UNITED NATIONS

GENERAL ASSEMBLY

Distr.
GENERAL

A/34/271 + add. 1
31 May 1979

ORIGINAL: ENGLISH

Thirty-fourth session
Item 102 of the preliminary list*

JOINT INSPECTION UNIT

Initial guidelines for internal evaluation systems of United Nations organizations

Note by the Secretary-General

1. The Secretary-General has the honour to transmit to the members of the General Assembly the report of the Joint Inspection Unit on the initial guidelines for internal evaluation systems of United Nations organizations (JIU/REP/79/2).

2. As indicated by the Joint Inspection Unit, the report is also of concern to the Economic and Social Council and the Committee for Programme and Co-ordination.

* A/34/50.

JIU/REP/79/2

INITIAL GUIDELINES FOR INTERNAL EVALUATION SYSTEMS OF UNITED NATIONS ORGANIZATIONS

Prepared by

Earl D. Sohm
and
Maurice Bertrand

Joint Inspection Unit

Geneva
February 1979

GE.79-0921

Table of Contents

SUMMARY

These guidelines, based on experience and consultations with the organizations, provide an initial common framework of principles and choices to assist United Nations organizations in developing and improving internal evaluation systems. Because United Nations system activities are so diverse and evaluation practice is still evolving, the guidelines are intended as a flexible framework and a stimulus to thinking, to be revised over time as experience is gained.

Chapter I briefly reviews the background of this guidelines effort, while Chapter II outlines the basic purposes and characteristics of evaluation systems on which there appears to be general agreement.

Chapter III discusses three major aspects of the effective integration of internal evaluation systems with the organizational decision-making process. Evaluation efforts should be well-integrated with organizational planning and programming processes and with current system-wide efforts to improve those processes. They should also be concerned with improving broad organizational strategies for attaining programme objectives, and they can better assist decision-making when conducted under an orderly pattern of evaluation coverage and an approved overall evaluation plan.

Methodological concerns are discussed in Chapter IV. Clearer statements of objectives are needed, good indicators should be developed, and there is a need to consider the various levels of sophistication of evaluation (and associated techniques) which United Nations organizations can actually use. Each of these areas requires much work and a long-term development process, but they will be important factors in improving internal evaluation and overall decision-making in the United Nations system.

Organizational considerations are also important, as discussed in Chapter V. An operating approach or combined approach must be chosen from among self-evaluation, "task-force" or "peer-group" evaluation, a central evaluation unit, or outside evaluation options. In addition, the composition, location, and responsibilities of an evaluation unit must be established, policies and procedures for evaluation reporting and follow-up developed, and evaluation training requirements determined.

The Annexes provide a list of key questions which might be asked during an evaluation, and a partial bibliography of evaluation guidelines which have been developed within the United Nations system.

In Chapter VI these general guidelines are proposed as the initial guidelines for internal evaluation activities in the United Nations system. It is recommended that they be considered by the Administrative Committee on Coordination and the appropriate intergovernmental bodies,and that each United Nations organization report to its executive or governing body on its activities with regard to the internal evaluation system topics discussed in this Report.

I. BACKGROUND

1. In its 1977 report on the history, nature, and status of evaluation efforts in the United Nations system[1/], the Joint Inspection Unit (JIU) described the various organizations' current efforts to develop new or revised internal evaluation systems. It concluded that this large volume of new evaluation activities and initiatives indicated strong new interest in and support for evaluation by executive heads and inter-governmental bodies. At the same time, however, the diversity of approaches reflected not only the particular requirements of each organization but also a lack of common agreement as to the purposes and functions of evaluation. The JIU recommended, among other things, that an attempt be made to develop broad guidelines for internal evaluation systems which could be used flexibly by each United Nations organization.

2. Similarly, the Committee for Programme and Co-ordination (CPC) recommended in 1977 that efforts should be intensified to define a methodology for evaluation based on a maximum degree of common principles and guidelines, as developed through co-operation among United Nations organizations and the JIU[2/]. The Administrative Committee on Co-ordination (ACC), in its comments on the 1977 JIU report, agreed that evaluation guidelines would be most useful, provided that they be flexible and that the United Nations organizations participate in their development[3/].

3. This report is an initial attempt at a broad common framework for internal evaluation in the United Nations system. It is based on the 1977 JIU report, on a synthesis of past and present internal evaluation guidance utilized by various United Nations organizations (see Annex II), and on consultations with the organizations during 1977 and 1978.

4. Widely-differing structures, subject matter, and activities have led United Nations organizations to develop a variety of internal evaluation approaches, structures, and techniques. Consequently, it does not seem possible, at least at present, to formulate a standard approach for the creation and operation of United Nations internal evaluation systems.

1/ "Report on Evaluation in the United Nations System", JIU/REP/77/1, March 1977 (ECOSOC document E/6003).

2/ Committee for Programme and Co-ordination, Report on the work of its seventeenth session, Official Records of the General Assembly, Thirty-Second Session, Supplement No. 38 (A/32/78), paragraph 6.

3/ ECOSOC document, Comments of ACC, E/1978/12 of 31 January 1978, paragraph 18.

5. Nevertheless, it is useful to establish an initial basic framework of agreed principles and alternative choices which each United Nations internal evaluation system should take into account. Such a common framework can be useful to: assist United Nations organizations in creating or revising internal evaluation systems; facilitate further common understandings of United Nations evaluation activities; develop insights into the effectiveness of various evaluation techniques and strategies; and to achieve more compatible evaluation systems and better co-ordination of inter-agency evaluation activities.

6. Moreover, the development of a common frame of reference for internal evaluation within the United Nations system is the natural and timely complement of the system-wide efforts - to which the JIU has contributed over the years - to develop and implement effective planning and programming systems. As will be seen, the inter-relationships between planning, programming and evaluation are one of the main themes that runs through this report.

7. The consultations between the JIU and the various agencies of the UN system have shown that a general agreement seems to exist on the basic purposes and characteristics of evaluation systems. However, each organization also needs to consider a number of major issues or choices in order to establish, revise, or clearly specify its internal evaluation system. Consequently, this report will deal with four main topics:

(a) the purposes and desirable characteristics of United Nations internal evaluation systems, on which there already appears to be substantial agreement (Chapter II);

(b) three broad areas of concern - integration with the programming system, improvement of policies, and coverage and work plans - which, when well developed, can help the internal evaluation system to achieve its proper role within the organizational decision-making process (Chapter III);

(c) major methodological considerations which each United Nations organization should address (Chapter IV); and

(d) Organizational alternatives pertaining to the location and operation of an internal evaluation system (Chapter V).

8. These guidelines are therefore not intended as a rigid set of instructions. Their purpose is to stimulate thinking, rather than to serve as a substitute for it, by providing a broad common guidance framework to be applied flexibly and pragmatically to the many diverse evaluation situations which United Nations organizations face. It is proposed in Chapter VI that these general guidelines serve as the initial guidelines for internal evaluation activities within the United Nations system, that they be considered by the Administrative Committee on Coordination and the appropriate intergovernmental bodies, and that each United Nations organization report to its executive or governing body on its activities concerning the internal evaluation aspects discussed in this Report. Over time and after more experience is gained, the guidelines can be revised and might be made more specific for certain areas of evaluation as well. The JIU plans to review in 1980 the progress being made by the organizations in specifying and implementing their evaluation systems.

II. PURPOSES AND DESIRABLE CHARACTERISTICS OF EVALUATION

A. Purposes of evaluation

9. Evaluation, broadly stated, is an organizational process which uses past experience to improve current and future decision-making and planning. Since the activities of any organization almost never turn out precisely as planned, and in fact proceed from considerable initial uncertainty, evaluation provides an analysis of lessons learned from both successes and failures as feedback information for decision-makers.

10. More specifically, evaluation may be defined as a process which attempts to determine as systematically and objectively as possible the relevance, effectiveness and impact of organizational activities in light of their objectives. It serves as a learning and action-oriented management tool and organizational process for improving both activities still in progress and future planning, programming, and decision-making. Evaluation thus differs from appraisal (the critical assessment of an activity before deciding to undertake it), monitoring (continuous oversight of activity implementation to ensure that it is proceeding as planned), inspection (a special on-the-spot investigation of activity problems), and audit (review and reporting on an activity's conformance with pre-determined standards or criteria).[4/]

11. Evaluation involves a series of simple but basic questions, such as:

(a) Where did we want to go and why?

(b) How did we plan to get there?

(c) What criteria did we use to determine whether we got there?

(d) Where are we actually going/did we actually go?

(e) What factors are associated with or have caused success or failure?

(f) How much change has been/was produced and for whose benefit?

(g) How much have/did our efforts cost ourselves and others?

(h) What decisions about current and future objectives, resources policies, planning and programming should be made in light of the answers to the above questions?

4/ Definitions of key evaluation terms together with concrete examples of their meaning, as presently used within the United Nations system, are included in the JIU Report on "Glossary of Evaluation Terms" (JIU//REP/78/5 of November 1978).

