

**EVALUATION OF UNITED NATIONS
TECHNICAL CO-OPERATION
PROJECT EVALUATION SYSTEMS**

Part I

Prepared by

Alain Gourdon

Joint Inspection Unit



Geneva

1988

EVALUATION OF UNITED NATIONS TECHNICAL CO-OPERATION
PROJECT EVALUATION SYSTEMS

PART I

by Alain Gourdon
Joint Inspection Unit

TABLE OF CONTENTS

	<u>Paragraphs</u>	<u>Page</u>
INTRODUCTION	1 - 12	1
I. <u>Project evaluation and logic</u>	13 - 42	3
Relationship between the evaluator and the evaluated	15 - 19	3
Degrees and levels of project evaluability	20 - 28	4
Limitations and risks from the standpoint of the project	29 - 35	6
Evaluation objectivity and functionality	36 - 42	8
II. <u>The project approach and the country approach</u>	43 - 77	9
Assessment of sustainability and impact	43 - 55	9
Impact co-ordination and competition	56 - 62	11
The need for a country-specific approach	63 - 77	13
III. <u>The improvement of evaluation procedures</u>	78 - 113	16
The time-lag between manuals and theory	79 - 87	16
The timeliness of procedures	88 - 97	17
The adjustment of the various types of evaluation	98 - 107	19
Modifications to guidelines	108 - 113	21

INTRODUCTION

1. This study does not purport to review all the problems which - it is claimed - have arisen in connection with the evaluation of projects undertaken by or for the United Nations in a particular region during the past few years. It does not seek to identify the results of the evaluations carried out or to consider the conclusions that may be drawn from them. The aim has been more limited yet more ambitious, namely to evaluate the evaluation systems used in the United Nations context. It is thus concerned with projects rather than programmes, which are in fact dealt with in another JIU report. It is an overall examination of:

(a) The concept, methods and practice of evaluating country projects, ignoring regional projects;

(b) The implicit philosophy and the content of the "guidelines" for the evaluation of such projects and the interaction of the evaluator, the evaluated and the inferences regarding evaluability;

(c) The conjunctions and/or disjunctions of monitoring and evaluation, their respective importance and the areas they cover;

(d) The new constraints affecting all evaluation systems, as a result of the increasing recourse to donor Governments for project execution;

(e) The activities and organization of the services entrusted with evaluation and any reforms or changes that might be required in order to improve their administration and functioning in the United Nations framework.

2. To ensure that the study was not purely theoretical, and that it would result in recommendations for less rigid procedures and for organizational reform, it seemed necessary to take a practical approach and to illustrate it by examples taken from the experience of several executing agencies: FAO, UNIDO, UNESCO, UNICEF, and UNHCR. These agencies, which work in co-operation with UNDP, were selected not only because of the role they play but also because of the variety, the heterogeneity, in fact, of their functions, aims and approaches. The Asian and Pacific region context was chosen because it offered the greatest possible diversity of economic systems, territorial sizes, geographical situations, socio-cultural contexts and levels of development.

3. The report contains no fixed terms of reference. The investigations of a JIU inspector charged with a given mandate could not be made to conform to the procedures governing the drafting of reports by international functionaries subject to the hierarchy to which they belong. The value of the research conducted by an inspector does not derive from the observance of given terms of reference which, in the case of an individual rather than a team exercise, he alone would be properly qualified to impose on himself. The main value of such studies comes from discoveries made in the course of the investigation and from the inspector's ability to change the framework or direction of his inquiry, without of course losing sight of his set goal.

4. Although no claim is made to an encyclopedic knowledge of all evaluation systems, which would actually be beyond the resources of even the most experienced specialist, the report's recommendations are based on an analysis of a wealth of administration and academic material, and in particular on observations made in the field and visits to projects, and on interviews with a hundred or more officials of recipient Governments and specialized agencies in a dozen developing countries in South-East Asia and the Pacific region during a one-month mission in 1987.

5. One criticism made of the preliminary draft of this report was that it did not contain large numbers of quotations or texts of resolutions or justify

each observation by a reference to some text or other. Suffice it to say that there can be no question of overloading a report of this type with documentation; that may be necessary for the support of a thesis, but not when the aim is to produce practical recommendations. The study relies more heavily on sources or conversations to which certain considerations of confidentiality apply, rather than on reams of written questionnaires which, in most cases, elicit replies that are too diplomatic or too rudimentary to be usable.

6. The preliminary draft was also felt in some quarters to be subjective or tendentious. Such critics have simply forgotten that this so-called "subjectivity" stemmed from the fact that the discussions were off-the-record and needed editing. True objectivity, whose conclusions rarely find their way back to headquarters and which is most easily recognized by a JIU inspector in the field, is more often found in the spoken than in the written word.

7. The present report appears at a time when the United Nations evaluations systems seem to be fairly new yet sufficiently established for their methodologies and orientation to be examined without fear of calling into question the reason for their existence. The report follows the principles adopted by JIU and, in particular, by Inspector Sohm in his 1979 and 1985 reports on evaluation questions. The recommendations made are therefore of a quite general nature and cut across the dividing lines between executing agencies. Evaluation is not treated as an isolated management tool, but as a technique which forms part of a set of decision-making procedures within the institutions studied.

8. The time now seems to have definitely passed when the pursuit of objectives defined in unduly vague terms and efforts to set up financing for operations and to provide inputs took precedence over such matters as the appropriateness and effectiveness of projects and consideration of the nature and true interest of their beneficiaries. The events which have promoted this awareness of the usefulness of, and even need for, evaluation are well known and may be recalled only briefly. The budgetary difficulties experienced by donor countries following the "oil shocks" made them follow more closely the use that was made of their contributions to multilateral aid. As a result of the relative decline in the funds made available, the donor countries began to pay more attention to the advantage to them of a reduction in aid. At least, their Governments were frequently urged to consider this aspect while, at the same time, a number of developing countries were themselves becoming net donors.

9. It is now accepted as a fact that evaluation is a normal and effective project management tool. In this regard, the UNDP administration has played an important role. There have also been many reports by JIU, which have done much to secure acceptance of the idea of evaluation, to popularize its vocabulary and to extend the introduction of its methods and support systems in the executing agencies. The financial crisis of the United Nations and the recommendations of the Group of 18 aimed at strengthening the machinery of the Organization have only made more imperative recourse to what had first been a palliative and has since become a necessary element of project activities, an increasingly felt requirement, a crucial administrative need, in some respects a fashion and sometimes even an alibi.

10. Since 1981, there has been no lack of General Assembly resolutions affirming and reaffirming support for the implementation and organization of evaluation capacity in the executing agencies. Resolution 36/228 of 1981 encouraged the agencies to co-operate with JIU in this regard to that end.

The annex to resolution 37/234 of 1982 defined evaluation as a system which allows for continuing critical review of achievements and collective thinking thereon. Resolution 38/227 of 1983 requested the Secretary-General to review all the possibilities available to strengthen the system thus defined. By its resolutions 40/259 of 1985 and 41/213 of 1986, the General Assembly reiterated even more strongly the determination it had repeatedly expressed and placed special emphasis on the necessity of using the evaluation capacity of the United Nations to provide Member States with a basis for more informed decision-making.

