

# PART III

## **Monitoring and Evaluating Performance**

<b>Chapter 6. Performance Measurement</b>	<b>63</b>
A. Performance measurement	63
B. Selecting indicators	66
C. Using indicators	71

**Figures, Tables and Boxes for Part III**

Figure 1. Dimensions of performance assessment	64
Figure 2. Progress towards results	72
Table 1. How to select indicators	68
Table 2. How indicators might fit into an outcome monitoring plan	70
Table 3. Indicators and the monitoring of results	72
Box 1. Key principles for data collection and analysis using indicators	71

# Performance Measurement

This chapter covers methods used in performance measurement. It introduces the use of indicators, including use of baseline data, setting targets, data collection systems and quantitative and qualitative analysis. The chapter should help users to apply indicators in ways that enhance the ability to judge progress towards results and performance when monitoring and evaluating UNDP-supported assistance.<sup>13</sup>

This chapter covers:

- A. Performance Measurement
  - Rating system
- B. Selecting Indicators
  - Key steps in selecting indicators
  - Indicator planning
- C. Using indicators
  - Involving stakeholders
  - Using indicators for monitoring

## A. Performance Measurement

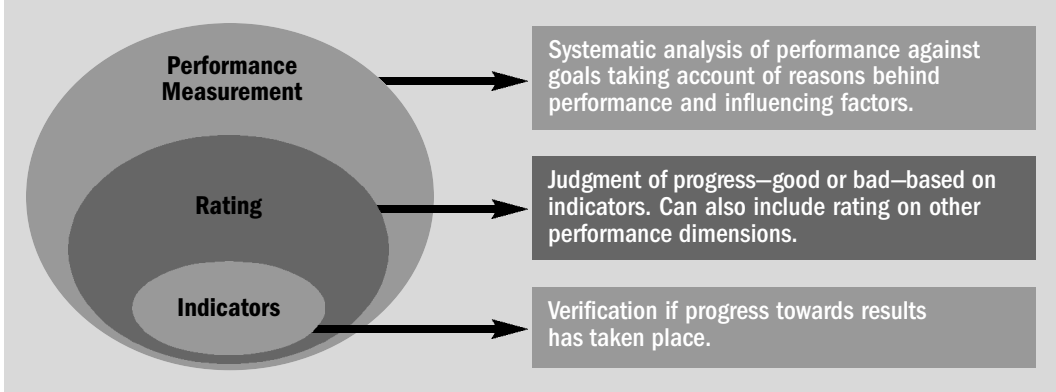
Indicators are part of performance measurement but they are not the only part. To assess performance, it is necessary to know about more than actual achievements. Also required is information about how they were achieved, factors that influenced this positively or negatively, whether the achievements were exceptionally good or bad, who was mainly responsible and so on.

Traditionally, it has been easier to measure financial or administrative performance, such as efficiency. Results-based management today lays the basis for **substantive** accountability and performance assessment or effectiveness. The APR, evaluations and the ROAR provide the means to assess performance at the country office level.

---

<sup>13</sup> Indicators of performance for corporate planning within UNDP are outside the scope of this *Handbook*.

**FIGURE 1. DIMENSIONS OF PERFORMANCE ASSESSMENT**



In any given country office, results-based management may also inform the assessment of performance of projects, programmes, programme areas, groups of staff and individuals—should the CO find it useful. Figure 1 illustrates the linkages between performance measurement, rating and indicators as elements of performance assessment.

## **RATING SYSTEM**

The growing internalization of RBM within UNDP is gradually allowing for an expanded use of reported performance results for internal management and oversight functions. A key area for such expanded use involves the development of a **common rating system** for all results reported by the organization under the RBM framework. Such a system allows UNDP to rate performance at the results level, and to analyze and compare trends by thematic category (e.g., governance, poverty or environment); level of intervention (e.g., project, output or outcome); geographic area (e.g., Africa, Asia or Latin America); or organizational unit (e.g., country offices, regional bureaux or functional bureaux). A common rating system would build on the three-point rating system developed in the ROAR.<sup>2</sup>

A common rating system may be used for all key monitoring and evaluation tools (the ROAR, the APR, field visits and outcome evaluations) to compare performance across results. With this approach, there are two kinds of ratings: self-ratings (as in the APR) and independent ratings (as in outcome evaluations). Having two kinds of ratings that use the same rating criteria allows a richer picture of how progress towards results is perceived by the UNDP CO, by UNDP HQ and by independent assessors. It also provides the basis for dialogue within and between the CO and HQ if ratings for the same outputs or outcomes vary.

