

PN-AAM-194

ISN=15613

DISSENT, OBJECTIVITY AND ETHICS  
IN PROGRAM EVALUATION

Prepared for the  
Office of Evaluation  
Bureau of Program and Policy Coordination  
Agency for International Development

Robert L. Hubbell

May 1979

DISSENT, OBJECTIVITY AND ETHICS  
IN PROGRAM EVALUATION

TABLE OF CONTENTS

I.	Statement of Problem	1
II.	Review of Literature	
	Evaluation vs. Research	5
	Judgments	9
	Influence of Values	10
	Dissent, Adversary View, Alternative Perspectives	12
	Effect of Participation	15
	Independence of Evaluators	17
	Purposes of Evaluation	20
	Scopes of Work, Monitoring Contracts	23
	Reports	26
	Publication	27
III.	Codes of Ethics of Professional Associations	
	Treatment of People	31
	Confidentiality	31
	Host Governments and Secrecy	32
	Publication in General	33
	Objectivity	34
IV.	Proposed Code of Ethics in Evaluation	
	Values in Evaluation	37
	Work on Standards	37
	Mutual Responsibilities	38
	Scope of Work - Purpose of Evaluation	38
	Contract	39
	Collecting Data	40
	Reports	41
	Publication and Dissemination	41
	Re-Analysis	42
	Share Basic Knowledge	42

## Table of Contents (continued)

V.	Information from Interviews	
	Dissent	43
	Use of Contractors	48
	Composition of Teams	50
	Techniques for Objectivity	52
	Review of Evaluation	54
	Scope of Work	55
	Training of A.I.D. Personnel	61
	Contract Monitoring	62
VI.	Analysis of Issues and Recommendations	
	Aspects of Concern	64
	Approach to Recommendations	67
	Better Understanding of Evaluation	68
	Collaborative Evaluations and Formal Reviews	71
	Scopes of Work	75
	Resources for Evaluation	78
	Benefits for People	80
APPENDICES		
	A. Scope of Work	88
	B. Annotated Bibliography	90
	C. Code of Ethics	124
	D. A Draft Statement of Ethical Respons bilities	138
	E. Instructions on Scope of Work	142
	F. Development Support Bureau Evaluatio Guidance	145
	G. Legislation on Evaluation	150
	H. People Interviewed	154
	I. Recommendations	155

I. STATEMENT OF PROBLEM

Some time ago, an office in A.I.D./W commissioned an evaluation of a project which involved special training in the United States at several institutions and also promotion of similar training overseas. A non-profit association was hired to do the job. It in turn recruited two specialists in the field. It is not known whether they were also experienced in program evaluation procedures. The evaluators were given about three weeks for the job.

One of the evaluators wanted to look at aspects of the project which the A.I.D. monitor considered beyond the scope of work. He was also more critical of parts of the project than the other evaluator (they did some of the work together and some alone). In a debriefing session, each submitted reports which were seen by several A.I.D. officials. However, the evaluators were unable to arrange time from their regular work to prepare a unified report. An official report was finally prepared by the A.I.D. monitor, association officer, and less-critical (and geographically closer) evaluator. It contained the major points of the critical evaluator but without detailed backup. It was duplicated by the association.

After some delay, the critical evaluator obtained a copy of the official report and then filed a complaint with the A.I.D. Administrator about irregularities. At the Administrator's request, the staff of the Auditor General investigated and concluded that the unified report did include the major critical points, and that the A.I.D. monitor did not violate Agency policies. At the same time, the Auditor General recommended and the Administrator agreed that A.I.D. should review its

procedures to determine whether changes would be advisable.

This is the genesis for the current assignment. Several issues were involved in the instance described. Did part of the confusion arise from a vague scope of work? Should the work of the evaluators have been more closely monitored by the headquarters of the nonprofit association? Did the A.I.D. project officer play too active a part?

Such questions imply that dissent is undesirable and could be avoided by better procedures. On the other hand, presentation of differing viewpoints may be a way of getting at the true situation. Should a scope of work actually encourage the members of an evaluation team to state their individual interpretations of data in the body of their report? Would this alert administrators to possible problems so that they would make more balanced decisions -- or would it unduly complicate the decision-making process and tend to encourage less attention to evaluation findings and recommendations?

Even if the Agency does not encourage differences of interpretation, should it ensure that dissents are published in accord with the traditions and ethics of scholarship? Or does the convenience of management override the scholarly imperative?

The commissioners of this review of evaluation procedures used the opportunity to request a consideration of issues which are related to but broader than the narrow problems of how to handle dissent. They were seeking ideas on how to get the broadest, most objective presentation for decision makers and then how to encourage decision makers to pay more attention to evaluation findings and recommendations.

Thus, the quality of evaluations may suffer from various kinds of pressures for consensus. These may be exerted by a team leader, who

seeks a unanimous report because it will be more convincing and because he so interprets his responsibility as leader. Or the pressures may come from project managers and their supervisors (either host government or donor) who want evaluations to endorse their work. Evaluators who are staff members may be influenced by organizational loyalties or ambition. Evaluators who are outsiders may submit acceptable reports in the hope of subsequent contracts.

More likely than unethical pressures are the influences of professional and ideological biases. Engineers may have an edifice complex. Social anthropologists are concerned about relationships. Economists look for favorable cost/benefit ratios. How can these various perspectives be filtered or reconciled to reach the best decision?

The problem of suppression may occur in the final stage of an evaluation. After the data have been analyzed and findings made (the stage of staffwork, whether by permanent staff or consultants) the decision stage of evaluation is reached. Policy officials may disregard or reject recommendations. Should the Agency have a process for reviewing decisions about evaluation recommendations, just as it responds in writing to audit findings?

Reflecting such considerations, the scope of work calls for a report analyzing the issues involved in promoting objectivity and candor in evaluation, in defining the ethics of dissent, and in protecting the rights of individual researchers and firms or institutions participating in evaluation research. Also to be discussed are measures needed to improve and protect the ability of the government to make effective use of evaluation results and the appropriate disposition of information gained on the basis of confidential relationships. (See Appendix A for text of the Scope of Work).

This scope might be interpreted so broadly that the contractor would discuss all methods for ensuring objectivity, including techniques for social science research. Similarly he might explore ramifications of assuring that administrators use evaluation finds or disseminate those findings to interested people.

Instead, guided by the Office of Evaluation in the Bureau of Program and Policy Coordination and mindful of the realities of how evaluation is conducted in A.I.D., the contractor has attempted to steer a reasonable course between the overly narrow problem of dissent and the unmanageably wide topics of objectivity and utilization.

The order of presentation in the report follows that of the scope of work: a review of literature on the subject (Chapter II), codes of ethic of key professional societies and consultant organizations (Chapters III and IV), information from interviews (Chapter V), analysis of issues and recommendations (Chapter VI). For convenience, recommendations are reproduced in Appendix I at the end of the report.

## II. REVIEW OF LITERATURE

Selected recent writings about evaluation which discussed aspects of the problem of objectivity and ethics in evaluation were reviewed. In many cases, these authors cited earlier writings, so that a fairly thorough overview of different authorities and opinions could be obtained without compiling an extremely long bibliography. (See Appendix B for an Annotated Bibliography)

The following sections synthesize the views of various writers on ten selected topics which seemed most important for this inquiry. For readability, titles and dates are omitted. The reader who wants such details can refer to the Appendix B.

### Evaluation vs. Research

The preceding chapter on "Statement of Problem" mentioned the traditions and ethics of scholarship which may require publications of dissent. It also suggested that the convenience of management might override such an obligation. Here we have a difference in approach which is crucial to the whole question of dissent, objectivity and ethics; namely, the different approaches of the researcher and the evaluator. Since many of those now earning their livelihood as evaluators began as researchers and since many of those who write about evaluation have a more scholarly bent than other practitioners, the literature pays a good deal of attention to these approaches.

This apparently esoteric, philosophical discussion is of some concern to the Agency for International Development for several reasons. First, it must decide what it wants from evaluation so that it can meet its responsibilities more effectively -- responsibilities for de-

velopment and for proper use of the taxpayers' money. At the same time that A.I.D. shapes evaluation to its own purposes, it must also follow accepted standards of evaluation well enough to be credible with the Congress and the public. Finally, A.I.D. must maintain a reputation of fairness so that it can attract good evaluators. Therefore, a summary of differing perception about research and evaluation is a good way to begin.

Researchers regard their primary function to be the discovery and publication of knowledge. Their professional ethics require them to adhere strictly to scientific methodology almost without regard to the possible influence such adherence may have on the operation of the program being evaluated. For instance, some researchers would refuse to modify an activity to increase its effectiveness because such a modification would also alter a research design. If their findings are ignored, researchers may feel compelled to publish them in an effort to marshal opinion.