11. It is in this light that the present report should be seen. Far from seeking to contradict the conclusions of earlier JIU studies, it sets out to continue the investigation already begun on the question of evaluation. Evaluation is not an immovable object or a fixed invariable dogma. It is thus quite natural that it should be subject to significant changes of direction, that its methods should be reviewed, its types separated out and new courses charted.

12. Experience has shown that the priority aims of evaluation and the obvious need for some types of evaluation have become less clear with the passing of the experimental and inaugural period. The forcefulness of the criticisms and approval with which the preliminary draft of this report was received are proof - if any was needed - not only of the importance and relevance of the questions raised, but also of the growing uncertainty and uneasiness which they arouse and which must be overcome.

I. PROJECT EVALUATION AND LOGIC

13. The evaluation of development operations within the United Nations, which is a historically recent innovation, coincides with a marked change in the approach to this process, both at the level of the social sciences and applied research and at the level of governments. It was the Governments of a number of industrialized countries that first spread this practice. This necessarily gave impetus to the use of newly developed methods and affected the design and implementation of projects executed by specialized agencies of the United Nations system.

14. Evaluation initially took the form of a survey. The survey, based on more or less detailed questionnaires, made it possible to learn the views and aspirations of a sampling, or of a social segment, and then to analyse the replies in which each respondent made a more or less personal judgement as to adequacy. The survey was an undeniable improvement over previous information-gathering activities, but its validity was nevertheless open to question. It had the particular disadvantage of being very costly without seeming to produce results commensurate with the expenditure incurred.

Relationship between the evaluator and the evaluated

15. From this survey concept, whose methodology seemed too subjective and too narrative in nature, it was seen fit to proceed to that of an evaluation system which attempted to determine "as systematically and objectively as possible" (JIU/REP/79/2), the relevance, effectiveness and impact of organizational activities in the light of their objectives. Aiming to be scientific, and more quantitative than qualitative, the new method stresses the relationship between cause and effect, means and end, and input and output. Its object is to monitor a posteriori the proper execution of the project by seeing to the smooth interaction of the resources, goods and

services with the activities, of the activities with the output and of the latter with the objectives.

16. The project, which is thus assimilated to a process, but whose results are nevertheless termed output, is described as being based on a logic which is partly mechanical and linear, partly analytical and also to some extent abstract. There is a presumption that serious planning and proper management will necessarily result in conversion of the inputs into output. The project is seen as a series of means and ends, a continuous chain of causes and effects, a hierarchy of levels, which follow and are superimposed on each other. It has the appearance of a sequence of intermediate ends, which are converted into means, of a composition and decomposition of successive phases and of effects which become causes. In it, the mechanical prevails over the organic and human.

17. Technical, tactical and compartmentalized, the projects, the subject of guidelines adopted by UNDP and the executing agencies, amount to a Taylor-style division of tasks. They function much more as isolatable entities than as the components of an integrated whole. It is as if the project evolved in an inert medium, as in a laboratory experiment, rather than in the climate of real life; as if the project, in taking form, were not destined to become a social reality and as if social realities were things; as if economic phenomena, in the developing countries, in particular, were not primarily cultural.

18. However, while evaluation can be applied to thematic subjects or, more rarely to countries, it will function most frequently in relation to a project, a group of projects or sometimes a programme, although it is not always possible to see what distinguishes the last-mentioned from a collection of projects linked together in mediocre fashion. Obviously, therefore, the design and construction of the project cannot fail to influence the conduct and performance of the evaluation. The evaluated is reflected in the evaluator even before the evaluator apprehends and shapes the evaluable.

19. In defining the method of using the project management tools that are design and appraisal, monitoring and evaluation, the manuals of UNDP, of the executing agencies considered and of ESCAP concern themselves with the evaluator and the evaluable, making only parenthetical reference to the differences which may result from the place and context of the exercise. The validity of the evaluation, according to the implicit philosophy of these guidelines, is neither connected with nor based on the specificity of the project to be evaluated. From the principle of the equality of Member States and the sacred necessity for development, postulated by the consensual United Nations viewpoint, the guidelines see fit to derive a logic from the universality of rules governing multilateral aid operations, as well as a correlative practice of applying standard procedures for evaluation.

Degrees and levels of project evaluability

20. Of course, the guidelines do not fail to take into account the external factors, which are defined as events, actions or situations which cannot be controlled or significantly affected by the project operators but which are nevertheless of importance for its success. They are, therefore, mentioned in the format of the Project Performance Evaluation Report (PPER) provided for in the UNDP procedure. Although divided into classes (institutional, political, socio-cultural, economic and "other"), these exogenous variables do not receive much attention. They are not mentioned at the design or appraisal

stage, but only at a certain phase of the evaluation, when the person in charge of the project is requested to record whether the project output is satisfactory or otherwise in relation to internal operational factors connected with the inputs. It is as if the external factors had to be taken into account downstream rather than upstream, in the course of the project, rather than at its inception, at the risk and results level rather than at the context and input levels.

21. The "external factors" are therefore matters of concern only as regards the monitoring and evaluation covered by the PPER. Although mentioned in the forms and working documents in the annex, they are not referred to in the text of UNDP's chapter on monitoring and evaluation. Limited to the PPER framework, they do not appear in the format dealing with the terms of reference of in-depth evaluation. Not only does the UNDP manual give them only a modest place, but their listing, which moreover does not have to be exhaustive, fails to mention some whose importance cannot be underestimated. Examples of these external factors are the influence of the parallel network of bilateral aid, the concurrent or complementary activities of NGOs, and the quality of the statistical and documentary machinery of the country in which the project is carried out.

22. From the project design and preliminary consideration stage, through monitoring to evaluation, the manuals of UNDP and the executing agencies tend overly to dissociate the projects from their socio-cultural context, from the level of development already achieved and from the orientation of the economic régime. Regardless of whether basic data collection is affected or not by the adequacy or otherwise of the information storage system, the solidity or otherwise of the administrative structure or the efficiency or otherwise of the internal communications, evaluation uses the same effectiveness standards. The manuals make no distinction between least developed countries, developing countries and newly industrializing countries, for which the evaluation should, without any alteration of its basic significance, at least apply relatively different assessment criteria and employ different progress indicators.

23. However, the format of the "Guidelines for project design" to be applied by UNIDO from 1988 provide for a context section which takes account of the description of the economic sector to which the project belongs, the development strategy of the host country, the programme and institutional environment and the administrative infrastructures. This format has the advantage of being upstream, but the shortcoming of not taking account of the socio-cultural context, or the demographic and political factors which inevitably have an influence on the project.

24. The UNDP guidelines also seem no longer to deal with the problem of the relative pre-determination of any evaluation by the very nature of the project to be evaluated, so that the possibility of a solution in the field has not received the attention it deserves. They do, of course, make a functional and objective distinction between projects according to whether they are aimed primarily at the setting up or strengthening of institutional structures (research, planning, education), at providing direct support (studies, surveys, advisory services) or at training (courses, seminars, workshops, fellowships, study travel). They also classify the operations as pilot or experimental projects, aid projects or projects connected with investment or preinvestment. They even provide for a special category of projects, which, under the cover of support for national projects, offer donor countries a means of securing administrative and logistic support. For the purposes of a

kind of functional programme budgeting at the United Nations scale, the manual's classification provides valuable information and enables useful breakdowns to be carried out.