### **Key elements of the common rating system:**

**For outcomes**, the rating system has three points: positive change, negative change and unchanged (no change). The three ratings reflect progress on outcomes, without

<sup>14</sup> A three-point rating system (positive change, negative change, unchanged) for outcomes and (achieved, not achieved, partially achieved) for annual output targets was originally developed for the ROAR. This system allowed UNDP to determine output and outcome performance rates and compare them across countries, regions, goals, sub-goals and so on. It was completed by analysts at HQ, was not typically shared with country offices and was not compared with country office rating of the same changes.

attributing the progress to any partner. The three ratings are meant to reflect the degree to which progress has been made towards or away from achieving the outcome. The methodology in all three ratings is to compare, as measured by outcome indicators, the evidence of movement from the baseline either towards or away from the end-SRF target.

- Positive change—positive movement from baseline to SRF target as measured by the outcome indicator(s);
- Negative change—reversal to a level below the baseline as measured by the outcome indicator(s);
- Unchanged—no perceptible change between baseline and SRF target as measured by the outcome indicator(s).

**For outputs**, the rating system also has three points: no, partial and yes. The three ratings reflect the degree to which an output’s targets have been met. This serves as a proxy assessment of how successful an organizational unit has been in achieving its SRF outputs. The three ratings are meant to reflect the degree of achievement of outputs by comparing baselines (the inexistence of the output) with the target (the production of the output). The “partially achieved” category is meant to capture those en route or particularly ambitious outputs that may take considerable inputs and time to come to fruition.

- No—not achieved;
- Partial—only if two-thirds or more of a quantitative target is achieved;
- Yes—achieved.

This results rating system applies to the ROAR, to evaluations and to select country office reports.

The **ROAR** rates progress towards outputs and outcomes. The rating is made annually by a ROAR analyst team at Headquarters. The ratings are used to report trends and progress to external stakeholders, as well as to identify weak areas for improvement by the organization. Feedback to the country offices on the ratings should stimulate debate on progress and needed action. Because country offices will receive information via other sources—outcome evaluations and key monitoring tools—using the same ratings, they will have a basis to discuss results, and perceptions of the progress towards them, with Headquarters.

**Evaluations**, at a minimum, rate outcome and output progress. Outcome evaluations, which are undertaken by independent assessment teams, may also rate key performance dimensions such as sustainability, relevance and cost-effectiveness. Other types of assessments should also provide ratings where appropriate, such as assessments of development results by the Evaluation Office. The ratings will be used for trend analysis and lessons learned corporately, as well as for validation of country-level results and debate on performance at country level. (See *Guidelines for Outcome Evaluators* in the companion series to this *Handbook* for more on the rating system to be used by outcome evaluators.)

Selected **country-level monitoring reports** rate outcome and output progress for projects, on a voluntary basis. For the Annual Project Report (APR), the rating on progress towards outputs is made annually by the Project Manager and the Programme Manager. It forms the basis of a dialogue in which consensus ratings for the **outputs** are produced. If there is disagreement between the project and programme staff on how

**outputs** are rated, both ratings are included in the report, with proper attribution. The rating on progress towards outcomes in the APR is made by the Programme Manager and/or other country office staff. For field visits, UNDP Programme Managers and/or other CO staff periodically rate progress towards both **outputs and outcomes**, discussing their ratings with the project staff. The ratings are used by the CO to assess project performance and for trend analysis and lessons learned. They may also be used corporately for validation and lessons learned.