12. The primary purposes of United Nations system internal evaluation processes are to help maximize the effectiveness of activities by:

(a) providing analytic information on results to Secretariats and intergovernmental bodies in order to improve the selection, adjustment, design, programming, and implementation of new or continuing projects, processes and programmes;

(b) providing accountability to intergovernmental bodies for the effective use of resources;;

(c) stimulating overall organizational interest in assessing operational experience and applying the lessons learned to future operations on a continuing basis.

B. Desirable Characteristics of an Evaluation System

13. In accordance with the above purposes and scope, an internal evaluation system should possess certain basic desirable qualities or characteristics. There are many such qualities and their importance may vary from organization to organization, but the following concepts appear to have achieved general acceptance within the United Nations system:

(a) an evaluation system needs strong and continuing direction and support from intergovernmental bodies and executive heads;

(b) the system should be participative and collaborative insofar as possible, to develop self-evaluation processes and constructive understanding of and support for evaluation throughout the organization;

(c) evaluation should be an integral part of the decision-making cycle, tied to the planning and programme budgeting system and to the implementation process;

(d) evaluation should be closely linked with management information systems and other basic information and reporting processes;

(e) the evaluation system should be co-ordinated and systematic in approach and as objective as possible, and it should also be flexible enough to meet diverse situations, respond to changing conditions and new decision-maker needs, and keep evaluations focused on priority areas;

(f) the evaluation system should contribute to constructive change, and should strive to identify programme successes and failures and appropriately apply the lessons learned rather than dwelling on operational short-comings without offering positive solutions;

(g) there should be careful advance planning of the purpose, scope, and resources available for individual evaluations;

(h) the evaluation process should emphasize timely communication of findings, both oral and written, in clear, concise reports which highlight matters of current concern and interest rather than providing stale historical data;

(i) the evaluation system should include an orderly follow-up process, in order to determine how reports were used;

(j) evaluation system policy should encourage cooperative evaluation efforts and information exchange with other United Nations organizations and bodies, Member Governments and their development assistance agencies, and other organizations, both to expand and strengthen evaluation activities in general and to foster the exchange of successful evaluation techniques and approaches.

III. ISSUES INVOLVED IN INTEGRATING THE EVALUATION SYSTEM WITH THE ORGANIZATIONAL DECISION-MAKING PROCESS

14. No internal evaluation system can be effective if it operates in isolation. Instead, it is extremely important that the evaluation system be an integral part of the organizational decision-making process, fulfilling its proper role in the larger planning-programming-budgeting-implementation-evaluation cycle. Three issues deserve particular consideration as part of this process of integration:

(a) the relationship between evaluation and the planning and programming system;

(b) the use of evaluation to improve policies or strategies;

(c) evaluation coverage and evaluation plans.

A. Relationship between Evaluation and the Planning and Programming System

15. As stated in the preceding section, the main objective of evaluation is to improve current and future decision-making and planning through the provision of analytical information on results. Since feedback of lessons learned is so critically important in effectively linking evaluation with decision-making, organizations should devote considerable attention to specific ways and means of ensuring this continuing process (as discussed further in Sections III.C. and V.C. of this Report).

16. At the same time the quality of planning and programming systems is itself an important factor in improving the effectiveness of the evaluation system. Good planning and activity design can greatly facilitate good evaluation, while poor planning and design make a proper evaluation very difficult. Activities should be designed with evaluation in mind - with clear and measurable objectives, specification of appropriate indicators, statements of assumptions made, a clear and logical implementation sequence, and identification of expected impacts and relationships with higher-level objectives.

17. Evaluation should also be firmly linked with programming, budgeting and implementation processes, so that corrective actions indicated by evaluation can be promptly considered and taken. The overall programming system should provide sufficiently explicit objectives, targets and indicators as benchmarks so that it can serve as a "built-in" component of the programming cycle. Programming, budgeting and monitoring systems should also provide sufficient details on inputs, costs, and outputs to allow evaluation of the quality and productivity of the implementation process.

18. The relationship between evaluation systems and planning and programming systems, therefore, is and should be a mutually beneficial one. Although evaluation and planning, programming, budgeting and implementation have not often been well-integrated in United Nations organizations in the past, recent efforts and recommendations to improve the overall quality of such processes within and among organizations may improve their coordination in the future.

19. In 1977, the General Assembly adopted resolution 32/197 on the restructuring of the economic and social sectors of the United Nations system. Among its concerns, the resolution cited the need to enhance the effectiveness of planning, programming, budgeting, and evaluation procedures within the United Nations system. It recommended thematic approaches to programming and budgeting to ensure implementation of the overall priorities established by the General Assembly, further harmonization of medium-term plans and programmes and consideration of their basic concepts, and better harmonization of budget presentations, classification and content. The resolution also called for prior consultation on work programmes and joint programming, improved coordination of activities at the intergovernmental level, and measures to improve the effectiveness of internal evaluation procedures in respect of programme implementation[5/].

20. As recently summarized by the Administrative Committee on Coordination, progress has been made in efforts to harmonize programme budgets and medium-term plans, although serious obstacles still remain. A series of proposals has been made for the further harmonization of programme budget presentations, including a common structure with improved programme narratives and tables and information annexes. In particular, a four-level programme hierarchy (major programme, programme, sub-programme and programme element) has been identified for the subdivision of the organizations' overall programmes.

21. With regard to medium-term planning, most of the larger organizations now have such plans. The common principles established for the plans stress the need for them to be problem- and issue-oriented; to include statements of objectives with targets, implementation approaches, and strategies and related objectives; and to identify cooperative, coordinative, or joint planning objectives. The principles also emphasize that objectives in medium-term plans should be directed toward external impact; contain indicators to verify their progress; establish explicit links between successive levels of objectives and related programme activities; and be reviewed and approved by policy-making organs who also participate in verifying progress in attaining them[6/].

5/ General Assembly resolution A/RES/32/197

6/ Administrative Committee on Coordination, Annual Report for 1977/1978, Part III, E/1978/43/Add.2 of 20 April 1978.

22. It is also clear, however, that much remains to be done to make such principles and reforms fully operational and effective throughout the United Nations system, and to realize the potential for improving internal evaluation systems and the other management system components to which they are intimately linked. For instance, a recent JIU report on Programming and Evaluation in the United Nations concluded that six gaps remain to be filled in the United Nations programming system:

(a) sub-programmes in the medium-term plan lack identifiable and precise objectives and target dates;

(b) outputs in the programme budget are not defined in full or sufficient detail;

(c) supplementary information on the work programme is unreliable and is not monitored, and there are no methods for determining actual costs;

(d) there is no sufficiently detailed planning for programme implementation at division and section levels, and programmes are never fully implemented within prescribed target periods;

(e) a performance report exists, but it monitors programme performance only from the budgetary aspect;

(f) the present evaluation system provides no systematic or regular evaluation of overall results, there is no means of using the evaluation exercises as a guide to the future, and objectives lack built-in indicators.

The JIU report makes specific recommendations for filling these gaps, and urges action to translate the existing programme budgeting procedures and documents into an effective and fully operational programming system 7/.

23. If internal evaluation systems are to be developed and improved, the process must occur not only through direct efforts by intergovernmental bodies and Secretariats to improve evaluation itself, but as part of their continuing efforts to improve the quality of policy and programming processes in the United Nations system. As the above report suggests, many problems must still be solved,and the establishment of fully operational internal evaluation systems will be intimately related to the achievement of more effective management systems overall.

7/ JIU/REP/78/1, March 1978

B. The Improvement of Strategies or Policies through Evaluation

24. One of the key features of a good planning and programming system should be the definition and formulation of a strategy or policy for attaining programme objectives. In practice, however, it appears that too often such strategies are poorly defined or are even absent from the programmes of United Nations organizations.

25. Since the objective of evaluation is to provide analytical information on results,the assessment of programme strategies will be an important component of internal evaluation system activities, whether developed through the direct evaluation of a programme, an evaluation of activities in a functional field, or the evaluation of a set of projects. The quality of a programme strategy should even be of some concern in the evaluation of a single project, since one of the major purposes of project evaluation is to assess the way in which achievement of project objectives contributes to the achievement of higher-level objectives.

26. A well-developed programme strategy should include:

(a) a clear statement of the existing situation in the field of the programme, the problems to be solved, and the extent to which a solution is considered feasible;

(b) a statement of the objective or objectives to be reached (the objective is the converse of the problem), the time-span required, and the intermediate objectives or targets which can be reached in a medium-term period;

(c) identification, if possible, of the various alternative strategies which might be used to achieve the objectives, and the rationale for the strategy chosen, taking into account the means of action at the disposal of the organization (manpower, money, etc.);

(d) a discussion of the resources, processes, and activities which are needed to carry out the programme;

(e) identification of the specific role of the organization and the roles of the governments of Member States and of other organizations concerned.

27. If such a programme strategy is described in a medium-term plan or other document approved by intergovernmental bodies, it provides an important source which evaluators can use to much more effectively assess the successes and problems of programme conduct, understand and evaluate the strategy applied, and attempt to determine the effectiveness, relevance, and impact of the programmes or projects being evaluated. If there is no guiding strategy, the evaluation process is made much more difficult, and the evaluators will have to devote much time and effort to identifying and developing the information which a good strategy statement should have provided.