25. However, it does not provide any outline document which is sufficiently indicative, without being too restrictive or detailed. It is not combined with a formula for determining relative measurability or with a model of project evaluatability levels which takes account of the project's objective, role and dominant aspect. What is worse, a classification of this kind ignores the differences between agriculture and industry, between administration and research, between culture and health and between the vocational and the social. It therefore inevitably neglects the specific impact of the diversity of these aspects and characteristics on the feasibility and effectiveness of evaluation.

26. The manuals should pay rather more attention to the relationship between the particular nature of the project and the project's evaluation. It is not possible always to follow the same pattern or to place in the same mould projects which have very different goals and whose implementation processes do not move at the same pace or are not adapted to the same methods. The amount of design, monitoring, preliminary study and a posteriori evaluation should vary according to the activity and sector involved.

27. Simple examples will serve to illustrate these points. Whether airports, bridges or telecommunications are involved, the projects require a small number of partners, thorough feasibility studies are made, the competence of those involved in the work is known and, unless there is a major error or exceptional construction defect, the project's execution will follow a predictable course. Monitoring will be more important than an evaluation which, occurring too late, will be able only to note failure or any possible damage. It will be less important than design and preliminary study, since the major difficulties arise at the stage of preparation and not during the execution of the project.

28. In the case of agriculture or health projects, the multiplicity of possible partners and beneficiaries, the variety of obstacles that are unforeseeable at the outset make it necessary to concentrate less on the project design and to place greater emphasis on learning by doing and on monitoring. Integrated rural development projects have been rightly compared to sailing uncharted seas where the course and heading have to be altered according to the discoveries made. The project design and identification of requirements must not be neglected, but will in this case be of less importance than experimentation and corrections during the operation, the safeguarding of options and the search for flexible solutions. In this case, it is on monitoring rather than on evaluation that stress will have to be placed.

Limitations and risks from the standpoint of the project

29. The nature of the project influences the type of evaluation and the weight that must be given to design and preliminary study, to monitoring and evaluation, whether detailed or otherwise, to ex ante and to ex post, to corrections made in the course of execution and to preparation, impact and follow-up. Overly standardized rules, without sufficient attention to the specificity of projects, prevents any conclusive project evaluation.

30. The logic of a project and of evaluation, on an individual project basis, which is its corollary, places the context in parenthesis and ignores particular specificities, which means that an accessory factor or a negligible

quality may be treated as a nuisance or constraint. This may be vital in determining the success or failure of the operations. This same logic, by restricting the coverage of an evaluation to a particular project or group of projects, limits the evaluator's aim and offers him only a distorted view of what he is required to assess.

31. The project's execution, however suitable and effective it may be in the eyes of its evaluator, may in fact conceal problems and make their solution more difficult. By concentrating the donor's attention on a privileged implementation and by allocating the available national resources to him on a priority basis, it aggravates difficulties which do not surface, but which nevertheless hover in the shadows. A project-by-project approach helps to raise barriers between the ministries concerned in the beneficiary country and to complicate the financing of measures, whose cost and point of impact have not been the subject of planning. It tends to confiscate for its benefit resources for recurrent expenditure once the project has been completed and aid has ceased. However, the evaluator will not take account of what may have been abandoned or lost in less privileged sectors owing to the execution of the project he has to evaluate.

32. It will, of course, be claimed that there is coherence and continuity in the donor Government's planning, which should in principle avoid this kind of distortion and bottlenecks. However, the situation with regard to plan management in a number of developing countries and the rivalry between departments (finance, planning, technical ministries) are well known. There is general awareness of the necessarily disturbing influence of projects on a country's planning and of the competition which then takes place between donors of all kinds. This is naturally prejudicial to the smooth preparation of programmes and the harmonious development of project design.

33. From the sectoral standpoint of international organizations and their specialized staff, the execution of the project and the rate of disbursement of funds may well seem excellent. On the basis of the forms and questionnaires the evaluator will conclude that the project has been a success. However, the actual result may in fact prove much more debatable as the mobilization of the financial, moral and administrative resources of a target group is sometimes underestimated. No steps may have been taken to associate the beneficiaries with the undertaking and their willingness to participate in the development process will have been discouraged. Thus, in one Asian country, it was noted that the local traditions of community development and sense of initiative had greatly deteriorated in a peasant environment which multilateral assistance had somewhat paralysed instead of making it more dynamic.

34. If the donor and the project management team seek to play an unduly pre-eminent role and aim only to lighten the burden on the recipient country, by writing off or reckoning as nil the potential contribution of its officials and its target groups, they will deprive them of a source of knowledge and experience, of know-how and of a psychological input which are inherently unassessable, or at least impossible to evaluate from the standpoint of the project.

35. Similarly, not only national administrations, but also local firms, should be involved in the planning and technical consultations, even if their effectiveness measured in terms of cost and benefit proves lower. Development does not consist in meeting the needs of target groups through more or less massive injections of aid, whether multilateral or not. It is a social process and not an output. It therefore has a price, which evaluation, a

prisoner of the project's logic, is unable to assess. Social facts are also psychological facts, and evaluation loses all status once development eludes the logic of the project.

Evaluation objectivity and functionality

36. Since evaluation aims to be as rational as the projects to which it is applied, it seeks ways of enhancing its objectivity by extending indefinitely the scope of its applications. For example, the preliminary study of a project may be treated as an ex ante evaluation and one hears of "self-evaluation" by someone who is required to be both judge and party in his case. The types of evaluation are classified according to the intensity of the exercise, the origin and type of the evaluator, the subject of the evaluation and the amount of resources allocated to it. An evaluation may also be recapitulative or retrospective, periodic or ad hoc. Sometimes the name "evaluation" may be used for what is only an "assessment", or may function as such, while assuming the title "programme audit".

37. In view of the interest in evaluation that seems to be developing in the United Nations and its specialized agencies, there is now a tendency to employ this term to cover an ever-increasing number of reports and studies which have little connection with genuine evaluation. Evaluation, to some extent a victim of its own success, seems both a cure-all and an administrative expedient. When everything is called an evaluation, the real meaning of the term becomes doubtful. Oscillating between the myth and the dictionary definition, the concept only loses in precision what it gains in scope.

38. Under these circumstances, if the postulated objectivity of evaluation is to be retained, the corollaries must be observed. It is necessary to ensure that the useful information reaches the proper person at the right time or at pre-established intervals, to ensure that urgent or important matters do not reach those concerned too late, and to prevent superfluous or secondary information from submerging those who are only indirectly concerned; requirements such as these are simple common sense. This does not mean, however, that they have therefore become general practice, even though they constitute the necessary complements of an objective evaluation.