In contrast to this research viewpoint are other writings which point out that evaluation is a process of management. As such, it should be oriented to specific policy or program decisions. Not only is evaluation limited in the questions asked, it is also limited in the time available -- its results serve no purpose if they are not on hand before the decision. Since management operates in a context of political responsibility, evaluation may be only one factor influencing the decision.

Evaluation uses social science methods to help make action decisions more accurate and objective. The social science methods involve establishing criteria for success, collecting evidence about those

and then drawing conclusions. The process differs from usual management practices of basing decisions on trained sensibility or from traditional political methods of marshalling opinion. It takes more time and money than using intuition.

(At this point, we can digress briefly to consider definitions. The Office of Management and Budget defines evaluation as systematic analysis and evaluation research as strict adherence to experimental design. Most authors use the words evaluation and evaluation research as synonymous, recognizing that many research techniques are usable in evaluation.)

Although using social science methods, evaluation takes place in a context which is intrinsically inhospitable to scientific methods. Aspects of time and money and of differing approaches to decisions have been mentioned. In addition, as Freeman points out, the major impetus for rational policy making and program development are political and pragmatic. Thus it is the politician, the planner, and the foundation executive who exercise the effective leadership in the evaluation research field, not the researcher.

The precision of definition needed for good research comes from the policy maker. Rossi and Wright point out that the independent variable of the program goals, and the dependent variables of the targets and criteria for success are set for the evaluator, although he may help clarify them. Illustrative of the setting of criteria is the investigation of the impact of guaranteed annual income on work incentive rather than on family stability or health. Moreover, the research design is often affected by the nature of the problem, which may prevent use of control groups, for example. Thus evaluation is

more judgmental than other research.

Because of these characteristics of evaluation, early writers were apologetic. Researchers had accused them of abandoning their principles. Now, evaluators are beginning to regard evaluation as something different from research, a profession with its own standards. Thus they have coined the term evaluation research and have formed a professional association. Weiss proudly asserts that evaluation research can be worthwhile when there are complex outcomes, when decisions are important and expensive, and when evidence is needed to convince people about the validity of conclusions.

A.I.D.'s approach to evaluation has always been pragmatic. For evaluation of single projects, A.I.D. decided that the responsibility for the majority of evaluations would rest with the field mission or A.I.D./W office managing the projects. (To derive guidance for future program policy, comparative evaluations are directed from headquarters.)

Decentralization of evaluations has both advantages and disadvantages. On the positive side are coverage of more projects, greater knowledge about projects, and easier use of findings to modify project design or execution. On the negative side are possible losses in objectivity and less informative evaluation reports.

A.I.D. has adopted several procedures to help overcome the threats to objectivity posed by devolution of evaluation responsibility. One is promotion of the logical framework approach to project design, so that criteria for evaluation (targets and indicators) are agreed in advance and arrangements made to collect data as the project proceeds. Another is the policy for collaborative project design and later collaborative evaluation, so that the different perspectives of host and donor are taken into account. A third is the use of outside consultants for

data collection and analysis, or objectivity to its own efforts.

It was an instance in the use of outsiders that led to this assignment and it is most likely that questions about dissent will arise when there is a team of outsiders. Internally, there may be suppression rather than dissent. But use of outsiders does not in itself assure objectivity, as review of subsequent topics from literature will show. Outsiders may bring their own biases or may offer "expert" opinion instead of solid evidence.

Thus A.I.D. has a continuing tension between the practicalities of management and the desirability of objectivity. The absolutes of research and unrealistic, but constant attention is needed to the application of procedures such as those cited above.

#### Judgments

The intrinsic meaning of evaluation is that something is judged -- that a value is put on it. Yet some argue that evaluators should confine themselves to description, leaving the interpretation or judgment to decision makers. It is thought that in this fashion evaluators can remain scientists. The difficulty with such a holier-than-thou attitude is that evaluators would not be discharging their responsibility to help the decision maker, who may lack the skill and certainly lacks the time to interpret data and reach conclusions. Moreover, the problem of value biases still is not obviated by refraining from judgment, as is discussed below under the heading of "Influence of Values."

Weiss argues that evaluators owe the organization that funds their work not only objectivity but as much usefulness as they can devise. But when the evaluators make recommendations, they should state which are supported by data and which draw on knowledge and values from outside the study.

A.I.D. has a mixed record concerning evaluation judgments, which come at two stages in the evaluation process. First, judgments come at the staff work stage of analyzing data, sorting out issues, and devising recommendations. Next is the review of staff work by decisions and accepting, modifying, or rejecting recommendations. Sometimes, when the staff work is done by an outside team, the team is given so little time and so little previously collected data that it has no choice but to offer unsupported opinions. At other times, decision makers have chosen to ignore well-supported findings and recommendations. They may do this because of a broader and valid view of political and social forces or they may simply be stubborn advocates of their own programs.

#### Influence of Values

Many writers point out that social sciences borrowed from physical sciences the myth that science is value-free. Gradually, scholars have come to realize that even physical science is not as value-free as it claims. As for program evaluation, it takes place in a context full of values, both for evaluators and for program managers.

Lynd and Kelman both declare that a controlling factor in any science is the way the problem is stated. This reflects values, either professional or ideological. For instance two people might study poverty because they are against it. A psychologist might look at the motivations of the poor and the sociologist at their access to opportunities.

Anderson and Bell sort values into three categories. One is ideologies. Evaluators may not be aware of their ideologies -- such as the "rightness" of pre-natal care or more education and the "wrongness" of corporal punishment or limited access to museums. E. R. House advocates an ideological approach which seeks a "just" distribution of resources maximizing satisfaction for all, not just the elite. Hudson propounds a "law": "The

greater the ideological relevance of research, the greater the likelihood that the research worker... will pay selective attention to the evidence he collects." Another category of values is personal preferences, which are not unlike ideologies. Finally, professional values are reflected in choices of methods.

Weiss points out that evaluation incorporates a set of political assumptions on the desirability of meeting stated goals and the probability of reaching goals. By accepting experimental variables, evaluation conveys that other elements in the situation are unimportant. Evaluation usually looks at services, not social structure. It is reformist, modifying programs without drastic change. Sjoberg says that the evaluator belongs to a social order and accepts indicators compatible with it.

How can the influence of values be handled? Campbell proposes methodological protections discussed in the next section. Myrdal said that values must be explicitly stated. Weiss hopes that the evaluator will at least distinguish which recommendations are supported by data and which derive from knowledge or values outside the study. In legal terms, the obiter dicta should be kept separate. Anderson and Ball urge that evaluators examine their values, inform evaluation sponsors of values that may influence the planning of the evaluation, and to the extent possible without creating confusion, inform audiences how results reflect a particular approach. But they conclude that there is no convenient means for making values explicit and that the evaluator must generally depend upon a simple accounting of decisions made during the evaluation, the reasons for these decisions and the major alternatives.

One implication for A.I.D. of these observations about the pervasiveness of values is that A.I.D. itself should set many of the controlling

values guiding evaluations, rather than leaving them to outsiders. A number of values, such as emphasis on helping the poor, have been set by the Foreign Assistance Act. Others are contained in A.I.D. policies. Scopes of works should specify areas of concern. At the same time, outsiders may usefully help overcome the limitations Weiss cites of ignoring some variables or of accepting social structure.

#### Dissent, Adversary Views, Alternative Perspectives

In an effort to protect against such biases as may arise from the evaluators' values, the choice of problems to investigate and criteria to judge by, or the relationship of evaluators to program and sponsors, some authors have argued for methods which encourage different viewpoints.

Thus Marcia Guttentag, the late president of the Evaluation Research Society, said that the application to evaluation of methods developed for other problems leads to difficulties in producing useful information. For example, experimental designs and classical statistics can force a set of assumptions on a program or attempt to change the program. Economic analysis stresses input and output but neglects process and the subject. Since the perspectives of various participants differ, research ought to seek a combination of information bearing on all participants' values rather than mere assessment of behavioral changes, she maintained. She was working on what she called a "decision theoretic" model, which she characterized as a simplified form of multi-attribute scaling of utilities using Bayesian statistics. The only example I found was for a program selection rather than evaluation. Guttentag also urged the use of multiple sources of information, e.g. evaluation of employment training by both the Department of Labor and the National Academy of Science.

Diener says that one can work against a possible bias by considering alternative perspectives.

Campbell advocated various methodologies to promote objectivity such as simultaneous replications of social experiments evaluated by two contractors, critical re-analysis of data by another researcher, encouraging internal criticism through minority reports and alternative analyses. He thought arranging for checks and balances could overcome not only biases from values but also cognitive biases of several kinds: tendency to oversimplify the environment and hence get too little information about it; a "perceptual set" so that when a generalization is formed, additional information is perceived in a way to support rather than deny it; and a contrast bias to notice the unusual rather than the central tendency.

