

10. Effectiveness (or Efficacy)

Principles and Norms

DEFINITION

10.1 *Effectiveness (or efficacy)* is the extent to which the program has achieved, or is expected to achieve, its objectives, taking into account their relative importance.⁵⁵ Shortcomings in the achievement of objectives have to do either with the *number* of objectives that have not been achieved (or are not expected to be achieved) or with the *extent* to which one or more objectives have not been achieved (or are not expected to be achieved). Positive unintended results may also be regarded as additional achievements if convincingly documented.

Based on DAC Glossary and IEG evaluation criteria

NEED FOR GRPP EVALUATIONS TO ASSESS EFFECTIVENESS

10.2 All GRPP evaluations need to include an assessment of the effectiveness of the program in order to demonstrate to stakeholders (a) the degree to which the original objectives are being met, (b) whether the program should adjust or restate its objectives or strategies to reflect changing circumstances, or (c) whether the program needs to put in place additional safeguards or compensatory measures to mitigate any negative unintended results. Depending on the findings of the assessment, the governing body may wish to consider expanding the program or increasing its reach, changing its geographical coverage, devolving some its activities, or even phasing out some or all activities. An assessment of effectiveness is also important to provide accountability to the international community. Given scarce development aid and many alternative uses for constituent taxes and other resources, the evaluation should compare the achievement of the program's objectives not only to the original expectations but also, to the extent possible, to the outcomes from alternative uses of resources.

Draws on IEG's experience with reviewing GRPPs

55. As noted in the OECD/DAC *Glossary of Key Terms in Evaluation and Results Based Management*, 2002, "effectiveness" is also used as a broader, aggregate measure — encompassing relevance and efficiency as well — of the overall outcome of a development intervention such as a GRPP. This chapter uses the term "effectiveness" in the narrow sense, which is synonymous with the use of the term "efficacy" in a number of development organizations such as the World Bank.

Draws on IEG's
experience with
reviewing
GRPPs

Standards and Guidelines

OBJECTIVES-BASED ASSESSMENT

10.3 The evaluation should first assess the achievement of the stated objectives — objective by objective — and the extent to which each objective has been achieved (or is expected to be achieved). The evaluators should also determine if the program has unstated objectives, since the objectives often differ from the perspective of different partners and other stakeholders, and since objectives are dynamic and change over time. (See also paragraph 9.7 under relevance.)

10.4 The ability to undertake a systematic assessment of the achievement of each objective will depend on the maturity of the program and the existence of a good monitoring framework, including a structured set of qualitative or quantitative input, output, outcome, and impact indicators. When the program is young (less than four years old), it will be more difficult to make a summative assessment of effectiveness. (See also paragraph 6.7.) When the program has not established a good monitoring framework, the evaluators could provide guidance to the secretariat in establishing one. (See also paragraphs 2.22–2.34 on establishing an M&E framework for GRPPs.)

UNINTENDED OUTCOMES

10.5 The assessment of effectiveness should not be limited to the achievement of expected outputs and outcomes, but should also cover unintended outcomes, whether negative or positive. These would include the unintended results of the program's activities as well as of the partnership itself, such as any harmonization of procedures or effects on aid coordination outside of the partnership itself.

10.6 The assessment of effectiveness should also include how the objectives and strategies of the program have evolved in response to (a) learning from experience or (b) the risks and opportunities arising from a new external environment, technology, or emerging target group. For instance, it may become important for the program to provide compensatory measures if unintended negative outcomes are occurring in relation to the program's safeguard objectives.

EVIDENCE-BASED CONCLUSIONS

10.7 The assessment of the achievement of objectives, and of other unintended results, should be evidence-based. Evidence-based conclusions distinguish an evaluation report from an expert consultant report, which is based primarily on expert judgments. Evidence-based conclusions and internal consistency among findings based on more than one type of evidence — or triangulation — have the added benefit of helping to ensure independence, regardless of organizational ar-

rangements. The ability to provide evidence-based conclusions depends on the use of measurable indicators, as laid out below.