39. The manuals of UNDP and the executing agencies do not seem, upon examination, to have paid sufficient attention to the question of ensuring that the questionnaires and forms are functional and practical to use. The filling of boxes in questionnaires automatically provides answers to questions that the operators might have neglected or forgotten. They help to feed data into the computers at headquarters. However, it is not always possible to answer with a yes or a no, or to give an appropriate reply by ticking one box rather than another. Excessive disdain of narrative and over-reliance on standard forms involve the risk that only the trivial may be recorded. In the absence of spot checks, of which there is rarely any fear, and without a general inspection system, which has so far been lacking at the institutional level, the validity of the replies given can be doubted in many cases.

40. A simplistic standardization of forms, which may be justified for airport immigration checks or materials control purposes, is inherently detrimental to an objective evaluation. Over-sophistication should likewise be avoided. At the end of 1986, UNICEF published an "Audit Guide" to be used in training courses for auxiliary field staff. This 150-page guide makes provision for all eventualities and seems to take it for granted that the countries subject to audit possess administrative and statistical facilities that it would sometimes be difficult to find in the most highly industrialized countries.

41. This exhaustive guide, which takes the form of a questionnaire, must have taken a long time to prepare. Its authors simply forgot, however, that a correct reply would necessitate several months of full-time work on the part of the field staff concerned, to the detriment of other equally urgent tasks. Moreover, its authors have not taken sufficient care to make the evaluation manual usable. The dominant trends and the principal facts to be established disappear in the multiplicity of questions asked, which are not arranged in any order of priority. This example, albeit extreme, is not an isolated one, but illustrates a general trend which must be brought to an abrupt halt.

42. There can be no objective evaluation without respect for priorities, which implies an orderly classification of conclusions. Although there is obviously a need for clear answers, it is equally necessary that the questionnaires and forms should require the evaluator to highlight the salient points and significant figures. If the forms are over-simplified in order to establish only what is essential, only the measurable and the commonplace is recorded, with a risk of the most interesting aspects being neglected. The converse may occur if the forms are over-complicated. To say everything in rudimentary fashion is tantamount to saying nothing at all. To require that everything be said is another way of agreeing to be told nothing.

II. THE PROJECT APPROACH AND THE COUNTRY APPROACH

Assessment of sustainability and impact

43. UNDP is justifiably anxious to determine whether the output of the projects and programmes it finances is lasting in nature. Working on the assumption that the "sustainability" of a project "depends on the ability of the target groups or institutions ... to benefit from the project results" in a continuous and effective manner, UNDP has reached the conclusion that this constitutes an objective that is affected by numerous variables, which must be taken into account at both the project design and the project implementation stages, as well as after project completion. It is necessary to establish the extent to which the design of the project may have compromised or favoured this "sustainability" and how a retrospective assessment may be of use in the thematic evaluation of its implementation. It is also necessary to obtain a better understanding of and make a more thorough study of how and why the results of a project are continuing, or ceasing, to contribute to the overall development effort.

44. The governing bodies of executing agencies, and the FAO Council in particular, are increasingly looking into the question of project "sustainability" or viability. The first lessons drawn from this "sustainability" approach seem to have proved fairly useful, although the practical consequences are still to be determined. They have shown that it is worth stressing the question of "sustainability" not during the implementation of the project, but at the preparation and design stage. In the case of agricultural projects, for example, production, marketing and the socio-cultural context must be taken into account just as much as biotechnology and vocational training.

45. It has been noted (see "Comparative results of UNFPA evaluation", May 1986) that many projects would have benefited greatly from a better understanding of recipients political decision-making machinery. It has also been found that the donors' willingness to co-operate is the most important factor in a project's success, but that their underestimation of the complexity of the administrative and logistic problems and their relative inability to make essential contributions, significantly affect effective

utilization of the project output. It has further been observed that training projects should be evaluated within the context of the overall economy, since their effects depend to a large extent on infrastructures and on manpower utilization policies in the light of the national plan and/or the market situation.

46. The UNDP officials are thus forced to acknowledge that evaluation is inevitably limited in scope if it is adapted to the project logic and viewed within its perspective. Since in-depth evaluations are carried out only in the case of a number of projects selected on the basis of various criteria relating to their usefulness, complexity, type and cost, they have to admit that, in so far as these criteria imply a project-by-project approach to evaluation planning, the conclusions of the studies undertaken vary and the results are not easy to classify. They therefore have to concede that even a representative set of evaluations for a given country does not provide a means of identifying the major problems of general policy to be dealt with within the context of a programme (DP/1987/21).

47. Although they are aware of the importance of project "sustainability" and even have some idea of the relative, limited and even distorting effect of an evaluation that is confined within its own framework and logic, the UNDP officials do not seem to have fully grasped the practical implications of their observations.

48. According to the theory of evaluation generally applied by the United Nations, a good project is one that makes it possible to achieve an objective as rapidly and as correctly as possible, at the lowest cost and with maximum yield, in such a way as to optimize multilateral aid and strengthen the basis and means of economic development in the beneficiary States. The requirements thus laid down are necessary, yet nevertheless insufficient. The in-depth evaluation which may be carried out on project completion, and which can obviously be very costly, may prove to be of little benefit if it disregards or fails to overcome the problems of impact.

49. In fact, a project may meet the cardinal criteria of relevance, efficiency and effectiveness and may have been executed in the most appropriate manner and on time, from the standpoint of its objectives and specifications, but nevertheless deteriorate very rapidly and lose all its impact, because the recurrent expenditure that it entails is too heavy or not provided for in the beneficiary country's budget, because the political intentions of the government have changed, because the socio-cultural environment nullifies its medium-term effects, or because the administrative and technical means employed or the human resources used are unable to ensure the managerial follow-up, which is the responsibility of the recipient State. It should also be borne in mind that the impact of projects is frequently affected by a lack of foreign exchange or an adverse balance of payments. This observation is confirmed by the following examples.

50. A project for the eradication of encroaching weeds may be implemented to perfection and its results may even exceed normal expectations. However, its evaluation will be of no benefit whatsoever if, in the meantime, the Government's priorities have been reoriented from the subsistence to the commercial agriculture sector, with the result that it neglects maintenance of the technical control system that has been established, thereby nullifying the results previously obtained.

51. A research project may produce the desired effects. Its economic and technical parameters will be considered satisfactory. From the standpoint of

relevance and efficiency, it will be irreproachable. However, its results will remain questionable since an ex post evaluation will not be able to indicate how the results of the applied research will find a market, or how they will contribute to the substitution of one product demand for another.

52. A training project may achieve its objectives and the work done may initially lead to satisfactory results, yet the medium-term outcome may be a failure. The recipients of fellowships, who will have derived full benefit from the excellent training given to them, will leave after a few months to seek work abroad where they can gain financial benefit from training that was intended to promote development in their home country. In such a case, the impact will be nil. It may also be diminished or deflected if the recipient of the fellowship, even if he does not join the brain drain, is given a promotion as a result of which the post for which he was trained is left vacant.

53. When reduced to its object and its limits, a major agricultural project that proves highly commendable and efficient will, in the long term, have a disastrous effect on the social environment and will be the indirect cause of political disturbances whose cost will be out of proportion to the project's anticipated economic benefit. In contrast, in the case of a rudimentary and extremely low-cost irrigation system created elsewhere by experts from another country in the same region, the demonstrative effect of this inexpensive project, which will not have been taken into account in any evaluation, will be very far-reaching.