Sjoberg points out that there are several kinds of logic or forms of inquiry. Most research tests hypotheses -- a logico-deductive method. Another logic is analogy. This is legal reasoning, with emphasis on precedent. A third logic is dialectic -- weighing and reconciling contradictory arguments. Sjoberg thinks this can be useful for some evaluations. It would mean, in effect, that new categories would be devised. For example, one could explore the idea of community service rather than fines for people guilty of misdemeanors. Another example of "counter-system analysis" would be to balance negative features against positive ones, e.g. school failures against successes.

Several people have advocated the arranging of adversary confrontations in connection with evaluation. This could take several forms. Bandl, a Minnesota legislator, comments that, for a legislator, an evaluator is perceived as another interest group. The evaluator seeks "truth with rigor, explicitness, and replicability." The politician seeks "good". He

believes in the process of the political market place, as in Federalist paper ten, seeing that society poses issues for which there are accommodations, not solutions. If evaluations were complete and definitive, they would replace politicians.

The adversary model proposed by Kourilsky, Levine, Churchman and others had its beginnings in the legal profession. Someone is appointed to "cross-examine" the evidence or to play the role of "devil's advocate." This requires decision makers who are willing to consider both positive and negative evidence. Anderson says it was formerly feared that the better debater could prevail over the evidence but now admits that the adversary method may be best when results are equivocal and much is at stake. It is seldom tried, perhaps because it raises costs and adds complexity.

Another way to get the benefit of differing views is to have one researcher re-analyse the data of another. It was suggested by Campbell. The Government Accounting Office has been doing this for some evaluations of interest to Congress so that it can advise about the credibility or reliability of findings. Re-analysis of groups of evaluations has also been done on various occasions by scholars seeking to learn about the effectiveness of various evaluation techniques. Re-analysis, obviously, does not change findings but may encourage more care in future evaluations.

A.I.D. has recognized that bringing differing points of view to bear may facilitate a balanced decision. Thus its policy is that both project design and evaluation should involve collaboration between host and donor. It has also encouraged formal review sessions to consider the findings of staff or outside evaluators. The encouragement of dissenting views in written reports is discussed in Chapter V, Information from Interviews.

## Effect of Participation

Gurel points out that the characteristics and interactions of the various parties to evaluation affect its outcome.

Campbell remarked that "trapped administrators" who cannot risk more objective evaluative data should use voluntary testimonials from those who have had the treatment, since human courtesy and gratitude will ensure favorable evaluation. At the same time, Campbell suggested that "qualitative knowing" from staff, clients, and families can add important information to more rigorous methods.

Scheirer says that much more than courtesy is involved in the bias of program recipients toward favorable evaluation. The bias is a function of participation and is also found in program staff and in evaluators who are participants. She cites social psychological theory and research, including the Hawthorne effect, (Roethlisberger and Dickson 1939), social desirability response bias (Edwards 1957) ingratiation attempts (Jones 1964), experimenter bias (Rosenthal 1966), social exchange theory (Homans 1974), placebo effect, etc.

Scheirer also cites the theory of "cognitive consistency", namely that once people engage in an activity, they report positive feelings. This even extends to hazing. Moreover, to change one's mind would admit that initial impressions were not perceptive. Shaw found that decision makers would put in more resources when initial consequences were negative and would put in the most when they were responsible for the negative result.

Scheirer reminds that the role of an administrator is to create success. This is a political creation, not scientific. The administrator assembles authorizations, resources, personnel and clients.

It is legitimate to gather support.

The bias is extensive, Scheirer says. Without measurable behavioral changes, favorable reactions were obtained for Head Start, for seven educational models of Follow Through, for juvenile delinquency experiments, and for encounter groups. She cites Gordon and Morse, who reviewed 93 published evaluations in Sociological Abstracts and found that when evaluators were affiliated with programs, 58% were reported as successes while only 14% were rated successes by independent evaluators. Similarly, studies with adequate methodologies found 22% success, and those with poor methodologies found 52% success. Scheirer concludes, "When looking for behavioral change, do not believe anyone's subjective impressions, including your own." This idea is contrary to the political theory underlying democratic government, Scheirer comments. The colloquial wisdom that a recipient knows where the shoe pinches is belied by the frequent choice of uncomfortable shoes for fashion. Long-term efforts of popular decisions may be delayed -- as when we belatedly realize that high school graduates cannot read.

Possible limits to the positive bias of participants, according to Scheirer, are when there are high attrition rates, when there is reorganization rather than new resources, and when participation is involuntary. She suggests that negative assessments by participants probably indicate organizational or implementation problems.

The Urban Affairs Institute strongly advocates that clients of a project should be interviewed from time to time in order to check the perceptions of project personnel and their administrative records of progress. There are many kinds of potential beneficiaries for A.I.D. projects - borrowers, patients, trainees, highway users, customers,

parents, farmers, etc. In several instances, A.I.D. has found that such interviews can uncover reasons for lack of participation or progress. This is consistent with Scheirer's comment that negative assessments can be quite indicative.

#### Independence of Evaluators

Considerable debate occurs about the possible inhibitions to objectivity when evaluators are dependent upon program managers.

Anderson and Ball say that evaluators are usually considered independent if they do not work directly for the people in charge of the program under scrutiny. They construct a grid showing varying degrees of administrative and financial dependence. They cite Renzulli who declared that as long as program people select and fund evaluators, the notion of independence is a mockery. Scriven argued that evaluation funds should not come from program budget or even from the same agency. He said that if an outside overseer like the Government Accounting Office is impractical, at least different people should monitor the program and the evaluations. He also proposed adding external consultants to internal evaluations and the rotation of internal evaluators.

Weiss notes that much evaluation is sponsored by some higher organization level in an agency, such as an assistant secretary. She says that these higher level decision makers can be more open-minded, being less concerned with organizational survival and more concerned about efficiency and impact.

In their case study about the New York City evaluations of the Gary education system, the Levines identified various kinds of influences such as the stake of decision makers in innovation, the tendency of supporters to minimize problems, the financial and political independence and career

interests of evaluators. Bernstein and Freeman found that the quality of evaluation studies (as judged by experts) decreased as dependence increase

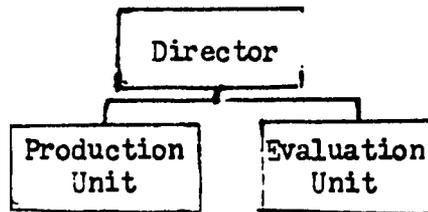
The debate about independence occurs because, despite the kind of arguments and evidence cited above, effective evaluation often benefits from close relationships between the evaluators and the program. Benefits are of two kinds, better information and better utilization. Thus Renzulli feared that complete independence would widen a gap between the evaluator and people whose honesty, trust and cooperation are needed. A dependent relationship may enhance the evaluator's responsivity to the significance of certain information and to program needs. Anderson quotes Longvold and Simmel who argue for a "lively personal involvement" in the organization whose practical problems are the social scientist's intellectual problems. They believe that objectivity could be retained by a strong identification with professional standards, through regular professional contacts and exposure to professional criticism.

At the Stanford Evaluation Consortium in 1976 two evaluation approaches were described by the analogy of someone riding parallel to a train and making observations through the windows or boarding the train and influencing the engineer, conductor, and passengers. Freeman is a strong advocate of participatory evaluations in order to influence program design by sharpening targets and to enhance progress by strengthening actions. Weiss says that to have immediate and direct influence on decisions, "inside" evaluation has a vital place and may stretch the decision-makers' sights a bit. Anderson and Ball strongly prefer some internal organizational role for formative evaluations (designed to improve programs while still in progress) but think the need for independence is clear for summative evaluations (which look to final results)

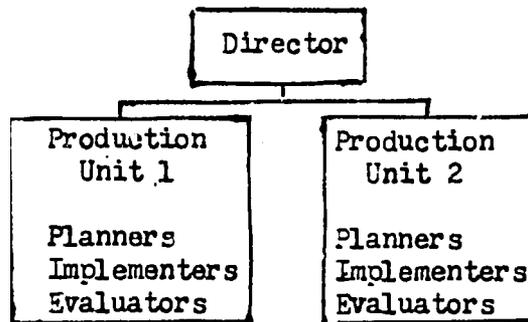
when credibility is important.

<sup>e</sup>  
Even for evaluators within an organization, there are variations in dependence. Two models shown by Scriven are:

A.



B.



Anderson and Ball prefer Model A. because it gives the evaluator better status and also access to the decision-maker. They cite Joseph Wholey that the responsibility for evaluation should be at a level appropriate to the decisions which the evaluation is to assist. They conclude, however, that whatever the degree of dependence, the evaluator should consider relationships that may be important, try to get clear agreement on lines of responsibility and build appropriate communications channels. Another device they suggest is an evaluation advisory board. They also point out practical considerations mentioned by Caro that the evaluator should have a prestigious position if in the same organization and strong professional credentials and organization backing

if outside the organization and by Warren that the evaluator should give an impression of technical competence, program sensitivity, and integrity.

