported. For example, although a housing project is listed under poverty reduction, the logical relationship between these two was not demonstrated. In four other cases the objectives identified were not backed with specific activities[21]. In other cases details of the activities, such as the type of intervention and amount of resources involved, were not stipulated.

3.46 In the case above the Bank did propose a substantial pipeline of activities. However, in other cases the connection between strategy and activities could not be demonstrated due to the fact that the Bank proposed very few specific activities in the strategy. This was usually the case in which the Bank attempted to draft a strategy during periods of crisis or when the Bank and the country could agree on very little. Strategy number 3 is an example of the former. Rather than identifying new operations the existing portfolio was the centerpiece of the Bank's programmatic response, as was pointed out in the strategy, "*La alineación de la cartera del Banco con la EBRP y los Objetivos de Desarrollo del Milenio constituye el principal activo del Banco para mantener una visión de mediano plazo focalizada en la consecución de metas sociales y económicas*". However, apart from references in the Strategy Matrix to a possible reformulation and a reference to a CPR mission, there were no specific proposals for formulation. The strategy did not make clear how this undisbursed balance would be brought to bear in support of the Bank and country's development objectives.

3.47 A second issue relating to the definition of activities is the choice of sector and choice of intervention. During the assessment of strategies it was found that the Bank often invoked its comparative advantage as a criterion for selecting areas in which it would (and would not) participate. The *Guidelines* state that the Bank's comparative advantage should be used as a criteria for selection: "*The strategy is designed on the basis of the Bank's experience in the country, the government's development program and commitment to MDGs, the assistance strategy of other development agencies and the Bank's estimation of its areas of **comparative advantage***" (emphasis added). In most cases strategies at least mentioned this principle as a basis for prioritization of activities, such as in the case of strategy number 9: "*Se concentra en las áreas prioritarias del Gobierno en las cuales el Banco tiene ventajas comparativas y coordina su acción con las correspondientes a otras fuentes de financiamiento del país, contribuyendo con ello a aumentar la eficiencia del proceso de inversión pública y de asignación de recursos.*" Similar statements are found in 7 of the 11 strategies reviewed. However, mention of sectors in which the Bank had a comparative advantage was only present in three strategies, and explicit criteria for selection of activities were only present in one of the 11 strategies reviewed[22]. In the remaining strategies there were no

[21] This exercise is made difficult given that activities identified in the Annex of the documents do not always match those that are listed in the Strategy's Matrix.

[22] The only exception to this is the case of strategy number 10. The strategy document explicitly states that the strategy will emphasize "*the use of criteria to help guide the country dialogue and setting of priorities for future Bank-supported programs during the period 2004-2007*." These criteria are *"(i) The extent to which they are likely to reduce the transaction costs in the economy and (ii) The extent to which they focus*

operative definitions of prioritization through comparative advantage or any other form of prioritization.

E. The Bank's Treatment of Risk

3.48 The Guidelines require that the strategies contain an assessment of risk and for measures to attenuate this risk, "For each identified risk to the implementation of the Bank's strategy for the country, the corrective or ameliorative measures, by the Bank of the country, need to be ascertained…"

3.49 In order to determine the degree to which strategies address risk the analysis will rely on the strategy-specific findings of each Evaluability Note and on the scoring results from the evaluability exercise. The results of the scoring for the three questions on risk are presented in the Table below.

Table 3.8: Frequency Distribution for Risk Questions

		Percentage scoring			
	Mean	1	2	3	4
Identification of risks					
Development Objectives (R1)	2.16	2%	80%	18%	0%
Strategy Component (R2)	1.27	73%	27%	0%	0%
Mitigation measures (R3)	1.64	36%	64%	0%	0%

1 "no content"; 2 "poor content"; 3 "adequate content"; 4 "superior content"

3.50 As can be seen all strategies complied with the minimum requirement of identifying some sort of risk for development objectives. This is reflected in the absence of virtually any scores in the '1' category. Nonetheless, further review shows that with the exception of strategy number 3 (see Annex Table B.7), risks were identified at very general levels, with no measure of degree, and with no justification based on experience, data, or other criteria. This has resulted in a concentration of scores around '2'. With respect to strategy-specific risks the results were worse. The review of strategies also revealed that risks are almost never specified with respect to strategy components, but rather at the general strategy level. This is reflected in high percentage (73%) of ratings of 'no content' for strategy component risks. In terms of measures to mitigate risks, although most strategies identify measures, they were also very general, not logically related to any development objective, and usually stated without any measure of degree or magnitude. These deficits are reflected in the high percentage of ratings (64%) in the category '2'. These issues are discussed in more detail below.

on the development of high value-added employment generating activities such as tourism and the financial services sectors."