54. In some cases, no evaluation will be able to establish even the medium-term catalytic or multiplier effect of a project. The intellectual level of a seminar or workshop may be correctly assessed by specialists. However, to establish whether the operation was a success or a failure, it is insufficient merely to ascertain, through an ex post evaluation, whether the trainee became qualified, whether the professional became retrained and whether the beneficiary State was satisfied. It is necessary to ascertain whether study groups were held, whether technologies were transferred and whether the instructors trained the teams needed to use them. Such effects can be determined only a fairly long time after any evaluation carried out which, since it is always carried out too early, will prove rather lacking in usefulness.

55. In situations and circumstances in which greater importance is attached to quantities and measurements than to the learning of lessons of very wide application or slow catalytic effects, the difficulties of evaluation, despite appearances, will be just as serious. It may seem relatively easy to evaluate the results of a distribution of medicines or of a vaccination campaign. However, a rate of distribution does not necessarily imply a particular rate of use, or even any use at all. A measurement of vaccinations is complicated by the calculation of the number of injections and, in some cases, there is good reason to hesitate between a laborious survey, which is disproportionate to the results sought, and a resort to sampling, with the risks of false extrapolation that this entails. By the same token, it will be easier to quantify the increase in production of a dwarf rice variety than to measure the impact of such an increase on the feed of livestock accustomed to non-dwarf varieties.

Impact co-ordination and competition

56. The logic of project-specific evaluation implies that greater interest should be shown in the particularity of the processes, output and effects of

the project than in its relationship with other projects. This logic leads to consideration of the intrinsic concatenation and the specific results rather than the environment and the consequences. Even an in-depth evaluation yields only incomplete information on a possibly significant aspect of a national economic situation. From the standpoint of the evaluation's logic, the fact that it is extremely difficult to dissociate the effects of one project from those of another has to be ignored. Being restricted to a single project, the evaluation cannot prevent questions of external effects and influences from assuming the same importance as questions of consequences and results.

57. The imputation of a specific impact to a given project must necessarily be conjectural. The output of any project will always be in a situation of interchangeability and interaction with another. An FAO agricultural project designed to improve the nutrition of a target group or a UNESCO literacy project may have more significant induced effects on the health of a population than a WHO primary health care programme undertaken jointly with, or independently of, UNICEF.

58. A report recently published by the Asian Development Bank recommends that the demographic aspect and impact of the economic projects to be financed should never be disregarded, whether the projects relate to family planning, health, nutrition, the exodus from rural areas or migration to medium-sized towns rather than to large cities. By interlinking these questions, with which each specialized agency of the United Nations will have to deal from the standpoint of its own expertise, the Asian Development Bank highlights the indissociable nature of the reciprocal effects of projects within a network of interrelationships that cannot be properly appreciated on the basis of evaluations which, however thorough they may be, are always compartmentalized and fragmentary.

59. Where it is proposed to carry out a retrospective evaluation of improvements in the health of children in a particular beneficiary State, to which more or less complementary but highly diverse United Nations projects have contributed, it will be almost impossible to determine what is attributable to a sanitation project, to a drinking water supply project or to the adoption of a technology conducive to more rational nutrition.

60. UNICEF experts who attended a meeting held in 1986, concluded that the problem was not merely to determine whether a decline in infant mortality and morbidity was due to the execution of a drinking water supply programme, whose beneficial effects would in fact extend to the entire population, including adults. The problem lay in the need to recognize that the direct impact of the project on the target group could not be assessed and that, in the project as well as in its effects, it was impossible to distinguish between what was attributable to the beneficiary Government, to bilateral aid, to NGOs or to the United Nations executing agencies.

61. Opinions will differ as to the extent to which certain bodies have contributed to the achievement of an identical objective by pursuing related but dissimilar aims in a number of projects. These differences may influence any evaluations that may be made, possibly at different times, and that will inevitably vary in their assessment of an impact viewed from the standpoint of a particular project. In addition to the interrelationship between impacts and the impossibility of breaking down the effects of related projects, competition between the donors of bilateral or multilateral aid and the rivalry or relative lack of co-ordination among United Nations executing agencies will make the evaluation increasingly less operative and more disappointing.

62. Although perfectly consistent from the standpoint of its design, the conversion of contributions into output and the correlation between objectives and anticipated effects, a project undertaken by a specialized agency may counteract, without being inconsistent with, one that is being carried out by another agency in a related field. For example, a UNICEF project might emphasize the health and humanitarian aspects while paying little attention to the demographic effects that it entails. At the same time and in the same region, UNESCO might be putting into operation an education programme to encourage family planning and restrict fertility. Although the two projects are not contradictory, they might have the effect of reducing, without actually neutralizing, each other's effects. The evaluation of each of them will not be very meaningful, since the combined effect of the forces that they bring into play is more significant than the outcome of either individual project. The initial lack of co-ordination of the projects will lead to an inevitable lack of co-ordination of the evaluations.

The need for a country-specific approach

63. By functioning in accordance with the logic of the project, regardless of the latter's degree of consistency and the extent of its integration with a programme, a subprogramme or a group of projects, the evaluation system cannot free itself of sectoral and partial constraints. It does not allow an overall view of the medium-term effects of multilateral aid granted to a particular country, for a given period, by the United Nations, notwithstanding the extreme importance of the lessons that could be learned from such observation and the fact that only consolidated results can be deemed significant.

64. It is conceivable that some might wish to conform to the project logic and the evaluation practice, with which it is consubstantial, with the sole aim of avoiding the need for accurate measurement of the overall impact of multilateral aid and sparing themselves the effort of self-criticism. They may be afraid of having to recognize a lack of co-ordination or a dissipation of effort where they would like to record the multiplying combinations of a synergy. With a project-specific approach to evaluation, it is possible to describe the situation in all its infinitely varied aspects, to draw balanced conclusions and to make substantial allowance for the vague and imprecise, whereas a country-specific approach offers the advantage and/or the disadvantage of dispelling illusions and portraying the situation in its true light before continuing, intensifying or abandoning the endeavour.

65. It is at the country level and in the light of the sum total of the country programmes undertaken by all the United Nations agencies that an in-depth evaluation is likely to provide the information most needed for the development of the beneficiary States. By focusing, not on the operations of a single specialized agency, but rather on the juxtaposed or superimposed activities of a number of them, which are sometimes concordant but may also duplicate and occasionally run counter to each other, such an evaluation would reveal both the deficiencies and excesses. In-depth evaluations (which, being the most costly, should be the most productive) are worthwhile only if a country-specific approach is adopted.

66. Although most tripartite studies and in-depth evaluations are project-related, UNDP also carries out evaluations of other types, relating to processes, institutions, themes and even country programmes. In the last-mentioned case, according to its Procedures Manual, the evaluation is entrusted primarily to external consultants with a view to drawing conclusions applicable to the implementation of programmes in other countries.

67. While efforts are made in this way to transcend the project logic, they remain limited in scope. Country programme evaluations have so far been concerned mainly with a very small number of LDCs in which medium-size activities and less complex operations are undertaken. This was so in the case of Nepal in 1987. Moreover, these evaluations are subject to approval by the Governments of the countries concerned, which share with UNDP the responsibility of the exercise. Finally, however decisive its role in the financing of United Nations projects may be, UNDP does not cover all the multilateral aid operations, some of which fall within the specific planning sphere of one of the executing agencies.