A.I.D. has been experimenting with varying degrees of independence for evaluators. Within Missions and AID /W Offices, the role and status of Evaluation Officers varies considerably. Some are systems managers or facilitators; others are evaluators. Some report to the Director and have experience and prestige. Others are junior subordinates to Program Officers. Gradually, the Agency has recognized that certain questions of interest to AID/W can be answered satisfactorily only by evaluators organized by AID/W and not by field Missions which have a more limited perspective about programmatic issues. The growing use of outsiders by both AID/W and Missions (reported in Chapter V) is another indication of a desire for independence and objectivity (as well as a reflection of inadequate staff time).

The practicalities of having so many types of projects in so many countries will continue to force A.I.D. to rely primarily on internal evaluations. Therefore, A.I.D. must continue to stress safeguards for objectivity other than independence of evaluators. Examples are targets agreed in advance, systematic collection of data, and bringing various viewpoints to bear. These same safeguards can also help assure that outsiders will be objective, rather than expert opinion givers. Effective use of consultants also requires careful consideration of the purposes of evaluation and of the scope of work, as discussed in the next two sections.

#### Purposes of Evaluation

Gurel states that evaluators must check the reasons for the evaluation to insure that they seek answers to the right questions. Too often, he says, evaluators do not respond to managers' needs for instant

information and a limited inquiry but instead impose their own preconception that a more central issue should be extensively investigated.

Although Guttentag urged that evaluations be designed to include multiple levels of decision making, Weiss and the Office of Management and Budget strongly aver that the all-purpose study simply does not exist. OMB says that the intended use of an evaluation is a fundamental concern. There must be elaboration on the nature and context of the decisions for which the evaluation effort is undertaken and who the decision maker is. If several potential decisions are indicated, the evaluator needs a sense of relative importance.

Several writers are concerned about the possible unethical purposes of evaluations. Among the kinds of non-informational purposes noted by Weiss are postponement of decision, ducking responsibility, better public relations, fulfilling requirements of a grant or loan, justifying a program to Congress, and increasing prestige. She argues that if the real purpose is not better decisions and there is little commitment to use findings, the program is a poor candidate for evaluation. But Anderson and Ball concede that the public interest can be served on some occasions by declared efforts to rally support or opposition, that is, if the evaluator does not sweep contrary evidence under the rug and collects evidence which had not been collected before. They warn against use of "expert" testimony in such cases.

Anderson and Ball point out that outsiders such as superior government agencies or elected officials often demand evaluation results before a program has had a fair chance to take effect. In such cases, Freeman suggests evaluation of the process of implementation. Often the real reason for lack of impact is that a program is not fully implemented and

process evaluation may reveal possible improvements.

Rossi and Wright comment that when a program is justified, strong claims of effectiveness are often advanced. Since the programs typically attack difficult problems, rapid and large improvements should not be expected but evaluations should be sensitive to small effects. Freeman encourages a "biasing" of evaluation for results by concentrating on targets most amenable to change and by stressing strong rather than minimal program interventions. Weiss says a responsible position for evaluation is to put goals in sensible perspective, compare a strong intensive program with the ordinary level, and locate positive programs so they can expand.

Apparently part of the problem in the case which led to the present assignment was that one of the evaluators wanted to go beyond what the contract monitor considered to be the issues for evaluation. At the same time the scope of work did not make clear the nature of decisions facing the A.I.D. administrators.

Although some writers are concerned about being asked to look for benefits or impact prematurely, A.I.D. should not assume that prompt benefits are unlikely and so limit evaluations to process analysis. Processes can be usefully analyzed to explain lack of impact or limited number of beneficiaries if that is the case, but A.I.D. needs to look harder for impact. The logical framework sequence of means/ends should help lift sights from delivery of inputs or production of preliminary outputs to the project purpose of solving a problem and the goal of helping people. The evaluation which started this assignment was confined almost entirely to looking at behavior of the training institutions in the United States and made no effort to learn about the impact or

use of the training overseas.

### Scopes of Work, Monitoring Contracts

In a background paper issued in 1975, the Office of Management and Budget points out that program evaluation is a potentially valuable tool for government decision making which is only rarely useful to key policy makers and program managers. The paper advocates better evaluation management to get objective and useful studies, including relating evaluations to impending decisions, directing evaluation where systematic analysis is most practicable, giving more attention to uses of data, and better monitoring of evaluation projects. One aspect of the latter is adequate written guidance to an evaluation contractor.

OMB says that a Request for Proposals (RFP) should include five critical parts. The clarity and adequacy of these parts may have considerable influence on the quality of the evaluation effort.

1. Statement of intended use and objectives of the study

Relate to the type of decisions for which the evaluation is needed and who will decide.

2. Description of the program to be evaluated

Objectives, scope, size, duration, mode of operation, cost, available documentation should be summarized and then what is to be measured should be stated.

3. Statement of scope and constraints

This deals with methodologies and study techniques. Is the agency designing the study and buying the staff it lacks or is it buying expertise and asking the contractor to design the study? Are interviews desired and is "informed consent" required?

#### 4. Statement of the level of effort

Whether any statement is included may depend on the degree of flexibility on methods. If cost limits are administratively imposed, it should be clear that their origin is non-technical.

#### 5. Statement of deliverables

This should describe whether draft reports are wanted. For final reports, the number of copies requested should reflect a pre-determined dissemination strategy in addition to meeting such needs as file copies, clearinghouse copies, etc.

The above advice has been expanded in an article by an OMB analyst, Donald Weidman. Concerning use of the study, he admonishes that someone must decide what questions are most important. As a rule, he says really good answers can be expected to no more than one or two major questions. If several different kinds of studies are desired -- management audit, efficiency measure, impact determination, policy analysis -- he urges several RFP's, arguing that few can put together a study team which is strong in more than one or two methodologies. Weidman is undoubtedly thinking of nationwide American programs, rather than smaller A.I.D. projects.

Weidman says the program description should specify at least one acceptable measure of accomplishment. Evaluation studies are often rejected for irrelevance because official statements of program objectives are ambiguous and evaluators guessed wrong on the specifics to measure.

Most of Weidman's 10 helpful hints concern the scope of work and constraints. If a certain type of comparison of program effects, e.g.

urban vs. rural, poor vs. non-poverty, large vs. small, is desired, it should be stated. It is important not to state any constraints unless genuinely needed. Favorite offenses are sample size, team size, intermediate steps. He urges asking for standard evaluation procedures, adjusting expectations to known data availability, and using opinion as measures only when programs are intended to change opinions.

In the evaluation report, Weidman would require a clear separation between results of data analysis and the evaluator's judgments. He cautions against requiring formal reports during the evaluation. Preparing such reports subtracts from time for evaluation and informal contacts are more effective for monitoring progress.

OMB says that monitoring evaluation contracts differs from other monitoring. It argues that competence on methodological and technical issues is necessary. There should be sufficient contact to be cognizant of progress and problems through oral reports, telephone calls, site visits, etc. It also says that adequate procedures are needed within the legal framework of contract requirements for making changes in the methods or tasks as an evaluation proceeds.

In testimony to the Senate Committee on Human Resources concerning evaluation, Deputy Assistant Secretary of Labor Donald Nichols reported that, before implementation, evaluation designs are reviewed by a Research Clearinghouse in the Labor Department. His staff monitors contracts through site visits, seminars, and meetings. It reviews evaluation reports, writes an assessment of quality and appropriateness and publishes a summary.

John Evans, who heads evaluation for the Office of Education,

testified that he has a staff of 40 working on over 60 studies. They design all studies in-house "to assure that they are methodologically sound and relevant to policy requirements" before contracting for the studies.

The scope of work for the evaluation which led to this study was poor, primarily because it sought answers for too many questions in too short a time. The general A.I.D. situation on scopes of work and their relative importance are discussed in Chapter V, Information from Interviews.

The evaluator who complained thought that the A.I.D. monitor had behaved improperly by intervening to get a final report. It would appear that this was a move of desperation because of lack of previous monitoring of the evaluation team by the non-profit organization which recruited them. Monitoring was difficult in a situation where the evaluators often worked independently of each other and had limited time.

### Reports

Cox asserts that most managers prefer verbal to written communication-- they receive it quicker and can exchange information easier. He says that a written report is necessary for documentation but that simple and frequent verbal communication is more effective. The evaluator should, he advises, find the salient points such as a change in trends.

Anderson and Ball say that simple reports are desirable so that the lay public can understand the evaluation message. The opposite, when an official encourages complex reports to obscure evaluation results, introduces ethical problems for the evaluator, the authors remark.

Finally, Anderson and Ball note that prompt, interpretable feedback may be more important than sophisticated research design or formal reports for evaluation studies which are oriented toward program improvement. Close working relationships and informal approaches between evaluators

and program staff may facilitate timely decisions about program installations and processes. The caveat to this is for impact evaluations, which should have a more formal design.