3.51 A main problem identified in the analysis of risks was that they were not relevant to the development objectives and the strategic objectives identified in the documents. This is reflected in the preponderance of ratings in the '2' category mentioned above. This was due, in part, to the very general nature in which they are identified, and also because they focused on issues related to the inability of the Bank to secure approval for the desired operations. An example of the later was given in strategy number 6. Here the identified risks were that the required political consensus would not be achieved and that consequently the Bank's program would not be approved. As a mitigation measure the strategy indicates that the bundling of reform into a hybrid loan *"permitirá reducir los tiempos de trámites de aprobaciones al aprovechar las economías de escala programáticas que brinda la fusión de programas por ejes de actividad"*. Two comments are in order here. The first is that the mitigation measure seemed to be designed in order to reduce the risk that a Bank program not be approved; there was no connection made between this risk and any possible development risk. The second is that the principle of concentrating assets into one instrument would seem, *prima facie*, to work towards increase risk by concentrating the proposed portfolio rather than diversifying it. The strategy did not elaborate on how this mitigation measure would in fact reduce and not exacerbate the development risks involved.

3.52 In some cases the formulation of both risks and objectives were not analytically distinct, meaning that both objectives and risks referred to the same concept. This is the case of strategy number 8. Here Bank objectives were established as strategy risks themselves. Three risks were identified. One of the risks identified was that the country might be unable to increase its revenue to the Fund's target. A second risk was that the Bank might be unable to disburse. Both of these are in fact elements of the implementation of the strategy and not risks associated with the implementation. Defining strategy components as risks for their own achievement generates a tautology in which the strategy's success is determined uniquely by the realization of risks.

3.53 A second finding is that the risks were developed with no underlying empirical or analytical foundation, or at least such foundation was not evident in the documents reviewed. This follows a trend that has been observed in all of the topics reviewed above. In the strategies analyzed there was no empirical characterization of risks. This was true for every risk in every strategy. Risks are stated without any measurement of degree. In other words, all risks were treated equally, as if their potential impact on development outcomes and on the Bank's portfolio and financial position were identical. At a more fundamental level, strategies did not specify how risks would be empirically characterized. i.e. what variables would characterize the realization of risks. Also, in most cases analyzed risks identified were also not specific to the strategy's development outcomes. They were identified as risks for the strategy, where the relevant variables on which they may have incidence were almost never specified. Likewise, the lack of analysis of the Bank's past experiences in the country applies to an even greater extent regarding risk. There was no evidence, from the review of the

Strategies, that the Bank's statements regarding risk were informed by its past experience.

3.54 Lastly, it should be mentioned that the strategies did not address how risks were identified, and have likewise did not identify how mitigation measures were chosen. The review of strategies exercise did not produce evidence of (i) instances in which past Bank experience is quoted as a justification for risk identification, (ii) instances that document the potential impact of these factors on development, so as to justify their selection as a risk, (iii) instances in which the probability of the occurrence of these factors is characterized, so as to justify their identification as a risk. In the absence of evidence or methodology to allow for identification of risks it is not clear how the identified risks were relevant, and how the identified mitigation measures could be expected to mitigate the occurrence and severity of risks.

IV. CONCLUSIONS, RECOMMENDATIONS, AND THE NEW LENDING FRAMEWORK

4.1 The preceding analysis has shown that the Bank's country strategies presented over the past two years have made progress toward the improved evaluability envisioned in the *Guidelines*. It also shows, however, that the strategies comply with only the <u>minimum</u> standards set by the *Guidelines*. OVE's review found that many of the criteria established in the guidelines were not met by the recently completed group of country strategies. In specific: (i) diagnostics are of poor quality and in many cases unrelated to the Bank's strategic objectives and programmatic response, (ii) there is little systematic assessment of the effectiveness of the Bank's past engagement with the country, (iii) the strategic objectives and the strategic response are in many cases not logically related, (iv) the strategies generally do not provide an adequate treatment of risk.

4.2 Running throughout this assessment is the issue of the use of empirical evidence in the production of the country strategies. The strategies reviewed are highly general and normative, stating problems and advancing solutions that are not reliably supported by data and evidence. In other words, from the diagnostic, to the selection of objectives and activities, to the assessment of risk, the strategies do not reflect strategic choices informed by evidence, be that sector work, past Bank experience, or evidence from third parties.

4.3 OVE's findings on recent country strategies closely parallel the separate assessment of programming made by Management in its review of the lending framework in operation between 2002-2004. That review, presented to the Governors in 2005 as part of the process of developing the New Lending framework concluded that: "*…despite marked progress, there remains room for improvement in the integration of evaluation results, further development of logical frameworks and systematizing of results indicators". (3.43)*

4.4 At the same time, the New Lending Framework document calls upon Management to re-examine the *Guidelines* themselves, in order to ensure that they are appropriately designed to reflect the goals, objectives and instruments that are to be available to the Bank in the future[23]. As input to this anticipated review, the remaining section of this report will attempt to draw recommendations from the evaluability assessment of past strategies as to how the country strategy guidelines might be adapted to the programming requirements of the New Lending Framework.

A. Country Programming in the NLF

4.5 The New Lending Framework launches the Bank into a process of transforming its basic business model from one based on funding specific projects built, to a greater or lesser degree, by the staff of the Bank to one based on supporting broad programs designed and executed by the borrowing member countries. The programming process of the Bank will need to be closely re-examined in light of the changing expectations articulated in the NLF. Four aspects of the NLF are particularly important in this context.