THE NEED TO MEASURE INPUTS, THE PROGRESS OF ACTIVITIES, OUTPUTS, OUTCOMES, AND IMPACTS TO THE EXTENT POSSIBLE

10.8 An evaluation should measure inputs, the progress of activities, outputs, outcomes and impacts to the extent possible (or an appropriate rationale should be given as to why not). Findings regarding inputs should be distinguished clearly from those regarding outputs, outcomes, and impacts. Outcomes and impacts should include any multiplier or downstream effects attributable to the GRPP and – as noted above – any unintended effects, whether positive or negative. To the extent possible, each of these should be measured either quantitatively or qualitatively and compared to benchmarks.

Based on UNEG Standard 4.12, paras. 19 and 20

10.9 In addition to quantitatively measurable inputs, such as budgets and staffing, the assessment should also consider other causal factors that have an effect on the progress of activities, outputs, and outcomes, such as changes in the location, the legal structure, or the governance processes of the program during the time period of the evaluation.

10.10 For GRPPs, it is also important to measure the program's inputs, progress of activities, outputs, outcomes, and impacts at all levels – global, regional, national, and local – and to find a way to present in summary form the results from the local and national levels and the way in which they affect results at the regional and global levels. A simple aggregation of results may not be ideal if this obscures causal relationships. It is better if the results are presented in a way that highlights the factors that have influenced success or failure in a variety of conditions.⁵⁶

10.11 In addition, outcomes related to the unique contribution of the partnership itself – such as the scale or joint activities made possible by its organizational setup as a GRPP, or its institutional linkages to a host organization – should be measured and assessed. What is the value added of the GRPP relative to what could have been achieved by intervening only at the country or local level, taking into consideration the leadership of the partnership, the roles and responsibilities of the various partners, and the degree of trust developed among the partners?

56. Sometimes ratings are used to facilitate aggregation of activity results to the country, regional, or global level. If ratings are used for such purposes, it is essential to distinguish ratings of performance (such as effort and inputs) from ratings of results.

10.12 The M&E system of a GRPP needs to be able to take into account the evolving nature of its portfolio. It is important that organizations or individuals proposing activities for financing at the country or local level not only list expected outcomes in their proposals but also link them to measurable indicators, so that GRPP management can take steps to incorporate these indicators into the program's M&E system to facilitate the later assessment of the effectiveness of the interventions.

SPECIAL CONSIDERATIONS IN ASSESSING THE EFFECTIVENESS OF GRPPS

Draws on IEG's
experience with
reviewing
GRPPs

10.13 To assess effectiveness, an evaluation of a GRPP must first attempt to define the *boundaries of the program's impact*, which may be difficult, particularly if these are expected to change over time as the program grows in scale or reach, or if these vary by activity. Defining the boundaries of the potential impact of environmental program activities can be particularly relevant, but the same is also true for programs providing social services, which may have the potential to serve a large population. For example, an assessment of the effectiveness of pilot health service interventions may need to consider both the outcome of the actual pilot, with its limited scale, and also the degree to which it yields useful information on the likelihood of its success under alternative conditions or at larger scale.

10.14 For many mature programs, the *large scale* of the program itself presents complications in assessing effectiveness. Choice of a representative sample of activities becomes very important. The diversity of country conditions that need to be captured may be larger, and finding an appropriate modality for presenting diverse results may be a challenge — going beyond mere aggregation and capturing the different factors affecting success and failure. The use of ratings, which may facilitate aggregation, may not be appropriate if the basis for the ratings is not articulated or understood, or if the raters are diverse.

10.15 That GRPPs typically support activities *at different levels* may create complexities if the objectives of stakeholders at different levels are different, or even in conflict. For instance, the global/regional public goods benefits of some environmental actions may be associated with disproportionate costs relative to benefits for some implementing countries. Thus, it is important to indicate from whose perspectives the results are being assessed and to assess trade-offs of costs and benefits to the various stakeholders.