A.I.D. has tried not to overemphasize reports in its evaluation process. The Project Appraisal Report (PAR) was regarded more as an administrative control device than as a transmitter of information. More recently, the Project Evaluation Summary (PES) asks for more content, so that Missions and AID/W Offices with projects will have a better baseline record for subsequent evaluations and so that AID/W can pass on more information to other Missions with similar projects. The PES consists of a printed cover sheet and then a narrative on various prescribed topics, most of which relate to elements of the logical framework approach to project design. It still encourages brevity of about a half page per topic.

In the case which provoked this study, the project monitor apparently sought a brief report of findings and recommendations that could be circulated to several institutions in the United States and overseas, rather than a detailed description of evidence obtained.

The kind of report expected from consultants varies with scopes of work. It seems likely that some consultants would benefit from better acquaintance with the logical framework approach and with the outline of the PES.

### Publication

The necessity for reports is a different issue from whether they should be published.

Diener and Crandall argue that scientists have a positive responsibility to advance knowledge and communicate findings accurately to other scientists. The authors say that an agency sponsoring an investigation

may want to control dissemination of findings or even forbid publication of unwelcome results. They conclude that except for reasons of national security, such restrictions on dissemination of information should be strongly opposed on ethical grounds.

Carol Weiss comments that while dissemination of basic research is essential and unquestioned, the majority of evaluation reports go unpublished. She wryly notes that many evaluation studies are not worth publishing, either because they were poorly conducted or because they are so specific that they are not generalizable. Practical considerations also work against publication. Administrators and staff who wanted the information to answer their questions begrudge the staff time to edit material. Evaluators are pressed to complete one job and start another.

The wise also recognize that administrators may be hesitant to air negative findings. (This penchant for "cover up" led to the Freedom of Information Act. However, this does not require publication but implies attention to indexing and controls on storage, retrieval and access). Weiss says that the presentation of negative results or lack of positive ones poses a problem for the evaluator, who must be sure the negative outcome is due to the program rather than to the evaluation design. The evaluator must also seek to avoid the overgeneralization which would cause the public to conclude that all similar efforts are worthless. If misinterpretations are made, the evaluator must speak out.

Some writers argue that the evaluator has a responsibility not only to those who commission the evaluation but also to the public. This responsibility is described in various ways. One is that the "community" is concerned and deserves to know. Another is that the social scientist

is interested in improving well-being and should add to knowledge about factors affecting success or failure of social programs. This obligation need not require publishing specific evaluation reports but could be met by drawing lessons from cumulative knowledge of several reports).

Weiss even argues that the evaluator must communicate to a variety of audiences in a balanced, intellectually honest way what the findings of the evaluation are. She says that is the minimal effort consonant with an ethical stance. Others recognize that the one commissioning the evaluation receives the information and would be expected to finance dissemination. If there is any distortion by a sponsor of research findings, a researcher is obliged to make a public clarification. By implication, this might be applied also to evaluation.

Quite different from these social and professional responsibilities is a responsibility for blowing the whistle such as exposure of cost overruns or allegations about inadequate safety precautions. This duty of good citizenship may mean reporting to responsible authorities (such as a prosecutor) rather than publication.

Brandl, a politician professor, argues that evaluators have a responsibility to present their case strenuously, even embarrassingly. Public dissemination of evaluations can help keep politicians honest. But he cautions that politicians are sold by people, not reports, and that ideas require time to be accepted.

The argument about publication seems more applicable to large domestic programs where public opinion plays more of a role than for A.I.D. Where investigative reporters may be concerned, publication might be a force for objectivity. If A.I.D. however, tried to prepare most evaluation reports for publication, there would undoubtedly be pressures to mute criticisms, hedge findings, and inhibit objectivity.

The A.I.D. approach has been to list formal evaluation reports from contractors with the National Technical Information Service of the Department of Commerce. Thus their existence is known to interested scholars who can request a copy. A.I.D.'s own Development Information Service publishes an accessions list and abstracts reports, both from contractors and within the Agency, so that it can answer requests for specific bits of information. This is probably adequate for most A.I.D. purposes, although a few reports may merit publication so that universities and developing countries could make more use of them.

### III. CODES OF ETHICS OF PROFESSIONAL ASSOCIATIONS

The scope of work called for a review of codes of ethics of consultant organizations and professional societies. Abstracts of three code are presented in Appendix C. Here the background of these codes and the application of some of their provisions to the A.I.D. context is discussed

#### Treatment of People

Some of the codes react to revelations about some medical and psychological research which had treated human guinea pigs unfairly. Thus there are various provisions to make sure that people are treated as ends rather than means and that their well-being is safeguarded and their dignity respected. Beyond taboos against causing harm, the codes emphasize that research subjects should give "informed consent" and should be able to withdraw from continued participation. Also, their anonymity should be protected.

In a few cases, A.I.D. projects may involve research into effectiveness of alternate methods of delivering health services, of contraceptives, or of nutrition additives. In such cases, the entire project may be regarded as a type of medical/evaluation research. It is to be expected that these projects will be conducted by professional contractors who will adhere to professional standards and that the A.I.D. project monitors will also be concerned that such standards are followed. Moreover, the National Research Act of 1974 created a commission to identify requirements for informed consent regarding participation in biomedical and behavioral research. Thus new A.I.D. rules or procedures should not be necessary.

#### Confidentiality

In a larger number of cases, A.I.D. will encourage and assist host

government agencies to sponsor sample surveys of client populations such as small farmers, co-operative members, or clinic patients. Experience both here and abroad has shown that program evaluations which rely entirely on data obtained from administrative and statistical reports and which stress progress toward planned targets may fail to deal with omission of eligible and needy people, unexpected side effects, or malfeasance. Therefore, an occasional objective check with the planned beneficiaries is essential.

For such surveys, the ethical standards calling for confidentiality and anonymity will apply. The U.S. Privacy Act of 1974 (P.L. 93-579USC552a) will also apply. However, possible misuse of data in the United States poses little threat to respondents. Violations of anonymity or confidentiality here would have little effect on villagers in another country.

Rather, the danger lies in the behavior of host governments, some of whom have a bad record of making individual statistical or survey returns available to tax or police officials. A.I.D. must be concerned not only for the well-being of the respondents but also for the validity of the answers received. Unless respondents can be convinced about confidentiality, candor will suffer. Special procedures may need to be developed for coding the forms on which interviewers record answers so that names and addresses are not on the forms, for safeguarding the forms until they are tabulated and perhaps for destroying them after tabulation. It may be advantageous to have the survey organization be non-governmental (e.g., a market analysis firm).

#### Host governments and secrecy

Some of the codes were adopted or revised in the early 1970's after scholars decided that they had been misused by the Department of the

Army for Project Camelot or in Southeast Asia. Revelations about C.I.A. cover organizations had also come out. Thus the codes contain strictures against studies with covert purposes. By a somewhat tenuous analogy, one might conclude that scholars would be hesitant to do field studies for A.I.D. unless the host government gave either full endorsement or collaborative participation.

The American Anthropological Association code is the only one reviewed which gives specific attention to relations with host governments, although the Sociologists' code says its principles apply to research either within or outside the United States. The Anthropologists admonish that every effort should be exerted to cooperate with members of the host society in the planning and execution of research projects. It calls for the research anthropologist to be honest and candid in relation to both his own and his host government, demanding assurance that he will not be required to compromise his professional responsibilities and ethics. Specifically, no secret research, reports or debriefing of any kind should be agreed to or given. The ban on debriefing could disappoint an A.I.D. Mission which might hope a consultant would be willing to talk more than he would write. In accordance with the Association's general position on clandestine research, no reports should be provided to sponsors that are not also available to the general public and, where practicable, to the population studied.

#### Publication in General

Scholars and their codes assume that publication will occur. Thus, except for the statement above from the Anthropologists, codes deal not so much with the right to publish as with particular aspects of publication such as protecting confidentiality or giving credit to student assistants.

It would seem that A.I.D. has an obligation to make evaluation reports available to host governments, contractors, and other donors with related projects. Beyond that, the operation of the Development Information Service, with its published accessions lists, together with the provisions of the Freedom of Information Act (93-502), adequately meet the need for access to information about development by scholars. A problem on publication which may arise concerns the publication of dissent, which is discussed after the topic of objectivity.

### Objectivity

The necessity for objectivity is a recurring theme throughout the codes. Thus it appears in connection with each of the topics discussed above -- treatment of people, confidentiality, host government relations, publication. In all cases, objectivity is regarded as one of the defining characteristics of a professional practitioner. (The origin of the word "profession" may be worth remembering. It began with people such as clergymen, doctors, and lawyers who took oaths or "professions" about their obligations and behavior. While today it sometimes refers to any occupation requiring extensive education, the word usually carries with it a connotation of adherence to scientific, historic, or scholarly principles and of maintenance of high standards of achievement and conduct, according to Webster's Third International Dictionary).