28. Executive heads and intergovernmental bodies are interested in evaluation only if its findings provide useful information for improved planning, programming, and strategy or policy decisions. It is therefore important that evaluations, particularly those dealing with broader programme issues, provide an objective questioning and assessment of programme strategies and of the translation of these strategies into projects, processes and activities. However, the serious risk always exists that evaluation might become merely an instrument for justifying the existing state of affairs. To overcome any tendency toward self-serving evaluation, reports to executive heads and intergovernmental bodies should carefully consider and explain the methodological and organizational techniques and policies used.

C. Evaluation Coverage and Evaluation Plans

29. To best meet decision-making and programming needs, internal evaluation system efforts should also be guided by a well-thought out plan of organizational coverage which effectively applies limited evaluation resources. This requires consideration of the types of activities to be evaluated, and the determination of which, how many and how often activities should be evaluated in an organization in a given period of time. One of the following basic patterns or a combination of them might be adopted.

(a) "built-in" evaluation, in which evaluation procedures are made a standard element of the design and implementation of organizational activities;

(b) periodic evaluation, in which evaluation is conducted on a regularly-scheduled basis for selected activities (perhaps at the mid-point of implementation, before a decision is made to begin a new phase, upon completion, or in a "rolling" pattern of coverage in which a certain number of activities are selected for evaluation each year);

(c) ad hoc evaluations, more flexibly scheduled to permit evaluations of activities as the need arises.

30. Built-in evaluation can provide the broadest and most orderly coverage of organizational activities, but it can require considerable resources when conducted organization-wide and may provide evaluations in less depth. To be fully effective it requires the use of achievement indicators, and some time will be necessary for it to produce positive results. Built-in evaluation can be supplemented or replaced by periodically scheduled evaluations, which have the potential for somewhat more detailed analysis, or by ad hoc evaluation coverage, which potentially can provide the most careful and detailed studies. The critical determinants are the degree and depth of evaluation coverage that is deemed necessary for effective decision-making and the overall organizational resources available.

31. Because resources for evaluation are inevitably limited, organizations should atempt to identify and concentrate on priority areas, with particular awareness of changing situations or new decision-makers' interests, so that evaluations address today's and tomorrow's problems rather than yesterday's. The choice of the activities to be evaluated should also be governed by:

(a) a "cut-off" concept which recognizes the fact that because of their small size or other factors, certain types of activities need not be evaluated, or should be evaluated only very infrequently;

(b) some regular pattern or cycle whereby all areas of an organization receive at least some evaluation coverage over a period of several years;

(c) sufficient flexibility to respond to unanticipated, ad hoc, priority evaluation needs of decision-makers;

(d) sufficient flexibility to permit co-operative evaluation exercises with other evaluation or review groups inside or outside the organization.

32. In determining how best to apply scarce evaluation resources to specific activities, certain criteria could be developed to help select evaluations which might be of most assistance to decision-makers. The following considerations might provide some of the criteria for choice:

(a) activities entering a second phase or being considered for substantial expansion or reorganization;

(b) activities for which evaluation could provide guidance for other similar activities, e.g., institution-building projects;

(c) activities which may be of marginal interest in view of changing priorities;

(d) activities which may be encountering difficulties;

(e) activities whose cost-effectiveness is uncertain.

33. The criteria for choice, however, might equally well be based on a concept of orderly and expanding evaluation coverage. Certain categories of activities (defined in terms of subject matter content, financing arrangements, or size) could be selected for evaluation coverage. The number of such categories could gradually be increased as the evaluation capacity of the organization increases, with the ultimate goal of subjecting all types of activities to evaluation.

34. A final consideration is the appropriate mix of evaluation coverage. United Nations system evaluation efforts in the past have concentrated heavily on the evaluation of projects, with much less activity in the area of programme evaluation, and only limited attention given to the evaluation of administrative and support processes. Each of the three types, however, represents an important component of overall operations of the organization, and accordingly each should receive some form of evaluation coverage. Programme and administrative and support process evaluations may well increase in the future as the organizations further develop and refine their programme budgeting systems.

35. Each internal evaluation system should be guided by an "evaluation plan". Such plans would be prepared by the Secretariats and might set out alternative possibilities concerning the types and magnitude of activities to be evaluated in a given period and the degree of flexibility to be maintained. Secretariats, executive heads and intergovernmental bodies could then make a final choice of the appropriate evaluation pattern and coverage.

36. In time, as medium-term planning and programme budgeting procedures are further improved, it should also be possible to integrate these evaluation plans into the regular planning process.

IV. METHODOLOGICAL CONSIDERATIONS FOR INTERNAL EVALUATION SYSTEMS

37. Methodology is the particular set of analytical methods and techniques used to perform the evaluation of an activity. The methods which can be used for evaluation are drawn from a number of fields and have developed a variety of forms. They may involve different data-collection instruments (statistics, interviews, questionnaires, on-site observations, etc.). They may involve various analytical approaches (from empirical case studies, the most common type, to highly sophisticated experimental studies using control groups). They may use relatively simple techniques, such as interviews or analysis of reports, or very complex ones, such as econometric analysis, simulation, or cost-benefit analysis.

38. The possibility of using the more sophisticated instruments, approaches, and techniques hardly exists at present in the United Nations system, for obvious reasons such as the limited staff, the limitations of the current planning, programming, budgeting, management information and accounting systems, and the difficulty of applying elaborate analytical methods to programs as complex and diverse as those of the United Nations organizations. In the selection and use of more modest evaluation methodologies for United Nations internal evaluation systems, the following considerations seem particularly important: the use of objectives and indicators in evaluation; the concept of differing levels of evaluation; and the problems and challenges of improving United Nations systems evaluation methodologies in the future.

A. The Use of Objectives in Evaluation

39. Objectives, i.e. the purposes and aims of an activity, representing the desired state which the activity is expected to achieve, are highly important in evaluation efforts because they establish the intended purposes against which the relevance, effectiveness, and impact of the activity will be assessed. Well-structured statements of objectives, including the necessary back-up documentation, should meet a number of criteria. Ideally, they should:

(a) clearly state the specific situation in which the objective is to be obtained (baseline condition);

(b) clearly identify the priority needs to be met by the project or programme;

(c) provide an understanding of the intended beneficiaries and benefits, the expected level of attainment, and the time period or periods in which specific degrees of attainment are expected;

(d) be specific enough, with respect to the nature and direction of intended change, so that progress can be measured whenever possible by carefully chosen and objectively verifiable quantitative indicators which are included in the objectives documentation;

(e) take account of multiple objectives, which might be either complementary or conflicting (while attempting to keep statements of objectives as clear and simple as possible), and indicate the relationship of the specific objectives to the higher-level objectives or policies to which the activity is expected to contribute;

(f) identify, to the extent possible, the key assumptions (external factors which influence activity success but are largely uncontrollable), qualitative aspects, and indirect costs and benefits which may be involved in the activity.

40. In actual practice, of course, it is not easy to achieve such clarity of objectives. Within the United Nations system, progress has occured over the years in formulating definitions of projects, although further improvements can be made. For administrative and support processes, objectives, productivity standards, proposed work levels, and unit-cost reductions have not often been well-developed. They tend to be rather vague and general, with completion dates and accomplishment goals which are largely indeterminate. This obscures accountability and makes it difficult for executive heads and governing bodies to assess budgetary requirements. The most difficult problems, however, have arisen in efforts to develop clear and precise objectives for programmes.

41. A variety of factors make it difficult to develop good statements of objectives:

(a) the broader the scope and size of the project or programme, the more likely it is that the objectives will tend to be multiple, vague in language, difficult to measure clearly,ambiguous as to their inter-relationship and priority, and over-ambitious in their purpose relative to the results that will actually be achieved.

(b) many United Nations activities, especially in the area of technical cooperation, are experimental and exploratory in nature;

(c) it may be difficult to state objectives in such a way that precise and reliable indicators of progress may be used;

(d) measurable objectives might prove to be so narrow that they are actually misleading as to the activity's true purpose;

(e) objectives might be vaguely stated as a political technique to build support for an activity or to weaken accountability by keeping its purposes unclear;

(f) vague objectives may reflect disagreements among decision-makers about the intended costs, benefits, and results of an activity;

(g) unclear objectives may simply mean that objectives were not well-thought out at the design stage;

(h) it may be difficult to formulate objectives with precision because they often change over time;

(i) while project objectives are more definite, relatively short-term, and non-recurring, programme and administrative process objectives are more often continuing and long-term in nature.

42. Nevertheless, objectives should be stated as clearly, as simply but systematically, and in as objectively verifiable terms as is possible, in order to facilitate management, accountability, and evaluation. The difficulty posed for evaluators by poorly-stated objectives is not simply that they make it difficult to apply a proper methodology. It is also that they may lead the evaluators to address the wrong problems; may produce information that is too limited, too imprecise, and not relevant to the information needs of decision-makers; or may subject the evaluators to the charge that the resulting evaluation was itself too abstract or imprecise.