For an effective use of performance assessment and rating, the country office may want to:

- Debate and determine the use that will be made of the results to be obtained from the performance assessment system, such as analyzing trends, targeting projects for improvement, contributing to RCA ratings and extending the contracts of project staff;
- Share information with the relevant CO and project staff about how the rating system will be used so that there are no false expectations regarding its application;
- Determine organizational unit and individual responsibility for performance assessment and rating as well as control measures to mitigate biases by “raters”;
- Qualify quantitative rankings with qualitative performance assessments.

(See *Guidelines for Outcome Evaluators* in the companion series to this *Handbook* and the Annexes for more detailed discussion of the use of rating for relevant monitoring and evaluation tools—specifically, the APR, the field visit and the outcome evaluation.)

## B. Selecting Indicators

Much of the following information is more relevant to Programme Managers and Senior Country Office Managers than it is to Project Managers, who deal mainly with output indicators for which baselines, targets and information gathering are fairly clear-cut. Programme Managers and Senior CO Managers are concerned with more complex indicators that reflect progress towards outcomes. Some key steps for UNDP managers in working with indicators are outlined below.

### KEY STEPS IN SELECTING INDICATORS

**Set baseline data and target:** An outcome indicator has two components: a baseline and a target. The **baseline** is the situation before a programme or activity begins; it is the starting point for results monitoring. The **target** is what the situation is expected to be at the end of a programme or activity. (Output indicators rarely require a baseline since outputs are being newly produced and the baseline is that they do not exist.)

- If wider access to education is the intended result, for example, school enrollment may provide a good indicator. Monitoring of results may start with a baseline of 55 percent enrollment in 1997 and a target of 80 percent enrollment in 2002.

Between the baseline and the target there may be several **milestones** that correspond to expected performance at periodic intervals.

Baseline data provides information that can be used when designing and implementing interventions. It also provides an important set of data against which success (or at least change) can be compared, thereby making it possible to measure progress towards a result. The verification of results depends upon having an idea of change over time. It requires a clear understanding of the development problem to be addressed—before beginning any intervention. A thorough analysis of the key factors influencing a development problem complements the development of baseline data and target setting.

**What to do when no baseline is identified?** A baseline may exist even when none was specified at the time a programme or project was formulated. In some cases, it may be possible to find estimates of approximately where the baseline was when the programme started through the CCA and annual review exercises and national administrative sources.

- For example, implementation of a local governance project already has begun but no baseline data can be found. It still may be possible to obtain a measure of change over time. Ask a number of people: “Compared to three years ago do you now feel more or less involved in local decision-making?” A clear tendency among respondents either towards “more” or towards “less” provides an indication of whether or not change has occurred.

Sometimes it is not possible to ascertain any sense of change. In this case, establish a measure of the current situation so that an assessment of change may take place in the future. Refer to the project document (or PSD) for more information about context and problems to be resolved.

**Use proxy indicators when necessary:** Cost, complexity and/or the timeliness of data collection may prevent a result from being measured directly. In this case, proxy indicators may reveal performance trends and make managers aware of potential problems or areas of success. This is often the case for outcomes in policy dialogue, governance and other results that are difficult to measure.

- For example, the outcome “fair and efficient administration of justice” is often measured by surveying public confidence in the justice system. Although high public confidence does not prove that the system actually is fair, there is very likely a correlation. In another example, in an environmental protection programme where a target result is the improvement in the health of certain lakes, the level of toxins in duck eggs may serve as a proxy indicator of that improvement.<sup>15</sup>

**Use disaggregated data:** Good indicators are based on basic disaggregated data specifying location, gender, income level and social group. This is also necessary for good project and programme management. Such information, sometimes in the form of estimates, may be drawn from governmental and non-governmental administrative reports and surveys. Regular quality assessments using qualitative and participatory approaches may be used to corroborate, clarify and improve the quality of data from administrative sources.

- For the outcome “effective legal and policy framework for decentralization”, for example, the indicator “proportion of total public revenues allocated and managed

<sup>15</sup> Allen, John R., “Performance Measurement: How to Do It, How to Use It,” paper presented at a workshop sponsored by the American Evaluation Association, Atlanta, November 1996.

at sub-national level” may demonstrate an increased overall distribution of resources to the local level but hide large disparities in distribution to some regions.