4.6 First, the NLF places the country at the center of the programming process, with the Bank playing a largely supportive role.

> *"Countries must take a stronger role in the design, implementation and evaluation of their own development programs." 3.50*
>
> ***"****The programming system should play a key role in deciding the adequate mix of IDB instruments to be utilized by each country, taking into account its development goals, institutional progress, and financial needs." 4.54*

4.7 Programming must be "adequate" to the needs of the country, and Bank resources are "to be utilized by each country" in pursuing their objectives. Utility to the country is thus the principal objective to be attained by the programming process.

4.8 Second, to achieve the goal of utility to the country, the programming process needs to place great emphasis on in depth analysis. Recommendation 10 of the NLF directs the Bank as follows:

> *"The analytical underpinnings of Bank strategies, programs and projects should be strengthened in order to implement the recommendations presented above." 4.59*

4.9 Third, the Bank's program needs to be closely integrated with the resources of the country. The NLF instructs the Bank to "move in a direction of greater reliance on country systems with respect to budget allocation, program and project design, procurement, monitoring and evaluation. 4.40. Bank resources need to be fully

---

[23] See CA-450-1, paragraph 4.54

integrated with the budgetary resources of the country, and to that end, the NLF noted:

> *"It is also important to take into account the financial effects that the Bank's interventions are expected to have in each country as an integral part of the programming exercise. The financial effects of IDB programs and operations in borrowing countries should be an important element to be analyzed during the programming, design and execution of IDB-financed operations, both in terms of the country's net flows and debt situation, as well as the possibility of loan repayments in the longer term."4.2*

4.10 Fourth, the strengthening of country institutional capacity to budget and manage for results is seen as an integral part of the programming process.

> *"The need to strengthen public sector capacity in borrowing countries is an integral part of these developments. Countries must take a stronger role in the design, implementation and evaluation of their own development programs. Improvements in public sector management for results will place borrowing countries in a more advantageous position to secure increased flexibility in lending and to reduce transaction costs, as a result of the adoption of a programmatic approach. In this broader sense, the capacity to manage for results assumes not only stronger systems to monitor and evaluate, but also systems for budgeting, resource allocation and expenditures, procurement processes, financial management and independent auditing." 3.50*

B. Recommendations

4.11 Keeping in mind these four aspects of the NLF, this final section will try to draw out the lessons and recommendations from OVE's review of past country strategies and Guidelines to offer recommendations relevant to the programming process under the NLF.

4.12 The first recommendation concerns the role of Guidelines in the institution. The review of past practice shows that while the Guidelines approved by the Board set explicit and rather high standards for the content of country strategies, most of these high expectations were not met by the actual country strategies produced. The relatively wide gap between expectation and performance raises an important issue for the revised guidelines called for by the NLF. While some gap is likely inevitable, both Management and Board should strive to create new guidelines with expectations that are achievable. The kind of gap observed in this report serves to undermine the credibility of guidelines and serves neither Management nor the Board well.

4.13 A second recommendation relates to expectations for analytic work. OVE's review of past strategies found the diagnostics deficient, a problem that, should it persist, could undermine one of the key pillars of the NLF. Future guidelines should establish clear and reasonable standards for both the quantity and quality of analytical work required for the preparation and implementation of country strategies. Analytical work should be seen as an integral part of the program, and should be monitored and evaluated by both the country and the Bank over the course of execution.

4.14 The third recommendation relates to the nature of goals and objectives. Past strategies tended to have broad and vague goals, and often lacked specific indicators, benchmarks or targets. Under the NLF, Bank programming is designed to support country objectives, and future guidelines should require clear specification from the country on both the overall development objectives and the specific contribution expected from the Bank. These objectives should cover the entire portfolio, not merely new operations, since the largest quantitative contribution of the Bank will come from disbursements on the portfolio rather than on new approvals. This is the key recommendation for making future strategies evaluable.

4.15 The fourth recommendation relates to treatment of resource flows in future country strategies. OVE's review found only 11% of strategies scoring adequately on the analysis of these issues. Given the importance assigned to integrating Bank resource flows with those of the country under the NLF, future guidelines should probably contain a specific template for the presentation and analysis of resource flow and budgetary data in order to ensure that this critical dimension to programming is managed adequately.

4.16 This leads logically to a fifth recommendation regarding the treatment of risk. As the Bank moves to implement the NLF demand to rely more on country systems for programming and execution, the issue of risk assessment and mitigation assumes even greater importance. The Bank should be encouraged to accept the risks of reliance on country systems, but the review of past strategies shows a system for dealing with risk that places insufficient emphasis on the elaboration of risk or the intervention to mitigate risks once they appear. A significant strengthening of risk management and mitigation should be a key component of future guidelines.

4.17 Finally, the past Guidelines established clear expectations with regard to the assessment of institutional capacity, yet the OVE review found that most ratings on this dimension scored in the "poor" range. Given the emphasis on institutional capacity in the NLF, and the resources made available by the Bank under the PRODEV initiative, PRODEV should be thoroughly integrated into the country strategy formation process, and PRODEV interventions should be evaluated as part of the overall assessment of the Bank's programming activity.