The psychologists state that while they demand freedom of inquiry, they accept the responsibility of competence and objectivity. They call for planning research so as to minimize the possibility of misleading findings, for thorough discussions of limitations of data and alternative hypotheses, especially when their work touches on social policy. They never suppress disconfirming data. They attempt to moderate institutional

pressures which may distort reports or impede their proper use.

While an anthropologist's paramount responsibility is to those he studies, he is also responsible to the public -- all presumed consumers of his professional efforts. To them he owes candor and truth. He should not knowingly color his findings. In providing professional opinions, he is responsible not only for content but also for explaining their bases. He should maintain a level of integrity and rapport in the field such that he will not jeopardize future research there. This does not mean to report so as to offend no one but to conduct research with honesty and open inquiry, as well as concern for the welfare and privacy of individuals.

The sociologists' code calls for "scientific objectivity." They must present findings honestly and without distortion or omission of data which might significantly modify the interpretation of findings. They must report fully all sources of financial support.

The excerpts above recognize that objectivity may not always be easy to achieve. Thus there are references to the influence of methodology, of institutional pressures, of bases for opinions, of efforts to maintain rapport, of financial support. It can be seen that it is easy to move from questions of ethics to the whole context for social research -- the techniques for obtaining and analyzing data, the organization and financing of enterprise, the use and publication of findings. If it is difficult for research, it is more so for program evaluation, which Weiss points out applies the methods and tools of social research in an action context that is intrinsically inhospitable to them.

The sociologists point out that results of their research may challenge long-established beliefs, leading to demand for dilution of findings. Similarly such findings may be manipulated by those in power for their own

use. Therefore, the code affirms the autonomy of sociological inquiry and states that the sociologist is obliged to clarify publicly any distortion by a sponsor of the findings of a research project.

While this is the only code which specifically calls for publishing a dissent from a report with which the researcher disagrees, others imply that there may be an obligation. The problem would, of course, be obviated if the report contained both interpretations, as suggested by the psychologists. Possible alternatives to "going public" mentioned in codes are to resign from the research project or to appeal to a committee on ethics of the professional association. In general the codes apply to individuals, rather than group research or evaluation activities.

Because of the inherent difficulties of achieving objectivity, it is most important for A.I.D. to consider its organization and procedures for evaluation. A.I.D. may commission evaluation studies from a contractor or association or university. This organization in turn hires people, who may find themselves working alongside A.I.D. and host government personnel. Their work is supervised and monitored by the contractor, A.I.D., and perhaps the host government. The result merges and submerges individual contributions. Then this group report is reviewed by A.I.D. and host government decision makers, who may accept, modify, or reject recommendations. Under this confused situation, the codes do not clarify the obligations or rights of the various parties nor give guidance about review and appeal as well as publication of rejoinders.

#### IV. PROPOSED CODE OF ETHICS IN EVALUATION

Program evaluation is called an emerging profession in a book issued last year (Profession and Practice of Program Evaluation. Anderson and Ball). This recognizes that it is an activity with its own characteristics distinguishable from related social science professions. The most important distinguishing feature is probably the fact that evaluation is an element of management contributing to action decisions. Related social science professions, on the other hand, are concerned either with the advancement of knowledge for its own sake or with the treatment of individuals.

##### Values in Evaluation

As a profession, program evaluation can be expected to have values and ethics. Indeed, values probably play more of a role in evaluation than in some other professions. Consider the title itself. "Program" is defined as an activity for "improving" social and economic welfare, (Anderson). Both "improve" and "welfare" connote value judgments. "Evaluation" means judging the merits of a program in achieving that improvement (Weiss), again a value judgment. Moreover, philosophers of knowledge now recognize that science cannot claim to be "value free" and that social sciences are especially affected by values.

##### Work on Standards

Anderson and Ball comment that one of the most telling indications that a field is coming of age occurs when the idea of standards and sanctions for its activities is broached. In 1975 that began for educational evaluation when the Lilly Foundation gave a grant and a Committee to Develop Guidelines and Standards was set up with representatives from eleven educational and professional associations. A fourth draft of standards is now being reviewed in regional conferences.

Shortly after it was established in 1977, the Evaluation Research Society appointed a Standards Committee. A draft is expected in June.

Pending the issuance of an ethical statement from the appropriate professional organizations, Anderson and Ball have tried drafting one which they hope will serve as a starting point. It is reproduced in its entirety in Appendix D and discussed below. Since their book appeared, Anderson has been appointed to head an Ethics Committee for the Evaluation Research Society.

### Mutual Responsibilities

One feature which distinguishes the draft statement from codes of other professional groups is its recognition that several parties are involved and that each has responsibilities to the other. These parties are the person(s) or agency with major responsibility for securing the services of an evaluator (dubbed Commissioner of Evaluation), the person(s) or agency with major responsibility for planning, carrying out and reporting evaluation activities (the Evaluator), and administrators, staff, program recipients and others with a role in the program being evaluated (called Participants).

Despite this recognition of mutual obligations, the proposed code is drafted primarily in terms of the evaluator, with the others responding to him. This reveals the authors' experience as evaluators. In actuality, the agency which commissions the evaluation has the main responsibility for some phases of the process. However, this criticism is one of emphasis and presentation. Most of the important points seem to be covered.

### Scope of Work-Purposes of Evaluation

The process starts with the Commissioner of Evaluation providing as complete information as possible about the program, the expectations for

the evaluation, and the conditions and resources for carrying it out. The concern of the authors is that sometimes agencies seek evaluations for possibly unethical reasons such as rallying support or opposition, proving an argument, or postponing a decision. It would seem that a more important concern is that evaluation reports will not be relevant or useful to decision makers unless the evaluator understands what issues are pending. This was discussed under Scopes of Work in Chapter II.

Therefore a key point of a code of responsibilities is that there be an adequate scope of work. Even when an evaluation is conducted by agency staff rather than contractors, it is advisable for staff to understand what the decision maker needs. If the Commissioner of Evaluation does not provide adequate details, the Evaluator should seek clarification for his own protection.

Paralleling the responsibility of the Commissioner of Evaluation to provide information is a responsibility of the Evaluator to call attention to his values and orientation which may affect the evaluation.

### Contract

The next step in the process envisaged by the code is a contract or "agreement" that is ethically, legally and professionally sound. This is a mutual responsibility and provides mutual protection. The code then says that the Commissioner of Evaluation should not press for work to be performed in advance and the evaluator should refuse to work without a contract. After a contract is reached, both parties have obligations to fulfill their commitments to the best of their ability and to acquaint each other with any unexpected problems which arise, attempting to work out solutions.

### Collecting Data

The Evaluator then has a responsibility to adhere to professional standards in conducting the evaluation, including appropriate provisions for informed consent and confidentiality. The Commissioner supports this adherence.

At this point, the proposed code adds a touch of operating realism not found in other professional codes, namely that the Evaluator should carry out data collection with as little interference as practicable with program operation. In return, the Commissioner encourages cooperation by Participants and Participants should provide accurate responses to legitimate requests.

No recognition is made in the code of the role of the commissioning agency monitoring the contractor's work. This was one of the complaints in the case which provoked this study. The Office of Management and Budget advocates continuing and thorough monitoring.

### Special Responsibilities

The code states that the Evaluator has two special responsibilities to the Commissioner while working on an evaluation. One is to acquaint him with any aspects of program philosophy or operation which appear to violate ethics or law or are physically unsafe, even if such observations were not part of the Evaluator's specific charge. The Commissioner would be expected to recognize the evaluator's duty in noting such problems and to consider his observations seriously. Weiss advises that the Evaluator should also inform the Commissioner promptly about any "surprises" without waiting for a written report.

The other special responsibility is to tell the Commissioner about any requests received from superior agencies (e.g. Department head, OMB,

Congress, etc.) for information or testimony, to ascertain whether the requests are valid, and then to acquaint the Commissioner fully about the response.

### Reports

The code then points out the obligation of the Evaluator to present a "balanced" and timely report in a usable form. The report should also spell out limitations of the investigation and the evaluator's values. Weiss says the evaluator has an obligation for unqualified objectivity and for as much usefulness as he can devise.

The reciprocal obligation stated for the Commissioner is simply to avoid misuse of evaluation findings. In A.I.D., at least, there is a more positive obligation. A.I.D.'s "doctrine" is that an evaluation is not complete with the submittal of a report by a contractor or members. There should be a formal review of findings in a meeting of responsible officials (including host country officials for overseas evaluations). This review brings all points of view to bear and is a device for objectivity when evaluations have been cursory or conducted in-house. It also lays the foundation for decisions about follow-up actions which complete the evaluation process. Hence A.I.D.'s official evaluation report (the Project Evaluation Summary) lists action decisions before it summarizes findings.