**Involve stakeholders:** Participation should be encouraged in the selection of both output and outcome indicators. Participation tends to promote ownership of, and responsibility for, the planned results and agreement on their achievement. A preliminary list of **output** indicators should be selected at the project formulation stage, with the direct involvement of the institution designated to manage the project and with other stakeholders. Partners are involved in the selection of **outcome** indicators through the Country Programme and SRF formulation processes. It is important that partners agree on which indicators to use for monitoring and on respective responsibilities for data collection and analysis. This establishes a foundation for any future changes in the implementation strategy should the indicators show that progress is not on track.

**Distinguish between quantitative and qualitative indicators:** Both quantitative and qualitative indicators should be selected based on the nature of the particular aspects of the intended result. Efficiency lends itself easily to quantitative indicators, for example. Measuring dynamic sustainability, in contrast, necessitates some qualitative assessment of attitudes and behaviors because it involves people’s adaptability to a changing environment. Methodologies such as beneficiary assessment, rapid rural appraisal (RRA) and structured interviews may be used to convert qualitative indicators into quantitative indicators.

**Try to limit the number of indicators:** Too many indicators usually prove to be counter-productive. From the available information, develop a few credible and well-analyzed indicators that substantively capture positive changes in the development situation. The CO needs to select from among a variety of indicators since several projects may contribute to one strategic outcome. Be selective by striking a good balance between what **should be** and what **can be** measured. Narrow the list using the SMART principles and additional criteria to sharpen indicators. See Table 1 for a selection table developed by UNDP Mali.

**TABLE 1. HOW TO SELECT INDICATORS**

INTENDED RESULTS	PERFORMANCE INDICATORS	CLASSIFICATION OF INDICATORS						TOTAL SCORE	SELECTED
		A	B	C	D	E	F		
Impact	If any -								
Outcome 1	Indicator 1 Indicator 2...	[Rate 1 per satisfied criteria]							
Output 1	Indicator 1 Indicator 2...								

Select the 2 to 3 indicators with best score

**A = the meaning of the indicator is clear**  
**B = data are easily available**  
**C = the effort to collect the data is within the power of the project management and does not require experts for analysis**  
**D = the indicator is sufficiently representative for the total of the intended results (outcome or output)**  
**E = the indicator is tangible and can be observed**  
**F = the indicator is difficult to qualify but so important that it should be considered (proxy indicator)**

**Ensure timeliness:** The usefulness of an indicator depends on timeliness and clear actions so that an indicator target date corresponds to the expected progress of the assistance. If changes take place, such as the modification of outputs or outcomes, new sets of indicators would need to be established to reflect the actual targets.

## INDICATOR PLANNING

A critical test of an indicator is its practicality in monitoring results—that is, how easy it is to obtain and analyze data for the indicator. Obtaining “valid” and “representative” data can become a large, complex and costly undertaking. Asking everybody’s opinion through a survey is rarely possible, for example, and it would be easier to obtain a smaller but representative sample of respondents. Indicators should therefore be as simple and few as possible, while demonstrating some measure of progress or magnitude of change.

**It will be difficult to understand or analyze the indicators if they are too complex.**

Responsibility for data collection and analysis belongs primarily to the Programme Manager, who consults when necessary with the senior CO management and the project management. For **outcomes**, the Programme Manager’s responsibilities are:

- To ensure that baseline or situational data for outcomes is available at the programme formulation stage. This will allow time-series data to be compared with the baseline data to assess progress in achieving the desired outcomes. Use existing sources and processes such as the NHDR and the CCA. The majority of outcomes will be identified at the Country Programme or SRF formulation stage. Some, however, may assume a slightly different formulation in specific programmes and projects. Some outcomes may also feature slightly different and/or additional indicators than those found in the SRF. The priority is to ensure that adequate baseline or situational data is available at the time the outcome and the indicators are selected, regardless of their specificities.
- To ensure that data collection and analysis is planned for the outcome. A sample table to aid planning is illustrated in Table 2. Rely on existing national sources rather than collection by the CO or the project, to the extent possible. For the outcome “responsiveness of national policies to gender issues”, for example, the indicator “government policy statements and plans include targets from the improvement in the status of women” is measurable. However, someone would have to collect all the plans and then go through them to see if targets are set. This would have to be planned in advance.
- To ensure that efforts are made to obtain information from beneficiaries on outcomes. This might be accomplished through household surveys, individual surveys and consultations with members of the community and stakeholders.