43. To avoid these problems four positive steps can be taken. First, activities should be carefully considered during the design stage to assess the clarity and logical relationship of the objectives, targets, indicators, and implementation sequence involved. Second, evaluators should closely and critically examine stated objectives in the evaluation design stage, in order to determine the "evaluability" of the programme or project and anticipate serious problems of definition or measurement, and to attempt to either qualify evaluation work plans accordingly, or discuss evaluation problems with appropriate decision-makers, or both. Third, evaluators should discuss the stated objectives with decision-makers and planners at an early stage in the evaluation process in order to determine whether there is a common understanding of the actual project or programme objectives.

44. Finally, and perhaps most importantly, intergovernmental bodies and Secretariats should continually encourage and support further development of co-ordinated management processes. Such processes should emphasize the integration of planning, programming, implementation, evaluation and reporting, and the incorporation of clear and specific objective statements in all project, process and programme planning. Through such actions, specific and precise statements of objectives can be developed not only to facilitate specific internal evaluation efforts, but to make an important contribution to improved planning, programming,implementation, and accountability as well.

B. The Use of Indicators in Evaluation

45. Another important consideration in developing evaluation methodology is that of indicators - the objective and specific measures of changes or results expected from an activity. Indicators have received only limited attention in the United Nations system and their application is difficult, but there is increasing interest in indicators as an important element of sound management of activities. Indicators can also form a critical evaluation element, which relates the results achieved to the original objectives set out for a project, process or programme.

46. Ideally, the programming system of an organization should provide sufficiently explicit objectives and targets which make use of quantitative and qualitative progress indicators as benchmarks for evaluation. During the design stage planners should facilitate subsequent evaluation by establishing, as precisely as possible, the baseline data (i.e., starting conditions) which the project or programme is intended to change; the project objectives, in finite verifiable terms and targeted toward specific magnitudes of change at specific times; and the planning assumptions or external factors which lie outside the scope of the design but are important to its success. In this design process of definition and logical linkage of baseline conditions, inputs, outputs, targets, and objectives, progress indicators would serve to measure progress from the baseline conditions to the planned objectives. It is important that the indicators be established at the design stage so that change can be systematically observed and necessary data routinely collected.

47. If carefully formulated and applied, indicators can help fulfil the basic objectives of evaluation by:

(a) clarifying the fact, nature and rate of change that has occurred or is expected to occur;

(b) permitting comparison of the change with that planned, and its impact on higher-level objectives as well;

(c) assisting in the examination of input-output and cost-benefit relationships.

48. Various types of indicators might be used to measure such things as volume, output, progress, performance, user satisfaction and impact. Where possible, indicators should be quantitative. When it is not possible to measure change directly, indirect or "proxy" indicators may be found, although care must be taken to determine that such measures are relevant to the activity being evaluated. But however they are selected, the indicators should be plausible, independent, objectively verifiable, and capable of being translated into targets set at a specified time when the results or changes should be visible.

49. Despite their merits, indicators can cause many problems if they are poorly constructed or wrongly applied. Initially, there is a very real challenge inherent in establishing good indicators to measure the complex types of changes which United Nations activities so often address. Indicators might lead to overemphasis on a quantitative measurement process rather than to discussion of the equally important questions of how and why such change did or did not occur. Planners might select too many indicators, which would require excessive data collection efforts and costs for both managers and evaluators. Finally, indicator usage might lead to concentration on short-term, easily-achievable results, ignoring the possibility that truly significant achievements may take a longer time to emerge.

50. Nevertheless, since the most essential characteristic of evaluation work is the effort to measure results against objectives, it is clear that indicators are a vital part of the evaluation process. As with objectives, good indicators rely strongly on good planning and design of a project or programme. Objectives and measurements of results are interdependent: the quality of each depends on the other. Appropriate progress indicators not only allow the examination of progress toward objectives If progress is not being achieved, they can help lead back to the alternative causal factors, such as project or programme design, strategy, inputs, or performance.

51. In this regard, it may also be possible to define better achievement indicators for United Nations system activities if objectives and indicators are linked with the idea of "clientele"--the users or "audience" of a project or programme. The objectives of an activity envisage certain clienteles or target groups to whom the activity is directed. If these clienteles, both actual and potential, can be identified clearly, it should be possible to build indicators into the objectives to measure the degree to which the activity is meeting the needs of its users. The objectives-clientele-indicator relationship would require considerable research to develop a truly satisfactory method, and might focus more on outputs than on impacts. But the use of indicators integrated with objectives and directed toward measurement of fulfilment of clientele needs might prove to be a fundamental component of a good evaluation of a project, process or programme[8/].

8/ The nature of this relationship is more fully discussed in Chapter VI of the JIU Report on Programming and Evaluation in the United Nations (JIU/REP/78/1 of March 1978).

52. Thus, while indicators have received only limited attention in the United Nations system in the past, the challenging task of developing good indicators provides the opportunity to greatly improve future internal evaluation efforts and overall project, process, and programme design and performance.

C. Levels of Evaluation

53. Although evaluation is defined as the attempt to determine as systematically and objectively as possible the relevance, effectiveness, and impact of activities in light of their objectives, it is clear that even within this definition some types of evaluations are more difficult to perform than others, and that the various methods and techniques which might be used vary considerably in their degree of sophistication. One of the major problems in developing common concepts of "evaluation" has been that some people view it pragmatically as the use of a "best-possible" set of modest techniques, while others insist that it must be a highly precise and rigorous exercise using sophisticated techniques.

54. There is merit in the arguments of the evaluation "realists", who maintain that evaluation should not promise more than it can deliver, and in the arguments of the evaluation "scientists", who urge that evaluation must be highly objective and systematic to differentiate it from more casual and less reliable types of review. The improvement of evaluation activities in the United Nations system will probably require a balancing of both views (as noted in section D. following). At the same time, however, United Nations system evaluation activities could benefit from efforts to more carefully consider and determine the various levels of sophistication of evaluation and evaluation techniques which are currently being used or could be used.

55. A variety of factors affect the ambitions of a particular evaluation exercise and the methods and techniques chosen to carry it out. In addition to the very basic considerations of the amount of resources and time available, these factors would include:

(a) the type of activity evaluated (project evaluation methodologies have already been fairly well developed in the United Nations system, but programme evaluations may be even more demanding because of their breadth, complexity and diversity, and administrative process evaluation methodologies have generally not been fully utilized;

(b) the type of evaluators involved (full-time evaluators and experts and consultants may be able to utilize sophisticated methodologies, while staff members participating only occasionally in a self-evaluation may have to rely on simpler and more standardized methods);

(c) the various users of an evaluation (the nature of the methods and techniques chosen might vary considerably depending on whether the major users of the evaluation request an emphasis on the activity's results and impact, implementation, planning and design, or special problems, or whether the evaluation is being done for programme managers, technical specialists, executive heads, or intergovernmental bodies);

(d) the various timings of the evaluation (the mix of methods chosen may vary according to whether the evaluation will be one of a periodic series, performed during implementation while results are still emerging, or performed after the activity has been completed).

56. The quality of the planning and programming system, however, may be more important than any of the other factors listed above in facilitating or restricting the methods chosen. When the activity evaluated is not well-designed and objectives, indicators, and other basic data are lacking or are of poor quality, the evaluation methods will probably be limited to analysis of documents and reports, physical inspections, interviews, questionnaires, analysis of compliance, and expert opinions. But when an activity is well designed with well-formulated objectives and indicators and other basic data available, evaluation methods can include comparative analysis, measurement of results, cost and productivity analysis, opinion surveys, statistical sampling and inference techniques, and other indices, measures, and criteria.

57. In view of these various factors influencing the nature and sophistication of various evaluations undertaken, and in light of the developing nature of evaluation activities in the United Nations system, it is premature to establish any orderly set of levels of sophistication of evaluation and to determine the techniques associated with them. However, an initial and very tentative list of levels, moving progressively and cumulatively from the simpler to the more complex, might include the following:

(a) determining the nature and extent of inputs provided and outputs produced in light of the objectives (for example, processing and sales of a publication or instruction and number of students trained);

(b) obtaining the views of intended users of the activity (as to their satisfaction with the quality of the publication produced or the students trained);

(c) assessing the direct impact of the activity (the ways in which the publication or training activity has actually produced changes in the problem situation which it sought to address);

(d) assessing the effectiveness and efficiency of the activity (the extent to which the publication or training activities have achieved their objectives and the productivity of their implementation processes);

(e) critically examining the overall design and relevance of the activity (reassessing the rationale, objectives and strategy of the publication or training activities and their importance relative to long-range objectives or other priority needs and concerns);

(f) assessing the broader impact of the activity (the ways in which the publication or training activity has produced changes which contribute to achievement of the broader, long-range objectives of which it is a component part).