The proposed code says that the Evaluator should reserve the right to publish rejoinders to any misinterpretation or misuse of evaluation results. As discussed in the preceding section on other professional codes, this is an inadequate answer to the problem of dissent.

### Publication and Dissemination

The proposed code calls for the Evaluator to identify other groups

with a legitimate concern for the results of the evaluation and to make the results available, while the Commissioner simply advises about groups with an interest and encourages dissemination. This is a backwards statement of bureaucratic responsibilities -- the Commissioner should do the staff work and pay the costs of dissemination. Since most reports are not published, special effort is required to duplicate and distribute copies or summaries of reports. The provisions of the Freedom of Information Act, as CMB points out, have obvious implications for indexing, storage, retrieval and access control procedures.

#### Re-Analysis

Next, the proposed code deals with the unusual situation when someone wants to review the evaluation data -- either another scholar attempting to learn more about evaluation processes, or some outside agency such as the General Accounting Office trying to judge the validity of findings. The problems involved in this process were discussed by Baratz (see bibliography).

#### Share Basic Knowledge

Finally, the proposed code says the Evaluator should share knowledge derived about basic processes in education, administration, development, etc. This is limited to the scholarly approach. The anthropologists' and psychologists' codes go farther and include an obligation to improve the well-being of people. Weiss states an obligation for the evaluator to improve social change efforts.

## V. INFORMATION FROM INTERVIEWS

In an effort to understand the background of the case which led to this report, I talked with auditors Eugene Linck and Benjamin Stevens. They felt that the evaluators had not been given good criteria to use, commenting that the project paper had been fairly precise but that the grant document was vaguer. They thought that the evaluation had, in effect, consisted of hiring expert opinions. The non-profit organization which recruited the evaluators was not active in supervising them. The debriefing in which the evaluators presented their separate reports was attended by four A.I.D. officers.

As might be expected from an auditor's perspective, both Linck and Stevens were concerned that A.I.D.'s system of evaluating itself may inhibit objectivity. Linck thinks such a system requires fairly constant oversight and believes that the Program and Policy Coordination Bureau should evaluate evaluations. This might flush out inconsistencies between bureaus.

### Dissent

Only one other instance of dissent was discovered. This happened after the testing of the new Title II scope of work by a team fielded by Robert R. Nathan Associates, Inc. which was accompanied by Edwin Fox, Evaluation Officer for O/FFP. Fox reports that a nutrition expert hired for that assignment wrote Nathan Associates after the report was completed concerning some differences of interpretation about a technical point. Nathan concluded that the point was not important for the total evaluation or the recommendations. Fox thought the dissent might reflect some personal frictions apparent during the field work.

Both Russell Diltz and Michael Snyder of the Office of Contract

Management were definite that the contractor should be held responsible for his product. Diltz went so far as to say that if A.I.D. set out to arbitrate the differences of evaluation team members, it would, in effect, be using personal services contracts which are illegal. A contract should have a reasonable description of a scope of work, time requirements, and report specifications.

Snyder commented that the reputation of a contractor is behind a report. He knew of one instance in which an employee of a contractor submitted a critical report which the contractor tried to disown. Snyder insisted that the contractor had to stand behind his work and the report was re-written. On the other hand, Snyder would not object if a scope of work called for the presentation of various viewpoints from a multi-disciplinary team.

The Evaluation Officers reacted the same way as the Contract Officers. Thus Bernice Goldstein of LA said that a firm with a contract is responsible and should put a report together. She would expect the firm to work out differences although she would not object to a minority view submitted as part of a report, comparable to the frequent practice with special commissions. She thought that Missions were usually aware of shadings of interpretation, having discussed findings with a team or reviewed drafts. In one case where this probably did not happen a Mission rejected a report. In this case, the problem was probably unsatisfactory work rather than unwillingness to accept criticism.

Joan Silver of NE, who often helps put together mixed teams of experts recruited individually, takes care to get compatible people and has had no dissent. One person is designated the team chief. If there were some difference of opinion, Silver would expect it to be reported in the

official report rather than as a subsequent dissent. She does not feel that writing a specific request into a contract for separate views of different disciplines is advisable -- briefings cover it. She thinks a team is more likely to say it is of two minds over matters of opinions rather than of evidence. Silver cited a recent case of dissent which was not within a team. A Mission Director wrote to say that he thought one member of a two-person contract team had a professional bias. Consequently, the Director was inclined to ignore an opinion which differed from that in another evaluation report on a related project.

Charles Molfetto of DSB said there may be a recent case of disagreement within a team -- at least they were behind schedule with a report. With mixed teams of A.I.D. and outside people, DSB generally asks the contractor to chair the team and calls for reports within 45 days after on-site visits.

When Practical Concepts, Inc. engages an associate who is not a regular member of the firm, the employment contract makes clear that the associate has a right to his judgment and that if he does not agree he is encouraged to make a separate submission that PCI will forward with the report or later. This has yet to occur. The closest instance was not for A.I.D. and was methodological. A team had different perspectives on what should be done and consulted the client while the work was in process. The client preferred the dissenter's approach. PCI tries to be objective and tries to improve a program -- both approaches that tend to encourage consensus. Leon Rosenberg thinks that if A.I.D. were to institutionalize PCI's contract clause, the result might be clumsy.

Albert Brown could recall only one instance in which the American Technical Assistance Corp. had unreconciled viewpoints. Philosophically, he does not like to stifle different viewpoints and thinks a good report should surface two or more if they exist, so that the political decision

maker has the whole picture. Brown remarked that companies like to do this for self-preservation, especially on feasibility studies. Then if a project fails, the company can say that it gave both sides. At the same time, he does not like a report that is ambiguous. Therefore, he would work hard to get a recommendation based either on good back-up or on judgment as to the best course.

Paul Schwartz of American Institutes of Research was puzzled as to how the case described in Chapter I could have happened -- he felt it must reflect poor management by the contractor and poor monitoring by A.I.D. He says that a team almost always has some disagreement, either interpersonal or interdisciplinary, but thinks this should come at the beginning of a study and should concern procedural issues such as criteria or weights.

Schwartz said that AIR's concern goes not to differences within a team but rather to different reactions to a final report from varying interest groups. AIR recently had two major assignments for agencies other than AID where its teams fulfilled their assigned scopes of work with findings backed by excellent methodology and data. Yet the public outcry was anguished and vociferous. In one case AIR demonstrated that teaching Hispanic-American children in Spanish was not helping them either in reading skills or in adapting to regular classes in English. This answered HEW's question but the Hispanic-Americans were concerned about preserving a culture. In the other case, AIR concluded that a supervised "diversionary" program for juvenile offenders had no better results in decreasing recidivism than imprisonment. This was the concern of the Justice Department but various reform groups were also concerned about treatment accorded juveniles.

AIR has now concluded that "objectivity" means that programs must be considered from more than one point of view. Therefore, its policy about an evaluation on a controversial issue is now to do several things: hold a preliminary meeting with the funders of the evaluation and get agreement for broader criteria, establish an advisory panel, send the draft report to interest groups, and include comments from the groups in the final report.

Schwartz gave an A.I.D. example of an analogous approach although not an evaluation. He had been invited by David Scherer to a meeting in Mali on livestock in the Sahel. Schwartz was to talk about training for semi-literate herders. After three days listening he told the group he disagreed with their approach of optimizing cows instead of people. Later Mali invited him back to help design a project to improve the quality of life of herders -- a project not yet adopted.

AIR would not guarantee staff the right to dissent in a report. They have created spin-off firms and divested themselves of staff and contracts so that they can have a homogenous commitment that concern for the quality of life is first and "science" is second. They refuse jobs with which they disagree although they take "harmless" projects to pay the rent. Since the new policy with "irreverent" responses to RFP's, the AIR has had a better record of acceptance.

The approach of Development Alternatives, Inc. is not unlike that of AIR in that they too have a definite philosophy and emphasize the impact of development on people and the need for involvement of local people -- their interests, aspirations, and leadership.

Donald Micklewaite says if team members do not see eye to eye, the firm tries to work out a consensus. If a cable reports a problem, he goes to the field. Footnotes may say that there is a different way to look at a situation.

Donald Weidman of the Office of Management and Budget mentioned some techniques to take different points of view into account. One was that advocated by Marcia Guttentag which was described in Chapter II. Another approach is to get agreement before the evaluation on the relative priorities of different kinds of information. A third approach is to include a minority report, but with a careful distinction in both the majority and minority report between the facts and the interpretation of action implications. Oliver Taylor of OMB commented that users of evaluations should be widely informed and that different perspectives are worth noting but that good staff work requires recommendations.

#### Use of Contractors

Since this report deals primarily with evaluation work by outsiders (contract firms, individual consultants, non-profit organizations), some estimate of volume was sought.