For **outputs**, the Programme Manager’s responsibilities are:

- To ensure that data is collected through the projects themselves and that the project management is aware of its responsibilities. Data should also be collected from various administrative sources and national systems.

An **outcome monitoring plan** drawn up by the programme management helps to ensure that the CO continuously collects information on the outcome and periodically analyzes it to chart progress, as discussed in Chapter 3. Country offices that establish an outcome monitoring plan must be sure to include information on the outcome indicators so that (a) indicators are not selected unless data is likely to be available and (b) data will be available for the indicators that are selected.

It is recommended that an outcome monitoring plan, should a country office decide to use one, identify the following:

- Outcome to which the indicators apply;
- Outcome indicator(s);
- Data source(s);
- Method and frequency of collection and analysis.

The plan may also identify who will be responsible for data collection and analysis and who will use the resulting information. Table 2 provides a sample format that suggests how indicators might fit into an outcome monitoring plan.

**TABLE 2. HOW INDICATORS MIGHT FIT INTO AN OUTCOME MONITORING PLAN<sup>16</sup>**

OUTCOME	OUTCOME INDICATOR(S)	DATA SOURCE(S)	METHOD OF DATA COLLECTION/ ANALYSIS	FREQUENCY OF DATA COLLECTION/ ANALYSIS	WHO IS RESPONSIBLE	WHO WILL USE THE INFORMATION

Key generic principles and issues to be addressed when Programme Managers plan data collection and analysis using indicators are listed in Box 1. Simpler indicators, as noted above, are easier to understand and analyze.

When planning data collection and analysis using indicators, Programme Managers may realize that data are not immediately available. In such cases, Programme Managers should plan to collect data through alternative instruments and/or approaches, such as:

- Awareness/attitude surveys and questionnaires;
- Expert panels;
- Key informant interviews;
- Focus groups;
- Mapping techniques.

Attitude surveys allow for some quantification of qualitative change. For example, the proportion of people who perceive of local government management as “participatory” has gone up from 40 percent to 65 percent over a certain period of time. This statistic provides some measure of the **degree** of qualitative change.

<sup>16</sup> USAID Performance Monitoring Plan, TIPS No. 7

## BOX 1. KEY PRINCIPLES FOR DATA COLLECTION AND ANALYSIS USING INDICATORS

### **Rationale**

Build a conceptual framework into which ideas can be placed, giving definition, orientation and direction to available information and to your own ideas about the project or programme. For UNDP, this means a credible chain of results in terms of outcome, outputs and partnerships.

### **Statement of purpose or problem**

What is it that you plan to investigate? Define problems and issues, look for signals that have clear meaning, establish agreements with beneficiaries and partners. This involves the definition of the development problem at the SRF Sub-Goal level.

### **Questions to be answered**

When the data collection is finished, what are the major questions to which reasonable answers can be expected? What evidence can be used to verify the level of achievement of the indicator?

### **Statement of outcome**

Spell out the particular outcome(s) you will monitor, set targets that are realistic and keep in mind data sources and monitoring responsibilities.

### **Design and procedure**

State who will be the subjects of your interviews, surveys and focus groups and describe how they will be selected. Explain the conditions under which the data will be collected, what measuring instruments or data-gathering instruments will be used, and how the data will be analyzed and interpreted. Look for data that is easily available and avoid major data collection.

### **Assumptions**

What assumptions have you made about the nature of the issues you are investigating, about your methods and measurements, or about the relations of the investigation to other problems or situations?

### **Limitations**

What limitations exist in your methods or approach to internal and external validity?