58. The concept of various levels of sophistication of evaluation methodologies is one which those who are involved with United Nations system evaluation activities are well aware of, but also one which has received only limited attention. In developing and revising their internal evaluation systems, United Nations organizations could benefit their own evaluation efforts and those of the United Nations system as a whole by carefully analyzing the various levels and types of evaluation and methodologies which they use or might use. Such efforts could serve as indicators of progress and possibilities in the further development of evaluation activities, and to facilitate clearer system-wide understandings of what methods, techniques, and levels of evaluation may prove most useful within the United Nations system.

D. Further Development of Evaluation Methodologies

59. In light of the considerations discussed in the preceeding three subsections--the complexity involved in improving objectives, developing appropriate indicators, and evolving a concept of levels of evaluation--it is clear that much work remains to be done before evaluation methodologies will be well-developed in the United Nations system. A number of organizations have already undertaken extensive efforts to develop or revise their internal evaluation systems and to establish useful and effective methodologies for evaluation. As the JIU noted in its 1977 report on evaluation in the United Nations system[9/], however, further development of evaluation systems and methodologies will be a gradual and long-term process.

9/ JIU/REP/77/1 of March 1977

60. Ideally, evaluation methods provide a set of tools with which to carry out objective, systematic, and reasoned evaluations of many different types of activities. In practice, however, a multitude of operational problems and constraints prevent full use of more sophisticated methodologies in the United Nations system. In these circumstances, some key considerations in the choice and use of methodologies are:

(a) the particular mix of methods chosen should be the most appropriate for the particular activity under evaluation;

(b) the methods chosen should be as objective, rigorous, and systematic as is pragmatically possible;

(c) analysis and measurement should be quantified where possible, but good quantitative data might not be available;

(d) many activities simply may not be evaluable at present in terms of any degree of rigorous scientific methodology;

(e) because of the many operational constraints, the evaluation methodology chosen is almost inevitably much less precise than the evaluator would like;

(f) whatever the method chosen for a particular evaluation, it should be clearly noted and briefly discussed in any evaluation reporting, so that those receiving evaluation reports can understand the rationale for the methodology used.

61. Despite the difficulties noted in these considerations, all evaluation methods and techniques used should be viewed in light of the following basic questions.

(a) reliability - how dependable and consistent is the analytical information being gathered?

(b) validity - how much confidence is there that the findings measure what they purport to measure?

(c) relevance - will the findings in fact be used to answer decision-makers' questions?

(d) significance - will the findings go beyond what is apparent from direct observation and tell the decision-makers something new and important?

(e) efficiency - does the value of the insights gained exceed the cost of using the approach?

(f) timeliness - will the analytical information be available in time to meet decision-maker's schedules?

62. The complexity involved in applying evaluation methodologies should be viewed as a challenge rather than as an unavoidable situation. Over time, continuing efforts to identify, improve, and steadily develop more useful evaluation methodologies can and should make a substantial contribution to improved United Nations decision-making, planning, programming, and performance. Particularly at present, those who design and develop evaluation methodologies need also to be imaginative and creative. To be effective, evaluation methodologies must be carefully considered, appropriately chosen, and intelligently applied, using the most essential evaluation resource of all - sound human judgement.

V. ORGANIZATIONAL CONSIDERATIONS FOR INTERNAL EVALUATION SYSTEMS

63. The major organizational considerations for internal evaluation systems include the choice among various operating approaches to evaluation, decisions as to the location of the internal evaluation function and its responsibilities, policies and procedures for evaluation reporting and follow-up, and training in evaluation.

A. Operating Approaches to Internal Evaluation

64. Perhaps the most important operational question for the conduct of an internal evaluation system has to do with choosing an operating approach or combined approach from among the four following major options.

65. Self-evaluation is evaluation made by those conducting the activity. Self-evaluation can be useful because:

(a) it is a participative process which can spread evaluation coverage widely throughout the organization;

(b) by involving staff directly in setting objectives and evaluating progress made, it can increase their commitment to an activity and their sense of responsibility for it;

(c) it provides evaluations performed by those most familiar with the activity,and allows immediate and direct feedback of findings into redesign and improved execution.

The drawbacks of self-evaluation are that it requires simplified evaluation methods and raises the question of whether staff can objectively evaluate their own work.

66. "Task force" or "peer-group" evaluation is evaluation made by a team of staff members chosen from various parts of the organization to perform evaluations on an ad hoc or part-time basis. Such evaluations can be useful because:

(a) they also provide a participative process which spreads evaluation coverage and experience throughout the organization at managerial levels;

(b) they allow staff from various parts of the organization not only to improve their analytical and problem-solving skills but also to become familiar with operations and problems of other parts of the organization;

(c) they increase the potential for "cross-fertilization" of successful ideas, techniques and operational strategies among various parts of the organization.

The drawbacks of task-force or peer-group evaluations are that such dual work responsibilities may disrupt both the regular work of those chosen for the evaluation team and their evaluation work as well, and that they may participate so sporadically that they have difficulty learning and retaining evaluation skills.

67. Evaluation by a central evaluation unit. Central evaluation can be useful because:

(a) it allows evaluations to be made by specialists, who have expertise and continually-expanding experience in evaluation work and are not directly involved in the activities being evaluated;

(b) it allows a more cohesive and controlled programme of evaluation coverage;

(c) because of their specialized skills and higher-level perspective, central evaluators can make more sophisticated evaluations and evaluate more complex and broad-scope activities.

The drawbacks of central evaluation are that central evaluators might be viewed as distant, control-oriented "policemen" by other staff members, and that a central evaluation unit may be too small to effectively cover all areas of an organization's activities.

68. Outside evaluation is performed by consultants or experts hired on an ad hoc basis from outside the organization, normally under the organization's supervision and often as a part of one of the preceding approaches. Outside evaluation can be useful because:

(a) outside evaluators can bring a fresh perspective to their evaluation work and are relatively "independent" with regard to personal involvement in the activity being evaluated;

(b) outsiders can bring to an evaluation highly specialized skills or knowledge which may not be available within the organization;

(c) outsiders, by virtue of their temporary, ad hoc employment, can add considerable flexibility to evaluation activities to cover peak work periods or perform specialized work.

The drawbacks of using outside evaluators are that their lack of familiarity with the organization's operation, and perhaps with technical co-operation and the operations of the United Nations system as a whole, may slow their work, and that their detached status may weaken the credibility of their findings and suggestions among organization staff.

69. Most internal evaluation systems will probably have some combination of each of these approaches. The particular mix which is chosen will reflect basic decisions of the organization on centralization or de-centralization, the scope of evaluation coverage, the time and resources to be devoted to evaluation, and the extent to which such work will be a participative process.

70. Determination of an operating strategy for an organization's internal evaluation system must also involve decisions about the timing of evaluation activities. The basic patterns of "built-in" evaluation, periodic evaluation, and ad hoc evaluation have already been discussed in Chapter III.C. concerning evaluation coverage. In addition, however, the organization must make a choice or choose some mix of:

(a) "On-going" evaluation of activities during their implementation; or

(b) "Ex-post" evaluation after completion of an activity.

71. "On-going" evaluation allows timely information feedback for corrective actions, but it can also mean that outputs and impacts have not had a real chance to become visible. "Ex-post" evaluations allow more careful assessment of results achieved, but may provide the "lessons learned" too late to be of much use in current planning and programming. In addition, because of the uncertainty of whether intended results will actually occur (and how to measure them accurately), and because the timespan needed for results to occur varies tremendously among various projects and programmes, determining the appropriate time to conduct either "on-going" or "ex-post" evaluations is itself a difficult problem. The mix of these two types of evaluation and the patterns in which they are scheduled is another important choice to be made in designing an internal evaluation system.

B. Location and Nature of Responsibilities for Internal Evaluation

72. Each of the possible operating approaches to internal evaluation discussed in the preceeding section (self-evaluation, task-force or peer-group evaluation, evaluation by a central unit, or outside evaluation) also implies the existence of a focal-point "Evaluation Unit" responsible for guiding or coordinating internal evaluation activities. In United Nations organizations such units are normally small, but they vary in composition. The unit might be comprised of a single individual or a group of evaluators. The organization might also choose to have the unit members participate on a full-time, part-time, or ad hoc basis, or to supplement these staffing patterns with additional evaluation policy or advisory committees. The particular form, size, activities, and funding arrangements of an organization's evaluation unit, of course, would also be governed by general financial and staffing con siderations, the overall organizational structure, and the extent of the evaluation activities undertaken.

73. Whatever the size and composition of the evaluation "unit", however, a decision must be made on its place within the organization. It might be:

(a) attached to top decision-makers (for example, within the Cabinet of the Secretary-General or Director General);

(b) integrated with planning and programming activities (for example, as part of a Division of Planning and Budgeting);

(c) combined with administrative and financial services (for example, as part of a Controller's office or Management Services branch);

(d) and, for each of the three above options, either centralized as a distinct unit, decentralized with full- or part-time members located in other parts of the organization, or some combination of the two (for example, a full-time evaluation coordinator in a Division of Planning and Budgeting, with part-time evaluation coordinators also located in each of the major operating divisions or regional bureaus of the organization).