10.16 Unlike projects, GRPPs are programmatic and typically have no fixed end-point. Many of the expected results have a longer *time frame* than that of the interventions that contribute to achieving the results. One needs to consider not only the joint outcomes of global/country/local interventions, but also the “joint outcomes” or

cumulative effects of different interventions over time. To properly assess this, the program needs to have established a baseline and to have put in place arrangements to gather information at specific times in order to assess long-term results.

10.17 GRPPs have *more stakeholders* and more diversity among stakeholders than country and local-level programs and projects. Hence, there are more perspectives on the achievement of results and objectives that need to be taken into account. Again, it is very important for evaluators to obtain a representative sample of views and survey responses. (Evaluators should schedule interviews not just according to convenience or availability.) And they should always disclose the criteria that they used for selecting interviewees or survey respondents.

10.18 Finally, GRPPs differ from other programs and projects because they have *distinct governance mechanisms and processes* that affect results. It is important to regard the way in which these mechanisms and processes work in practice, as well as any changes in them over time, as a part of the results chain. For instance, any of the following can affect the achievement of results and objectives:

- Interruptions in the continuity of key management positions or of members of the governing body
- Changes in the frequency of governance meetings or in the types of decisions handled by the governing body, as opposed to management
- The processes for allocating resources and choosing activities to support
- Changes in the resource mobilization strategy that affect the scale of the program and, if there is earmarking, the allocation and use of funds
- The influence of host organization representatives or the need to comply with their policies.

ASSESSING EFFECTIVENESS OF DIFFERENT TYPES OF PROGRAMS

10.19 GRPPs support diverse types of activities. (See also paragraph 9.14.) Each type of activity (networking, advocacy, knowledge creation, technical assistance, or investments) presents methodological challenges with respect to the assessment of effectiveness. This is because the different types of activities contribute in different ways to the program's value added and leverage on domestic policy and institutional reform, human resource capacity, and total investments in the sector — as well as to other objectives such as poverty reduction and improvements in welfare.

10.20 For each type of activity, the program should have articulated a results chain or logframe with clear, agreed-upon indicators that allow evaluators to attribute results to the program.⁵⁷ (See also paragraphs 2.33, 2.34, 9.5, and 9.6.) The evaluators should aim to capture the distinct contributions of each type of activity toward the achievement of the program's objectives so that the objectives can also be adjusted to increase the program's impacts over the long term. Evaluators also need to understand the way in which the objectives are being achieved – whether through command and control within bureaucracies, through voluntary exchange in markets, through common interest in collective action, or through some combination of these. (See also paragraphs 9.15–9.18.) Unintended outcomes on markets or prices, such as “crowding out a market” or “catalyzing a market,” should be noted.

Draws on IEG's forthcoming review of regional programs

10.21 For regional partnership programs, it is also important to assess the distribution of the benefits and costs of the program among the beneficiary partners. Experience has shown that an inequitable distribution of net benefits can adversely affect the sustainability of the program.

10.22 Where feasible, evaluations should assess final welfare outcomes in relation to a counterfactual in order to isolate the effects of the program on those outcomes. This would include assessing how the outputs of the program have supported enhanced welfare outcomes in the sector and country in which the GRPP is operating. (See also Chapter 15, Impact Evaluation.)

ASSESSING LINKAGES BETWEEN GRPPS AND COUNTRY OR LOCAL-LEVEL ACTIVITIES

Draws on IEG's experience with reviewing GRPPs

10.23 For GRPPs, it is important to assess the effectiveness of their operational linkages with country or local-level activities, whether or not the latter are supported by donors. For most GRPPs, positive outcomes and impacts at the country or local level are a joint product of both the GRPP and country or local-level activities.