68. Under the country-specific approach, the in-depth evaluation would cover all the multilateral aid operations undertaken in the country during a given period by all the United Nations agencies. It would not be confined to the cases dealt with by UNDP, in which a beneficiary country may modify its development strategy and priorities to a considerable extent, decide that there is a major defect in the very design of the programme or encounter too many difficulties in securing the desired results. It should be periodic and should help to gather the information needed for establishment of a balance sheet. While it is not, of course, for States to carry out overall evaluations of their national performance, it is consistent with the mandate of the United Nations to make overall evaluations of the multilateral aid provided by its various bodies. This would not be a new departure, since it is already done in the case of bilateral aid and the World Bank.

69. An evaluation carried out on the basis of a country-specific approach would have the double advantage of ensuring a better understanding of the interrelationships among United Nations projects and programmes and a better appreciation of any competition and of the relationship with the donor Governments and NGOs providing bilateral aid. For purposes of international comparison, it would provide useful information and contribute to a better understanding of the causes of failure and success in a given regional zone.

70. With an evaluation thus designed it will be possible to obtain even more information by adapting the methodology according to whether the countries are insular or land-locked and whether they fall into the category of LDCs or NICs or that of continental-size countries or nations with a very small territorial area, classifications which are not mutually exclusive. In the continental-size countries receiving substantial United Nations aid and having a high IPF, the evaluation could focus on specific provinces or territorial entities. Elsewhere it could combine a country-specific approach with an approach by geographical region.

71. It is widely believed that "heavy" evaluations should be viewed with a certain amount of suspicion or indifference. Once a project has been completed, the specialized agencies often feel that they would do better to move on to other work. If they have not sought, or known how, to provide the necessary explanations to the beneficiary States, the latter will tend to mistrust what, several years later, may be viewed as intolerable interference in their internal affairs. UNDP itself also sometimes balks at the cost of some evaluations when it is equivalent to that of a major project.

72. It is a fact that the only country-approach evaluation carried out by JIU, for Sri Lanka, kept five inspectors and four research assistants occupied full-time for a year. It is also a fact that the Economic and Social Council undertook a number of country evaluations in 1969, that the World Bank conducted one in Colombia at the same time, that the experiment was repeated by UNDP in Nigeria in 1973 and that all of them found the exercise to be so

complex and costly that they refused to repeat it. Nevertheless, in 1987, UNDP again saw fit to undertake an evaluation of projects in Nepal using the country approach.

73. While the cost of impact evaluations on the basis of the country approach deserves consideration, the argument that the current financial crisis of the United Nations precludes such evaluations can easily be turned against those using it. Admittedly, the use of this type of evaluation initially results in increased expenditures under a budget item which, moreover, would have to be created. However, a financial crisis at least has the advantage of inducing the officials concerned to question the marginal effectiveness and relative usefulness of established structures and procedures and encourages or obliges them to effect, within an unchanged budgetary framework, reforms which they would never have dared to undertake or have considered undertaking beforehand. Once the principle is established that the system must remain frozen, that any redeployment of staff is out of the question and that any transfer from one item of expenditure to another is inconceivable, then any readjustment of evaluation procedures necessarily seems costly, since it does have financial implications, whereas inaction has none, there being no accounting procedure to isolate and calculate its actual cost.

74. The need to conduct a given type of evaluation cannot be dependent on the degree of facility with which the financial problem that it presents can be solved. The cost of in-depth mid-term evaluations is incorporated in its budget envelope. The cost of self-evaluations is included in the general overheads of executing agencies without any itemized accounting being done to show its actual magnitude. This partly explains why it is used fairly frequently, but is not enough to justify it as a practice. Conversely, the main reason for the infrequency of impact evaluations, in spite of their efficiency, must surely be that project documents and operating budgets do not allow for or furnish the means of financing them.

75. Although they appear to be the most costly, impact evaluations could prove to be the most cost-efficient. A UNDP report has pointed out that, by allowing an annual decision as to whether half a dozen large-scale projects should be terminated or restructured, the evaluation system justifies its cost, however high it may be. Moreover, at an inter-agency meeting held at Geneva in February 1988 at the initiative of the UNDP Evaluation Service, the representative of the World Bank laid considerable stress on the fact that no development agency could do without impact evaluations either for technical co-operation or for investment projects. A consistent and global impact evaluation based on a country-specific approach favours more economical project management from the standpoint of both yield and development.

76. There remains the question whether the argument that impact evaluations based on the country-specific approach are too complex can be accepted. While such evaluations are certainly difficult, this fact cannot be used to infer that they are unnecessary and, in particular, since those who have experimented with them are few and far between, it is not easy to find officials who can question the need for them on the basis of sound experience.

77. A number of obvious facts are inescapable. United Nations agencies, in executing projects, must necessarily consider viability and follow-up. Where the areas of competence of executing agencies adjoin or overlap, it is better to determine the interaction and combinations of effects between projects, than to identify the specific impact of each individual project. The involvement of several executing agencies in one evaluation process is bound

to cause rivalry and conflict. However, the all too frequently invoked, but by no means insurmountable, complexity factor must not be identified with fear of or refusal to face complications.

III. THE IMPROVEMENT OF EVALUATION PROCEDURES

78. The evaluation units of the United Nations agencies follow procedures and use methodologies that conform to guidelines and refer more or less implicitly to the UNDP manual and its chapter on monitoring and evaluation, the most recent version of which was issued in April 1987. This chapter, which concludes rather than initiates a series of reforms and is in the nature of an assessment, a source of guidance and a test of orthodoxy, represents the outcome of three years of work, and the provisions it contains are regarded as having been tested in the field, in co-operation with UNDP resident representatives, executing agencies and recipient Governments.

The time-lag between manuals and theory

79. The system developed is meant to be both comprehensive and manageable. The attention paid to results and the considerable care taken to establish effective retroaction and feedback machinery govern procedures that are valid not only for UNDP, but for the projects and programmes carried out by the executing agencies which UNDP is required to finance and support. The basic principle is therefore to make the evaluation exercise more cost-effective by ensuring a more coherent integration of the information it produces, in order to improve the design and execution of subsequent projects.

80. Whatever the intentions, the delicate negotiations associated with the preparation of the new chapter of the UNDP manual were unable to bring about a very substantial change in the system. It was tidied up to advantage, but the procedures selected in the main continue to reflect the concept of the almost mechanical conversion of inputs into output, as well as the analytical and sectoral design of individual project evaluation that it implies. The system continues to focus on achieving uniformity in respect of rules and on leaving the context aside. It has greater faith in rational terms of reference and in the concept of objectivity than in the quality of the evaluator and timeliness requirements.

81. Although the UNDP manual enjoys some pre-eminence and serves as a reference work, it does not eclipse the guidelines intended for the executing agencies and continues to coexist with them. To be sure, the efforts made to achieve uniformity in these documents have been maintained in recent years. The preliminary understanding on the basic evaluation vocabulary is a noteworthy achievement. It is also encouraging that the United Nations agencies as a whole have adopted the definitions developed by JIU and adopted by UNDP.