74. The specific responsibilities of those who guide and coordinate the internal evaluation system will of course vary depending on the location and composition choices made. However, the responsibilities of the designated person or committee or group should be clearly stated, with commensurate authority, and should involve some combination of:

(a) Coordination and leadership of evaluation activities;

(b) Participation with decision-makers in identifying and developing evaluation priorities and work programmes;

(c) Establishment and development of evaluation standards and methodologies;

(d) Monitoring of evaluation scheduling, implementation, and follow-up, including periodic reports on the results of the organization's evaluations;

(e) Liaison with other evaluation and review groups, inside and outside the organization;

(f) Provision of evaluation expertise and assistance when required or requested;

(g) Conduct of evaluation training and research;

(h) Performance of ad hoc evaluations as requested.

C. Evaluation Reporting and Follow-Up

75. The potential users of evaluations fall into several groups:

(a) the various intergovernmental bodies of the organization, the inter-agency bodies of the United Nations system, the organization's member governments and their representatives, and host governments;

(b) the Secretariat of the organization--the executive head, managers, planners and programmers, and field and other staff;

(c) to a lesser degree, those outside the organization--other organizations, the clientele or beneficiaries of the project or programme, and the academic and research communities.

76. Evaluators should attempt to determine which of these various groups or sub-groups are most directly interested in a particular evaluation or type of evaluation. If possible, the evaluators should also consider what particular purposes, time schedules, problem areas, or other facets of the evaluation the specific clientele is concerned with, and what evaluation strategy, format, and reporting process might be best suited to meet the information needs of these key groups. Through such discussions, the evaluators can work with managers, executive heads, and intergovernmental bodies to develop appropriate reporting formats which will facilitate the understanding and use of the evaluation reports, with particular attention to clear presentation of conclusions and recommendations.

77. Several other reporting considerations are also important to provide an effective internal evaluation process:

(a) Reporting policies should consider the need for consultation and discussion of draft reports with those involved, timely and appropriate distribution processes to get reports to intended users, and the need for careful scheduling when an evaluation report is to be considered by a series of intergovernmental and expert meetings and conferences;

(b) Evaluation reporting should be coordinated with other review, reporting, and information processes, so that overlap and duplication are held to a minimum and evaluation reports can provide timely and useful information not available from other channels;

(c) Formal written reports are not the only way to convey "lessons learned" - evaluators can also provide oral briefings, discussions, or interim reports to promote timely corrective action;

(d) Orderly processes also need to be established to ensure effective feedback and use of evaluation findings, provide systematic storage, comparison and dissemination of evaluation experience, and serve as a record for future reference and review.

78. Finally, evaluation reporting is never complete without appropriate follow-up. An orderly procedure should provide for periodic review of the implementation of corrective actions, which have resulted from evaluation conclusions and recommendations approved by managers, executive heads, or intergovernmental bodies. If possible, such periodic reports should be linked with performance reports, in-depth programme studies, or other status reports which the organization produces on a periodic basis.

D. Training in Evaluation

79. Since evaluation is still a relatively new and developing field, training is an important element in the development of a viable internal evaluation system. Training can be used both to train evaluators in evaluation techniques and to expand management and staff understanding of the potential of evaluation. Consequently, within its existing training programmes and policies, each organization should consider the possibilities of training in evaluation at two broad levels:

(a) training in the purposes, techniques, and strategies of evaluation for those who will be participating directly in evaluation;

(b) more general training for organizational managers and planning and programming staff on the purposes of evaluation, the relationship of evaluation to the organization's decision-making and programming processes, the various types, techniques,and methods of evaluation, and the ways in which evaluation and evaluation reports can be most useful to them.

VI. CONCLUSIONS AND RECOMMENDATIONS

80. The present and potential contributions of evaluation to improving the planning and execution of programmes and a better utilization of resources are now well recognized. However, because of the complexity of the activities of the United Nations system, because resources for evaluation are not yet adequate, because staff trained in evaluation are few in number, and because there is still a lack of full understanding of evaluation and therefore some resistance to it, results will continue to be modest for some time. Only after guidelines and principles are established, more systematic methodologies introduced, necessary minimum resources made available, and the internal evaluation systems of the organizations more fully developed can evaluation make its full contribution.

81. The initial guidelines presented in this report are intended as a flexible guiding framework for United Nations internal evaluation systems, and as a series of important choices and considerations which must be addressed in designing and operating such systems. Although an organization's own policies, situation and characteristics are the primary determinants of the particular internal evaluation system that it adopts, these guidelines have identified certain basic elements that every such system should provide.

82. At present, a number of United Nations agencies are making major efforts to revise their existing internal evaluation systems or to develop and implement new systems with evaluation policies which reflect many of the matters discussed in this report. Other agencies already have internal evaluation systems which address many of the basic evaluation system elements. Some of the smaller United Nations agencies have done little in the way of developing internal evaluation sytems, but may have less need for formal systems.

83. United Nations evaluation efforts on a system-wide basis could make substantial further progress toward their basic purpose of improving the quality of decision-making if each United Nations organization were to clearly specify the basic conditions and strategies of its own internal evaluation system. Without compromising the need for evaluation systems adapted to unique organizational situations and without foreclosing future improvement and innovation, such a policy-specification effort could:

(a) strengthen evaluation efforts within each agency and throughout the United Nations system by establishing a common framework of basic evaluation considerations;

(b) thereby facilitate cooperation among agencies on evaluation matters;

(a) strengthen evaluation efforts within each agency and throughout the United Nations system by establishing a common framework of basic evaluation considerations;

(b) thereby facilitate cooperation among agencies on evaluation matters;

(c) also facilitate comparisons and the development and transfer of successful evaluation methods and approaches; and

(d) make the nature, progress and utility of evaluation activities more visible to top decision-makers and governing bodies.

84. The Joint Inspection Unit proposes these general guidelines as the initial guidelines for internal evaluation activities in the United Nations system, and recommends that they be considered by the Administrative Committee on Coordination and the appropriate intergovernmental bodies. In this respect, the Joint Inspection Unit will also initiate periodic reviews of these guidelines with the expectation that changes will be made in light of subsequent experience.

85. It is also recommended that each United Nations organization report at an early date to its executive or governing body, as appropriate, on its policy, structure, and general operating procedures for each of the internal evaluation system aspects identified in the table on the following page.

86. The Joint Inspection Unit plans to make a report in 1980 on the status of evaluation activities in the United Nations system. The progress being made by the organizations in specifying, developing and implementing their evaluation systems would be one of the major subjects of this report.

	Organizational Actions Taken Concerning:	Relevant Chapter of this Report
1.	a statement of the basic purposes, principles, and nature of the organization's internal evaluation system, an analysis of the current overall status of the system, any plan and schedule for its further development (or, if there is no such system, why it is not feasible to establish one)	Chapter II paragraphs 9-13
2.	the integration of evaluation with the planning and programming system, and with efforts to improve that system	Chapter III paragraphs 15-23
3.	the use of evaluation efforts to improve organizational strategies or policies	Chapter III paragraphs 24-28
4.	the development of evaluation coverage patterns and evaluation plans	Chapter III paragraphs 29-36
5.	the improvement of statements of objectives	Chapter IV paragraphs 39-44
6.	the development of indicators	Chapter IV paragraphs 45-52
7.	analysis of the actual or possible levels of sophistication of organizational evaluation efforts and associated methodologies	Chapter IV paragraphs 53-58
8.	plans for further development of evaluation methodologies	Chapter IV paragraphs 59-62
9.	the operating approach or mix of approaches which has been chosen for internal evaluation	Chapter V paragraphs 64-71
10.	the location and nature of internal evaluation responsibilities	Chapter V paragraphs 72-74
11.	policies, processes, format, and procedures for evaluation reporting and followup	Chapter V paragraphs 75-78
12.	training in evaluation	Chapter V paragraph 79

ANNEX I

STEPS AND POSSIBLE QUESTIONS IN THE EVALUATION PROCESS

This Annex attempts to collect a representative list of the various questions which United Nations and other organizations have noted as particularly important for performing a good evaluation. The list forms a progression through the five general steps or stages of an evaluation process:

- evaluation design and terms-of-reference;
- re-appraisal of basic conception and design of the activity;
- analysis of implementation;
- evaluation of results; and
- reporting and follow-up.

While they might serve as an informal checklist, these questions are intended only as a reference list to stimulate thinking about those concerns which might be important to evaluators in designing and implementing an evaluation of a project, process, or programme of their organization. Every question would not apply to every evaluation; some would be much more important than others, depending on circumstances; and the list, although long, is by no means exhaustive.

> NOTE: In the following questions, as elsewhere in this report, "activity" is used throughout as a generic term for organizational undertakings, rather than constantly referring to the more precise but repetitive terms "project, process, or programme". Some organizations normally use "activity" to refer to a specific task or job carried out within a project, as part of the process of transforming project inputs into outputs.

A. Evaluation design and terms-of-reference

1. What is the subject of the evaluation? The purpose? The objective? Who are the potential users?

2. Does the evaluation have the potential for providing new information? New techniques? New procedures? New policies? Have a wide variety of potential findings been considered to help determine whether this evaluation effort will be worthwhile?