### **Delimitations**

On what basis have you narrowed the scope of data collection and analysis? Did you focus only on the selected aspects of the problems or outcome, certain areas of interest or a limited range of subjects?

### **Definition of terms**

List and define the principal terms you will use, particularly where terms have different meanings to different people. Emphasis should be placed on operational and/or behavioral definitions.

## C. Using Indicators

### INVOLVING STAKEHOLDERS

The country office—specifically, the Programme Manager in direct consultation with the senior CO management—should establish mechanisms for sharing information generated from indicators with primary stakeholders. This is particularly true for outcome indicators. This ensures that the analysis of progress is locally relevant using local knowledge, while fostering “ownership” and building group decision-making skills. It is worth noting, however, that stakeholder or partner participation in the analysis of the indicator data may significantly alter the interpretation of that data.

Participatory observation and in-depth participatory reviews with implementation partners and beneficiaries are integral to visual on-site verification of results—which is a reliable form of assessment. More “top down” and less participatory approaches to assessment may be used to achieve analytical rigor, independence, technical quality, uniformity and comparability. Ultimately, of course, the information gained through the analysis of indicators feeds into evaluations. This data helps assess progress towards outputs and outcomes, and includes a measure of stakeholder satisfaction with results.

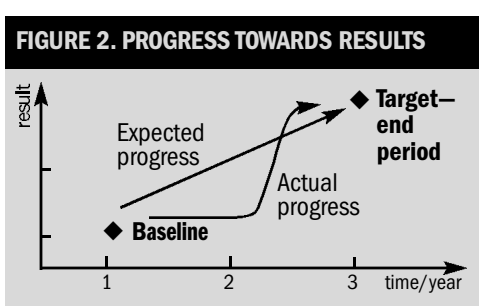
## USING INDICATORS FOR MONITORING

Results-oriented monitoring of development performance involves looking at results at the level of outputs, outcomes and, eventually, impact. Table 3 shows how indicators are used for each type of result and who is the primary user of the indicator.

TABLE 3. INDICATORS AND THE MONITORING OF RESULTS			
TYPE OF RESULT	WHAT IS MEASURED	INDICATORS	PRIMARY LEVEL OF USE
<b>Output</b>	Effort, or goods and services generated by projects and programmes	Implementation of activities	Project Management
<b>Outcome</b>	Effectiveness, or results in terms of access, usage and stakeholder satisfaction from goods and services generated by projects, programmes, partners and soft assistance	Use of outputs and sustained production of benefits	Programme Management
<b>Impact</b>	Effectiveness, or results in terms of the combined effect of a combination of outcome activities that improve development conditions at a national level	Use of outcomes and sustained positive development change	Senior CO Management

Indicators are used periodically to validate partners' perceptions of progress and achievement, to keep projects and programmes on track and to provide early warning signals of problems in progress. Indicators only **indicate**; they do not **explain**. Any interpretation of indicators is done through qualitative analysis. As indicated in Figure 2, qualitative analysis is needed to interpret what the indicators say about progress towards results.

**For output indicators**, the Programme Manager uses day-to-day monitoring to verify progress, as well as field visits and reports and/or information received from the project management. The Annual Project Report (APR) is too infrequent to allow early action in case there are delays or problems in the production of outputs.



**Note:** Progress is rarely linear. In this example, if one were to look at the indicator at the beginning of year 2, it would reveal that no progress has been made against expectations. Later, if one were to look at the indicator again at the end of year 2, it would reveal substantial progress even beyond expectations. Analysis is required to explain why progress was slow for the first year but picked up dramatically during Year 2.

**For outcome indicators**, annual monitoring is more appropriate and is accomplished through input from the technical project experts in the APR, discussions at the Steering Committee and the Annual Review. Since outcomes are less tangible than outputs, indicators are indispensable for an informed analysis of progress.

**For impact indicators** (also called situational indicators), discussion may take place annually if information is available but is often done less frequently. Discussions may be scheduled on the occasion of the CCA and Country Programme preparation, mid-term in the Country Programme and/or UNDAF, and towards the end of the Country Programme or SRF period.