10.24 Two types of linkages need to be assessed: (a) opportunities for direct linkages that are subject to the control of participants at both levels and (b) effects that may operate through markets or the behavior of agents external to the program, which may require a strengthening of safeguard measures or other compensation. The linkages in

57. However, much work still needs to be done to develop generic indicators for generic-type activities that are common to GRPPs such as advocacy, improving donor coordination, knowledge generation and dissemination, supporting national-level policy and institutional reforms, and capacity strengthening.

(a) and (b) may result in unambiguous win/win outcomes, or may bring to light trade-offs that need to be taken into account in considering the net benefits to different partners and participants.

10.25 With regard to the first set of linkages – which can be improved through conscious action – linkages in both directions are important. First, country and beneficiary representatives (whether public or private) need to have an effective means of communicating their constraints, requirements, and priorities to GRPP management, thereby potentially increasing the relevance, focus, ownership, and outcomes of the GRPP. The means of communication may be through direct participation on the governing body (important, but not always easy in practice),⁵⁸ through other periodic consultation mechanisms like workshops, through GRPP procedures that solicit proposals for assistance (and provide help in shaping them to be successful), and through deliberate exchange and discussion of government planning documents and donor assistance strategies relevant to the country or local beneficiaries. An evaluation of a GRPP should always assess the effectiveness of these various means of communications.

10.26 In addition, it is important to assess the actual outcomes and impacts of the GRPP activities on country or local-level priorities, activities, and deployment of human resources. To the extent possible, the benefits of participating in the GRPP – from the perspective of the beneficiary groups – should be compared with the costs, including increased reporting and compliance requirements and other demands on senior skilled implementers (whether public or private). Such an assessment of the opportunity costs of the participants' time and resources should ideally include participatory methods to directly obtain information on beneficiary group satisfaction, complaints, and suggestions for change.

10.27 An assessment should also be made of how well the GRPP acts on the information it obtains from beneficiary groups – both upfront information on needs and priorities that might influence strategy or allocation of funds, and periodic feedback that would provide opportunities for improving the outcomes and impacts of the program at the country or local level. GRPP management needs to fashion the program's support to add value to country or local-level activities by contributing new knowledge and technologies, facilitating exchange of good practice among beneficiary groups, and helping to mobilize additional resources or to channel existing resources to more productive activities. They should seek to ensure alignment of the country or local activities they support with country or local-level

58. See Chapter 12, Governance and Management, for treatment of the issue of including beneficiary groups in governing bodies.

plans and budgeting priorities and to ensure that complementary country or local-level inputs are available to make GRPP interventions effective. Regional partnership programs – which are often focused on specific cross-border issues and whose success is typically more dependent on country or local-level commitment and capacity – may require special attention to ensure that program priorities and requirements do not hamper the achievement of the rest of the countries' development agendas.

10.28 GRPP management needs to examine periodically the degree to which the GRPP activities and outputs (which may be inputs to country or local-level activities) are relevant to the needs of final beneficiary target groups. Linkages between the GRPP and donor representatives in decentralized country or local-level units are useful but not sufficient; direct dialogue with country or local implementers and beneficiary groups is also needed. At a minimum, this needs to be done through wide dissemination of monitoring reports, annual reports, and evaluations of the GRPPs to all existing and potential beneficiary groups or local implementers. Even better would be active dialogue to solicit the views of beneficiary groups on the responsiveness of the GRPP activities and outputs to their needs.

10.29 Unintended outcomes that operate through effects on trade, commercial markets, or the behavior of agents external to the planned program results may also need to be assessed. Ideally, the potential results would have been identified in the results chain or logframe in the planning stages of the program, thereby facilitating monitoring of such results. Alternatively, participatory methods of evaluation can identify cases where results are perceived to be due to the GRPP, and methods could be devised in the evaluation to test these hypotheses and recommend compensatory adjustments or additional safeguard measures, if needed.