3. Who,if anyone, requested the study? Are there particular aspects which the requestor is concerned with? Special guidance of any type about how, where and when the evaluation might be conducted? Is there a deadline for its completion? Will interim reports be needed? Can these matters be discussed with the requestors?

4. What are the legislative mandates and guidance for this activity and how and by whom have they been interpreted? Have any specific deadlines, strategies, outputs, users or targets been prescribed in this guidance?

5. What is the scope of the study? In-depth? Concerned with progress? Efficiency? Effectiveness? Impact? Have constraints or restrictions been placed on the scope of the study? If so, what is their nature and expected impact?

6. What are the major questions to be answered about the structure, process and consequences of the activity during the study? Have they been critically assessed by people involved in the planning and execution of the activity, and have these assessments been used to reformulate the evaluation approach where necessary?

7. Is this evaluation a high priority relative to other evaluation alternatives? Why?

8. Are there reporting and review documents or procedures for the activity which provide "built-in evaluation", monitoring data, or other review information which can serve as a sound basis for the evaluation?

9. Have other relevant evaluations or reviews been conducted on this subject? What did they find?

10. What are the information requirements for the study? Are they reasonable and achievable?

11. Does it appear that certain information will not be available because it is too hard to assemble? Does not exist in the desired format? Is not "official" data or is only a preliminary or working paper? Is confidential or privileged? Can action be taken to overcome these obstacles to data access, or can secondary data sources be found?

12. What are the data collection strategy and procedures to be used? Will they provide too little or too much data? Will they be expensive and/or time-consuming? Have alternative data sources or collection procedures been considered to assure that the most appropriate process has been picked? Will the data be adequate to perform the required analysis?

13. What specific methods and techniques are best suited to this evaluation situation? How can they best be combined to form an appropriate methodology? Will standard methods have to be adapted to meet special conditions of this study (i.e. different languages, etc.)?

14. Is there a clear and specific methodological strategy for the analysis to be performed? Have procedures for sampling, statistical analysis, and testing of the reliability and validity of the data been established? Will the analytical procedures produce meaningful statements leading to clearly-stated generalizations?

15. What resources are being made available for the study - money, calendar time, specialized skills, staff time? Will these resources be sufficient to fulfil study objectives, or must adjustments be made?

16. Has the composition of the evaluation group been considered in order to design a participative evaluation process which involves various viewpoints, kinds of experience, and skills to the extent possible? Is the evaluation seeking to create a collaborative effort rather than a judicial review? Should special expertise from elsewhere in the organization or outside the organization be enlisted for this evaluation?

17. Has a systematic preliminary work plan been prepared for this evaluation overall, covering such factors as key questions to be answered, methodology to be used, organizations involved, any expected constraints or complications, resources required, an implementation schedule, and specific responsibilities of the participants?

18. Is the overall evaluation study logically designed? Timely? Well-organized? Will anticipated results justify the costs of the evaluation?

B. Re-appraisal of basic conception and design of the activity

1. Was the activity based on an adequate diagnosis of the problem situation which it sought to address? If not, what was its origin and why?

2. Did the activity design provide a clear and reasonable overall justification of the purpose of the activity and a convincing rationale for undertaking it? Were alternative approaches, technologies and operating strategies considered before finalizing the activity design?

3. Are the objectives of the activity sufficiently explicit so that it can be determined whether they were achieved or not? Do they contribute to the solution of priority problems? Are they still relevant to priority concerns?

4. Do statements of the activity's objectives permit an understanding of the intended benefits and level of attainment? Identify recipients of unavoidable adverse consequences or indirect benefits? Include important qualitative aspects, even though they will be very difficult to measure? Take account of multiple objectives which may be complementary or conflicting?

5. Does this activity fit logically in the context of the larger activities of which it is a sub-part and relate meaningfully as a building-block for dealing with higher-level objectives, issues, and policies? Does it clearly identify and relate to similar, parallel, prior, or subsequent activities? Is it an appropriate activity for the organization in terms of goals, policies, resolutions, and procedures?

6. Are there specified indicators against which the attainment of objectives can be measured and impact determined? Do they appear reasonable, reliable, relevant, and valid? Were "baseline data" (original conditions prior to beginning the activity) adequately described?

7. What target group or clientele were to be affected by the activity? How were benefits to be distributed? What costs were involved and for whom? Do these cost and benefit concepts seem reasonable and achievable?

8. Are the measures to be used comprehensive, in that they quantify the extent of achievement of objectives ("effectiveness" measures)? Capture qualitative aspects of consequences ("intangible" measures)? Quantify unintended consequences where possible ("spillover" or "side-effect" measures)? And quantify differing impacts on beneficiaries and cost bearers ("distribution" measures)?

9. What major assumptions were made in designing the activity? Have these assumptions proved accurate? Were there other major considerations or problems which could have been foreseen but were not?

10. To what extent did the activity design anticipate and make provision for overcoming structural and environmental obstacles?

11. What other major changes - external, technological, political, institutional, economic - have occurred since the activity was conceived and designed which have a substantial impact on its conduct and expected results?

12. What kind and magnitude of outputs were to be produced with the inputs provided so that the objectives might be achieved?

13. Were inputs adequately described and adequate? Was the process and strategy for converting them into outputs plausible?

14. How plausible and logical was the overall design for the activity, in terms of linking inputs with outputs with intended objectives with expected impact? Were statements of causal relationships formulated and well-thought-out?

15. Were the various implementation actions and their sequence clearly defined and appropriate? Were various implementation strategies considered and a justification provided for the alternative which was chosen?

16. Was there a clear management plan for the activity, which described the work plan, budget, participation and support requirements, reporting requirements, review and evaluation procedures, time schedules, and completion and phase-out procedures involved?

17. Was the time lag between activity completion and achievement of activity purpose adequately identified? Were means of maintaining or enhancing the results after activity completion discussed?

18. What was the extent of consultation with and participation in activity design by the potential users of the activity's results? With other organizations involved in the activity?

19. Is the activity designed to permit the development of information for use in follow-up activities, institution-building, or for other purposes?

C. Analysis of implementation

1. Were there any pre-conditions to be met before activity implementation? Any specific design or testing phases? Was the activity to be implemented in a stage-by-stage sequence? Did any such actions take place as planned? If not, what were the consequences?

2. Have initial discussions been scheduled, if possible, with those responsible for the activity to get their ideas on problem areas or key matters for consideration in the evaluation? Will subsequent meetings be held to discuss findings of the evaluation with these people as the findings emerge, and to suggest possibilities for on-the-spot corrections where such opportunities exist?

3. Have inputs been provided as planned and work schedules and timetables adhered to? What has been the nature and effect of problems in providing resources (money, staff, and material) for the activity?

4. Have outputs been achieved as expected? How well have the outputs been measured? What has been their quality?

5. What was the adequacy, quality, timeliness and utilization of such factors as expert services, equipment, sub-contractor services, or fellowships?

6. If a host Government or Governments were involved, what was the adequacy, quality, and timeliness of governmental inputs, supervision, and cooperating agency performance? Of "counterpart" provision of buildings and physical plant; equipment and services; and technical, administrative and service personnel?

7. What monitoring, control and reporting systems are in place? What data do they provide for analysis of the operation of the activity? What is the quality and usefulness of this data?

8. Have appropriate policies, procedures, and standards been complied with during the implementation process? Has implementation suggested any need for re-consideration or revision of such guidance?

9. What were the problems causing delays in implementation and the consequences of such delays, if any?

10. Can operational problems be traced back to specific causal factors? Can these factors be determined to be problems which have significant negative impacts in other organizational activities as well?

11. Have any changes or corrections been made in the implementation strategy as originally planned? Why and with what effect?

12. Are there other important considerations, such as shifts in public opinion, communications processes, or personnel charateristics, which are not readily evident from normal activity data but may nevertheless be significant causes of the activity's operational success or difficulties?

13. Have any of the managerial aspects of the activity been handled particularly well or innovatively, so that they might be applied more widely to similar or other activities?

14. Are other organizations (United Nations agencies, or other international, regional, bilateral, or national organizations) concerned with the activity? Have they been consulted or directly involved in implementation? To what extent and with what effect?

15. Does the activity overlap with other activities? Could the activity be more efficiently carried out by another organization?

16. If implementation is still on-going, should objectives and/or targets be re-stated? Should progress indicators be made more precise? Should resource inputs be increased, decreased, or shifted? Are there alternative ways to operate the activity which require fewer or less expensive inputs, take less time, produce less expensive outputs, or produce greater impact for the same level of effort or expenditure? Should any such changes be made on an urgent basis or in the future?

17. Even though the results or impact may be small, is it possible that larger impacts are not brought out in the available data or would be achievable with considerable change in design or implementation of the activity?

18. Based on the results achieved, have a range of alternative possible decisions been considered and compared, including continuing, modifying, expanding, reducing, or abandoning the activity, as well as entirely new alternative activities?

19. Could improved technical or cost efficiency have been achieved in implementation through a different choice or mix of such elements as organizational level, methods, manpower, finances, facilities, participation, or managerial controls?