82. Certain discrepancies among the manuals still persist, however, due less to the sectoral specificity of the areas covered by the executing agencies than to differing views on the relative importance of the phases and methods of evaluation and monitoring. The diversity of administrative structures and of the positions adopted by central evaluation units also have an impact and tend to colour the general picture. But while the standardization exercise may be continued to advantage in defining terms and clarifying concepts, the harmonization of the rules of procedure seems quite advanced and does not require too much attention.

83. Although the latest version of the UNDP manual sets out the procedures which came into effect on 1 July 1987, it would seem premature and hazardous to envisage their review at the present time and to embark on their revision. It is most regrettable, however, that the new chapter fails to take sufficient account of the research conducted in the meantime by the most innovative evaluation theorists.

84. These theorists have, to some extent, taken a fresh look at the subject, and their conclusions tend to favour a country approach and an upgrading of the socio-political context of the evaluation exercise rather than a project approach and the reduction of the socio-political context to a fairly closely knit combination of assumptions and external factors.

85. In point of fact, these experts tend to regard evaluation as a socio-political process associated with a geographical framework and involving local values, target groups and partners inseparable from the context in which they evolve. They consider any generalization of the lessons drawn from the exercise to be dangerous and the detection of differences to be more important than any significance attributed to common features. As they believe that socio-cultural factors carry as much weight as technical factors, they would like them to be treated as inputs rather than bugbears. They would like the evaluator to be not merely a technician, narrator and judge, but a pedagogue and a mediator, a designer and an agent of change capable of humility and patience, open-mindedness and adaptability.

86. Showing less confidence in the accuracy and extension of the terms of reference than in the interpretation relationship to be established between the evaluating officer and the partners whose activities are being assessed, these theorists attach prime importance to the concepts of co-operation and negotiation. The human factor and the dialogue with the user or beneficiary are therefore more important than the invariably relative rationality of rules and procedures.

87. While not speaking, like these over-enthusiastic theorists, of a fourth-generation evaluation, and while not believing that the UNDP manuals are lagging behind by one type of evaluation, it is nevertheless considered worthwhile to extract from this research the elements that can be assimilated and that are likely to improve the guidelines currently in force.

The timeliness of procedures

88. The reliability of an evaluation should, in principle, be determined initially by the precision of the terms of reference used and not by the personality of the evaluating officer. Current experience shows, however, that an evaluation's value and the approach adopted also vary with the evaluating officer, his independence, the place he occupies in the hierarchy and his position inside or outside the agency concerned as much as with the validity of his terms of reference.

89. Needless to say, excessive subjectivity and a defective methodology can only weaken the conclusions of an evaluation. The UNDP and executing agency manuals, like previous JIU reports, rightly stress the objectivity of this exercise and revert repeatedly to the need for it. However, all appear to have paid insufficient attention to the need for timeliness, which is no less imperative. Objectivity is necessary without being sufficient, and the same can be said of timeliness.

90. An objective evaluation must above all be carried out at the right time, which will vary, depending on the types of project evaluated. If it is to prove useful, it has to be carried out fairly soon after the completion of operations, depending on whether the effect anticipated is more precise or more diffuse, simple or complex, and on whether in the medium- or long-term it fits decisively into an overall policy or not. Timeliness should also depend on the dominant functional characteristic of the project, depending on whether it consists of the creation or consolidation of an institution, an experiment or a pilot project, a training project, emergency assistance or logistical support.

91. The evaluation should also take into account the time-frame and the nature of the project. The time-frame in agriculture is not the same as that in industry, and a vaccination campaign cannot be evaluated in the same way as a teaching programme. The interdisciplinary and polyvalent elements in an integrated rural development project or a cultural campaign presuppose a much less rapid impact than the introduction of a more profitable variety of seed or the provision of latrines in a specific sector. The evaluation time-frame used for a project designed to achieve an overall reduction in domestic energy consumption will not be applied to a research and development project on energy saving which, in the long-term, may be very cost-effective and promote the adoption of new technologies, but which in the short-term has only a slight impact on the consumer and his behaviour.

92. Without imposing an unduly rigid time-frame for the in-depth evaluation, it seems that the decision to undertake such an exercise should take greater account of the period during which it can be conducted most effectively. The happy medium between the too early and the too late is elusive and no manual could indicate limits and thresholds. It would, however, be useful to explore the problem of differences in evaluation times in relation to the economic sectors covered and dominant functional characteristics.

93. The work performed ought not to lead to a report which soon becomes outdated owing to the fragility of its conclusions or to an intellectual exercise whose pretentiousness is equalled only by its lack of practical interest. It would therefore be advisable for UNDP and executing agency manuals to indicate more clearly, in the section describing their use, time-frames and margins so as to ensure the better use of time in respect of various types of project.

94. The evaluation must be based on a clear understanding not only of its purpose but also of the nature of the body for which it is being carried out. Skill in communicating a message is as important as the quality of the expertise. However excellent the design and conduct of the evaluation and however rational and unexceptionable the lessons drawn from it, users may not necessarily be capable of assimilating its conclusions or consider it expedient to make use of them. The usefulness of the information provided and the scope of the recommendations presented will depend less on the nature and value of the methodology employed than on the recipient's socio-cultural context which must be capable of digesting them and using them for decision-making purposes.

95. Despite the relatively precise detail of the terms of reference, evaluating officers are not always able to grasp what the users want. It is not enough for them to understand the imperatives that motivate agencies and govern the decision-making process. They have to be understood from the standpoint of the users - even if the evaluating officer considers that the latter do not understand them properly. Decisions are taken in a specific

socio-political context and not elsewhere. The interactive process that has to be set in motion requires sensitivity training and a knowledge of group dynamics that are only too often lacking in evaluations.

96. The evaluation is no more required to present the complete truth than to tell obvious lies. It must be as objective as it is timely, and for that reason must refrain from mentioning certain things. It must measure out what is communicable and bearable. It must sense the wishes of the users that are not implicitly stated in the terms of reference, and order its conclusions in terms of what will be easy, or relatively easy, to incorporate in the decision to be taken.

97. However objective and experienced he may be, the evaluating officer must resist the temptation of offering himself an intellectual treat. An evaluation is not a dissertation. The user must, after giving his consent, be convinced of the usefulness of the evaluation, even though the price to be paid for his conviction is the sophistication of the reasoning or the perfection of the evaluating officer's study. The publication of an evaluation report does not constitute an automatic and necessarily effective transfer of a certain amount of knowledge and experience. No mechanism can guarantee that a lesson, however relevant, will be understood or applied, but at least some thought can and should be given to the manner in which the user receives it and to the ways of making it assimilable and accessible.

The adjustment of the various types of evaluation

98. If, in point of fact, the United Nations attaches to evaluation the importance that, in the opinion of all those concerned it should be accorded, it would be advisable to strengthen the basis and improve co-ordination of procedures so as to allocate tasks in a way which will ensure a more productive use of results. We do not intend, nor are we able, to provide detailed solutions which would require the assistance of a large number of experts, an appropriate forum for whose studies could be created by institutionalizing the Inter-Agency Working Group on evaluation (IAWG). Nevertheless, it is our intention, without going beyond the scope of this report, to identify the general direction and to indicate the course and approach to be followed in reorganizing the types of evaluation.

99. The administration of evaluation at the higher and the lower levels, within and outside headquarters, in the executing agencies, among the donor Governments and in UNDP involves so many partners that one may legitimately question the ultimate usefulness of an exercise that is too often directed at an undifferentiated audience because it fails to distinguish properly between user and recipient, between day-to-day administration and long-term policy, and between the beneficiary State and the target group. The desire to communicate too much with too many partners could well have the result of not getting through effectively to any of them.

100. UNDP is already aware of this problem, and neither underestimates its difficulties nor fails to recognize its extent, but for the moment it is leaving unanswered a question that it has simply raised (DP/1987/27). Its updated and revised manual specifies the procedures for formulating and disseminating information obtained by evaluation. However, what still has to be done is to improve ways and means of identifying and consolidating the results thought to be of the greatest use, so as to apply them most profitably in accordance with the level of responsibility and the type of beneficiary.

101. Without claiming to provide a solution, it is nevertheless possible to suggest certain guidelines. In the recipient countries, where projects are carried out, what is needed above all by national co-ordinators, UNDP resident representatives, experts and executing agencies, is immediately usable and up-to-the-minute information that can best be obtained from first-echelon monitoring. It is in this context that the PPER and the Tripartite Review should continue to be the hard core of grass roots verification which, combining self-evaluation and monitoring, enables report conclusions to be closely linked with the decisions to be taken, and the results of field reports with the corrections to be made in the course of project execution.

102. At programme level, self-evaluation in itself can only be beneficial. At project level, its reliability and usefulness will depend primarily on increased use of the PPER and even more on the reorientation of the work of Tripartite Reviews, the agendas of which should in future include questions of project viability and sustainability, as suggested by the representative of FAO at the IAWG session held in Geneva in February 1988.

103. Some clarification is also needed to avoid unfortunate confusion between reporting, investigation and self-evaluation. In many instances, the term "self-evaluation" is still used to describe what is simply reporting or a routine report from subordinate to superior. Somewhat differently, but as part of a similar approach, the term "self-evaluation" is applied to something which, even in the opinion of its users, is merely a participatory procedure. In UNHCR, for example, responsibility for evaluation, thus defined, will devolve on the bodies with which an agreement has been reached, and "whenever possible" on the refugees themselves.

104. Very appreciable changes should also be made to the scope of in-depth evaluations. Firstly, the usefulness of mid-term in-depth evaluations for projects which are complex and/or cost more than one million dollars appears highly questionable. Since their financing is included in the project's budget envelope and they are not excessively costly, these evaluations can be used for mid-term assessment of the relevance, efficiency and efficacy of a project. However, they ignore questions of viability and sustainability which would go much further towards justifying their use. If the questions of viability and sustainability can be dealt with at the project formulation stage and be subject to a continuing process of review within the framework of PPER and Tripartite Reviews, the mid-term evaluation exercise could be done away with without any major loss.

105. Ex-post in-depth evaluations, the funding for which is not included in project budgets, may be of some use to the extent that they provide conclusions which can be used in the preparation of similar projects or in programme implementation. However, the experience of executing agencies shows that their rather high cost is not always justified by the practical use that can be made of them or by the anticipated feedback benefits. Provided that they do not resemble simple archeological reconstructions, ex-post evaluations can be useful for programmes of a certain size rather than for projects. In the case of projects, such evaluations should continue to be used selectively and to a limited extent.

106. Ideally, therefore, the scope of in-depth evaluations should include the thematic or institutional evaluations which are already in use. However, the main focus should be on country-specific or geographical-region impact evaluations which are concerned with project follow-up as well as execution, and can address the questions of cost-effectiveness and sustainability. A review of the cost-effectiveness and sustainability of a set of projects would

actually be the best way of giving meaning to the unnecessarily partial concepts of relevance, efficacy and efficiency which are applied to individual projects, but the assessment of which can provide only a partial or distorted view of United Nations technical co-operation activities.

107. In these conditions, evaluation in accordance with the country approach will make it much easier than evaluation in accordance with the logic of the project to grasp the essentials with a view to the study or review of multilateral assistance machinery and strategies. It will at the same time facilitate the transmission of the most important diagnoses and specific recommendations to the recipient Governments. Better, more independent and more reliable, even if it seems less diplomatic and more carping, this kind of evaluation will not merely suggest the decision, but also prepare the ground for it and, above all, adapt the message to the users, namely, Member States, most of which are also recipient countries.

Modifications to guidelines

108. The process of modifying rather than recasting United Nations manuals on procedure consequently implies the incorporation of certain points and the elaboration of others which, while not affecting the drafting or the substance of the chapter on evaluation and monitoring, would make its use more productive. Omissions to which UNDP officials have themselves drawn attention (DP/1987/27) would also be made good.

109. The guidelines could be improved by the incorporation of more instructions which would assist evaluating officers in their work. The auditing manuals used by a number of national auditing agencies cover the most important and/or frequent irregularities, as well as the most common types of errors and omissions. An annex to executing agency and UNDP manuals could contain check-lists of the most obvious difficulties and constraints which are encountered in project delivery and which must be assessed during the evaluation.

110. UNICEF recently prepared for its field staff a list of the defects and shortcomings observed most frequently in the area of training activities. This idea could well be adopted on a wide scale, and lists corresponding to the specific needs of various agencies compiled. They would indicate functions, objectives and external factors that should, in any event, be spelt out in greater detail. Similarly, manuals should contain lists of seasonal constraints (cyclones, monsoons, etc.) affecting the implementation of projects and their timetables, as well as a series of progress indicators which would facilitate the monitoring of successes and failures attributable to project components. These progress indicators would vary with fields of application and the types of activities with which the executing agencies have to deal.

111. In order to obtain improved information and a better grasp of the disparities and distortions likely to restrict the scope of the evaluation, were they to be overlooked, the manuals should also refer to certain analytical criteria which should be listed for information purposes. For example, the overall results of a health project would be rather insignificant if the evaluation neglected religious, ethnic, occupational or regional differences, income levels, types of habitat and natural resources endowment.

112. Ways and means should also be studied of achieving a better adaptation of the manuals to the nature, size and level of development of the countries where the evaluation is carried out. It seems unlikely that the same criteria

of efficiency are valid for the least developed countries and the newly industrialized countries, for continental-size countries, island nations and land-locked countries, and for the projects carried out by United Nations executing agencies or by recipient Governments acting as executing agencies.

113. Lastly, the manuals should, in an initial section containing instructions concerning their use, emphasize the need for greater punctuality. An inordinate desire to achieve perfection or an obsession with thoroughness should not hamper the success of the evaluation. Refinements in the planning and drafting of a study should be avoided if they are likely to have the effect of delaying its preparation. A less perfect report produced on time is more important than a highly sophisticated evaluation not completed within the prescribed period. The many reports whose belated preparation has affected the evaluation which they are supposed to enhance serve to emphasize the importance of observing time-frames and deadlines.