20. Overall, what where the strengths and weaknesses of the activity's implementation?

D. Evaluation of results

1. Was the activity effective in that it achieved its objectives? Did production of outputs lead to achievement of immediate objectives? Was it effective both in achieving its immediate objectives and in contributing to the achievement of higher-level, longer-range objectives? How can this be determined and demonstrated?

2. How relevant are the activity's results to longer-range objectives and priority needs? Do other activities which will utilize this activity's results exist, or will they be created in a timely fashion? Are the activity's results well-timed to help achieve the long-range objectives?

3. Did the objectives of the activity change over time? Did implicit objectives exist or emerge?

4. Has implementation produced the desired impact? How has this been determined? Are there alternative measures or criteria for assessing the impact achieved?

5. Does analysis of change which has occurred because of the activity consider not only its amount, but its nature, rate and direction - as well as unanticipated changes which have occurred?

6. What factors were most important in causing the results which occurred? How did these key factors affect the outcome?

7. To what extent have the changes which occurred been attributable to this United Nations system effort? To other causal factors? To what extent and how can these specific impacts be measured?

8. Does the Government or other beneficiary plan to use the United Nations system contribution? Were they satisfied with the results? How extensively did they cooperate with the activity?

9. Was the cost of achieving the objectives generally reasonable in light of the priority attached to the activity and the share of the organization's resources used?

10. Can an analysis of costs and benefits, whether precise or relatively imprecise, be made? Can any unintended costs and benefits be identified and analyzed?

11. Are there alternative ways of achieving the activity's objectives that are less costly and/or more effective? Can relevant costs and benefits be considered so that some ordering or rankings of the alternatives can be made, however rough it might be?

12. If data on results is conflicting, can the conflicting interpretations be reconciled? Are the differences due to data collection methodology? Can the significance of the difference itself be analyzed, perhaps by experts in the subject field?

13. Could additional information about the activity significantly change the evaluation of the results? How, when, and at what cost could such information be obtained? Is it worthwhile to postpone decisions while new evaluations, studies, or research efforts are undertaken?

14. Although clear results may not be obtained, can the evaluation findings be used to suggest a structure for further research and evaluation? Can the findings also suggest new or alternative strategies which should be assessed?

E. Reporting and follow-up

1. What are the key lessons which have been learned from this evaluation? How can these lessons be transferred in a timely fashion for the improvement of this activity or the planning and design of future activities?

2. What factors have been identified as significantly favourable or innovative in the conduct of this activity? Can they be applied or utilized elsewhere? What factors have been adverse to conduct of the activity? What corrective actions have already been taken, and what specific further actions are needed? Are these problems symptomatic of more widespread and systemic problems? Might they be found in other activities?

3. Has this evaluation study identified any new, emerging problem areas which should be noted in reporting or considered for future evaluation? Can it be appropriately linked to current issues of major concern to decision-makers?

4. Do the findings of this evaluation significantly support, strengthen, or contradict past evaluation studies, or do they suggest new solutions for problems which were identified in past evaluations?

5. Was the purpose of the evaluation found to be valid? Was the original scope, problem or issue involved re-defined as part of the study or as a result of the study?

6. Has full consideration been given to the information, analytical, and reporting needs of potential users of the evaluation?

7. How can reporting be done in the simplest and most expeditious way possible in order to feed information back to decision-makers, planners, and programmers in a timely and direct fashion?

8. Does the report contain an organized structure of findings, conclusions, and recommendations? Are the findings adequately supported? The conclusions logical and justified? Is the reporting clear and concise?

9. Does the reporting product clearly and succinctly state why the study was made? Who did the study, and when and where? How the study was conducted, what procedures were used, what information was collected and how was it analyzed?

10. Are there other ways in which the results of the evaluation can usefully be disseminated, in addition to written or oral reporting, such as publications, workshops,and seminars or training sessions?

11. Are the findings significant and relevant? Do the findings and conclusions agree with other studies or other evidence, and if not, why not? Do they answer key questions posed at the beginning of the study?

12. Were any special problems (conceptual or practical) found in using various measures? Was it necessary to employ alternative or "surrogate" measures and if so, why were they chosen?

13. Are recommendations specific rather than truisms? Are they action-oriented and feasible rather than mere exhortations for improved performance?

14. If the evaluation results still leave a high degree of uncertainty about the activity, has the uncertainty been clearly communicated and the recommendations appropriately qualified?

15. Can specific suggestions be made, based on the evaluation, to improve the linkage and coordination between and among organizational evaluation, planning, programming, budgeting, operational work plans, and information and performance reporting systems?

16. Were participants in the activity and other interested groups given the opportunity to review and comment on the evaluation findings? Are their comments reflected in the report?

17. What is still left to be studied? What new questions were raised that require further research or exploration? What research methods need to be developed or improved in order to make future evaluations more effective?

18. Will decision-makers need further assistance in interpreting the report? Getting answers to questions raised but not answered in the report? Undertaking further action? Is the evaluation staff prepared to provide such assistance, so that the full value of the evaluation may be achieved?

19. Have provisions been made for subsequent follow-up to determine that appropriate corrective actions have been taken?

ANNEX II

PARTIAL BIBLIOGRAPHY OF EVALUATION GUIDELINES WITHIN THE UNITED NATIONS SYSTEM

Administrative Committee on Coordination (ACC), Report of the Sixth Session of the Inter-Agency Study Group on Evaluation, ACC document CO-ORDINATION/R.925 of 29 February 1972.

Committee for Programme and Co-ordination (CPC), "Report of the Committee for Programme and Co-ordination on the Work of Its Seventeenth Session", Official Records of the General Assembly, Thirty-Second Session, Supplement No. 38, (A/32/78).

Committee for Programme and Co-ordination (CPC), "Report of the Committee on Programme and Co-ordination on the Work of its Eighteenth Session", Official Records of the General Assembly, Thirty-Third Session, Supplement No. 38 (A/33/38).

Economic and Social Council (ECOSOC), "Report of the Joint Inspection Unit on Evaluation in the United Nations System (JIU/REP/77/1): Comments of the Administrative Committee on Co-ordination", ECOSOC document E/1978/12 of 31 January 1978.

Food and Agriculture Organization (FAO), Evaluation Service, "Guidelines for Evaluation of Technical Assistance Activities", FAO document DDDE:INF/74/1 of February 1974.

International Labour Organization (ILO), "Procedures for the Design and Evaluation of ILO Projects", of January 1979.

International Trade Centre (ITC), "Enhancing the Effectiveness of Technical Assistance in Export Promotion: Proposals for a Project Evaluation System for the International Trade Centre UNCTAD/GATT", ITC document ITC/AG(VIII)/40 of November 1974.

Joint Inspection Unit (JIU), "Report on Evaluation in the United Nations System", JIU/REP/77/1 of March 1977. (and ECOSOC document E/6003)

United Nations, Report of the Secretary-General, "Programme Evaluation for the Biennium 1974-1975", (including Annex,"Office of Financial Services' Guidelines for the Evaluation Reports"), ECOSOC document E/AC.51/80 of 12 April 1977.

United Nations, Report of the Secretary-General, "Programme Evaluation for the Period 1974-1977", including Annex," Revised Guidelines for Internal Evaluation"), ECOSOC document E/AC.51/91 of 16 May 1978.

United Nations Development Programme (UNDP), "Report by the Administrator: Addendum: Joint UNDP/Agency Evaluation Programme", DP/319/Add.1, and Annexes I - III of 4 and 5 April 1978.

United Nations Educational, Scientific and Cultural Organization (UNESCO), "Preliminary Report of the Director-General on Studies Carried Out Concerning Evaluation of Programme Activities", UNESCO document 103 EX/10 of 11 August 1977 and ECOSOC document E/1978/111 of 3 July 1978, Annex III.

United Nations Educational, Scientific and Cultural Organization (UNESCO), (Constantin G. Soumelis), "Project Evaluation Methodologies and Techniques", UNESCO, Imprimerie Maury, 1977.

United Nations Children's Fund (UNICEF), "Criteria for Programme Appraisal, Monitoring, and Evaluation: Notes and Recommendations by the Executive Director", UNICEF document E/ICEF/P/L.1628 of 19 January 1976.

United Nations Industrial Development Organization (UNIDO), "Evaluation of Selected Activities of UNIDO", Report of the Expert Group Meeting on Approaches, Procedures and Methodology of Evaluation of UNIDO's Activities, UNIDO document ID/B/C.3/48/Add.2 of 6 April 1976.

United Nations Institute for Training and Research (UNITAR), "Criteria and Methods of Evaluation: Problems and Approaches", Unitar Series No. 1, ECOSOC document E/4649 of 2 May 1969 (also published as "UN Development Aid: Criteria and Methods of Evaluation", New York, Arno Press, 1971)

World Bank (IBRD), "Guidelines for Monitoring and Evaluation of Rural Development Projects", Agriculture and Rural Development Department, IBRD, of January 1977.

World Health Organization (WHO), "Development of Health Programme Evaluation", Report by the Director-General, WHO document A31/10 of 29 March 1978.