In the Latin American region, Missions are making increasing use of contracted assistance, according to Bernice Goldstein. She says this trend reflects staff limitations and also the increased complexity of evaluating large projects such as sector loans affecting many people. The Missions usually use Indefinite Quantity Contractors (IQC). The LA Bureau is also planning to use outside assistance for longitudinal analyses of development experience in individual countries and is preparing RFP's for a separate contracting competition outside the IQC framework.

Last year, the Development Services Bureau had 213 routine evaluations and 59 in-depth evaluations scheduled, Charles Molfetto reported. The routine evaluations are done by staff. The in-depth evaluations are generally done by a team of two to six people. The teams are often mixed, with some A.D. personnel and some outsiders. If the project involves

research, one of the "outsiders" is usually a member of A.I.D.'s Research Advisory Committee. Generally, DSB asks an outside contractor to chair the group. Molfetto said that the Office of Population uses contractors extensively. DSB seldom uses an ICQ firm.

The Asia Bureau has 107 evaluations scheduled, of which six are characterized as "special". Vance Elliott estimated that between a quarter and a third of these evaluations would use Americans travelling from the United States and in addition perhaps 75 to 80 local nationals who are not A.I.D. employees would be involved. Most of the Americans will be from AID/W. Elliott said more consultants would probably be used if the region had more Project Development and Support money (regional funds for research, feasibility, and evaluation studies). This use of direct hire personnel contrasts with the recent experience of the Office of Food for Peace, which has used contractors because it lacked travel funds for direct hire. Elliott pointed out that projects prepared for the fiscal years 1979 and 1980 are required to include money for evaluation so that the use of contractors can be expected to increase. However, the projects just started in 1979 may last for five years.

Elliott believes outside evaluators are preferable and hopes to encourage their use by having Missions plan fewer evaluations. This would mean scheduling evaluations at critical points of their projects rather than annually. Also he proposes that several small projects might be combined for one sub-sectoral evaluation. As part of affirmative action, evaluation teams are sometimes enlarged. Elliott prefers to issue new RFP's but does use ICQ firms when time is short.

Africa Bureau has recently revitalized its evaluation efforts with the naming of Paul Saenz as Evaluation Officer. Saenz was not sure about

volume of evaluation, since about 50% of the scheduled 400 were behind schedule. He estimated that about a quarter of the evaluations were proposed by AID/W and most of these would use outside contractors. The Africa Bureau has had its own group of six IQC firms. They have been used primarily for project design, although often design of a new project may start with evaluation of a related terminating project.

Joan Silver said the Near East region has about 30 evaluations a year which use outside personnel. In addition, teams go out for feasibility and project design studies. A close-out evaluation is linked with design of a follow-on project. The Bureau is using its Project Development Support money to finance an A.I.D. team member on inter-country studies of Food for Peace programs. The A.I.D. member facilitates comparison when different contractors are used in different countries. Otherwise, the Bureau must restrict travel by direct hire people. The Near East Bureau seldom uses IQC firms, finding that it can put together teams of individuals and avoid overhead charges.

#### Composition of Teams

Several people expressed concern that the composition of teams could affect their objectivity and the possibility of dissent, but considerable difference exists on the best composition.

Two bureaus, Development Support and Near East, like to put teams together of different individuals and perhaps also include an A.I.D. person. Silver said that she has located some unusually able people. She cited the example of a Palestinian on the Stanford faculty who is an engineer and has a Ph.D. in political science. His Arabic linguistic ability has been useful in North Africa. Silver likes to recruit new blood and then to use people for assignments in more than one country. Kolfetto said

the DSB often used the "old boy network for team members or that it used people "already broken to harness" -- seeking a grand guru to give a blessing. Don Micklewaite of Development Alternatives, Inc. says AID is afraid of "stuck-together" teams and refuses to put firm members on other teams.

Paul Saenz of the African Bureau was concerned that the same firms and the same people have reviewed a project several times. He thought the results were neither objective nor professional, since evaluations often dealt with such problems as late delivery of inputs rather than project re-design or impact.

On the other hand, Albert Brown of American Technical Assistance Corp. argued that ATAC has sometimes furnished useful continuity as it progressed from sector study, to implementation advice to evaluation while Mission personnel changed three times. He said that his firm has had arguments with Mission Directors because ATAC proved that a project was "lousy" but were asked back because they had advised on ways to improve the project. He cited such an instance with a Mission in Asia. For sector assessments, ATC uses a multi-discipline team; for evaluations it uses an economist/evaluator and a technical specialist if there is more than one person.

Michael Snyder of O/CM said in the current review of proposals to choose firms for new Indefinite Quantity Contracts, considerable weight was being given to firms which would use regular staff members rather than recruit outsiders for each job. This is the policy of Development Alternatives and American Institutes of Research and generally of Practical Concepts, Inc. but American Technical Assistance Corp. argues that it insures professional integrity by having a majority of a team be non-company.

Saenz was skeptical whether company members who are generalists would be able to ask searching technical questions. On the other hand, Vance Elliott of the Asia Bureau feared technicians who did not know evaluation techniques for establishing criteria and measuring indicators out were willing to give opinions after interviewing project personnel. He said some Missions have asked for evaluators by name and resisted having additional or new people on teams.

### Techniques for Objectivity

Various procedures for enhancing objectivity have been mentioned in the preceding sections of this chapter. In addition, some of the interviewees had other suggestions.

Leon Rosenberg of Practical Concepts, Inc. comments that A.I.D. evaluations are better than they were 10 years ago in that they usually include some solid evidence. A.I.D. is ahead of some agencies which have publicly stated that evaluation of some programs is impossible. Nevertheless, there is still much of the same discourse as before. Recently, Samuel Daines of PCI analyzed over 600 A.I.D. and IBRD projects concerned with rural poor people and found only two evaluations of impact. Rosenberg thinks that failure to focus on either intermediate or ultimate impacts (purpose or goal) leaves the evaluator to check on the original opinions of the project designers.

Micklewaite of Development Alternatives commented that checking on what the designer said several years ago can be terdentious. He wants his people to use an information system built into the project. If there are no data, they devise a few specific retrospective questions on such items as use of agricultural techniques to ask village leaders or householders.

Rosenberg cited a recent evaluation done under Daines' direction in Haiti (see Annotated Bibliography -- Appendix B). Two preceding contract evaluations had followed scopes of work and examined processes for mobilizing inputs and producing outputs of a project to provide credit to small coffee producers but had not attempted to examine impacts. Both made useful recommendations and expressed doubt about success in raising income of small farmers. Daines' interviewers found significant rises in income and well-being. Rosenberg suspects impact may often go undiscovered. He says that the sample need not be large and that a control group can usually be found in new clients for the project who did not get "treatment" during the period being evaluated.

Micklewaite commented that development is hard. He thought evaluators should look for the perception of small changes. Rapid reconnaissance surveys do not allow much time and put more stress on judgment. Even so, one can seek data. For example, on a seed multiplication project in which the project personnel did not know where the seeds were distributed or how they were used, the evaluators were able to talk to only 10 farmers but still obtained some clues. This was an instance of a project which dealt with only one element of a system and so had limited impact. Micklewaite likes comparative studies as a way of seeing what works. DAI just completed a study of 17 private voluntary organizations in two countries and found no activities which were optimum but could see what was reasonable. They have signed a contract with the Entente Fund for a series of small impact studies and will do two evaluations a year.

The Latin American Bureau has found comparative studies useful not only for future programming guidance but for improving the projects being studied. Vance Elliott said that Asia Bureau is planning to do such studies

of projects in agricultural research, potable water, and small scale irrigation.

In addition to its decision to look at several criteria reflecting various interest groups (described above under the heading Dissent) American Institutes of Research likes to look at intermediate changes when final impacts may not have occurred. Knowledge may be gained about some things of value and about some reasons for failures, while a classical experimental evaluation might yield no incremental knowledge. For example, in evaluating Law Enforcement Assistance Administration grants to citizen groups AIR will look at such observable things as whether groups got organized and what they did rather than solely at crime in the neighborhood. For A.I.D., Schwartz would decide what part of a logical framework could be achieved "and document the hell out of it."

#### Review of Evaluations

For its ambitious longitudinal studies of total experience in several countries, the Latin American Bureau is trying hard to get studies which will impress two audiences as valid. One is Congress which wants to know how effective development assistance has been. The other is the development "community" which wants to know what works. The Bureau has had a task force working on a scope of work and will recruit widely. Beyond that, it plans advisory teams in AID/W and the USAID and perhaps a panel of outside critics.

The Development Support Bureau, which uses a group review for a scope of work, usually does not convene a formal review meeting to consider an evaluation report, although the Agriculture Division recently had a meeting with all bureaus. If a DSB Officer Director rejects evaluation recommendations, he must tell why in an attachment to the Project Evaluation Summary and the Deputy Assistant Administrator has the final decision. The

Evaluation Officer follows up on implementation of action decisions.

Vance Elliott and Paul Saenz have re-instituted reviews of evaluation reports (both Project Evaluation Summaries and special reports) in the Asia and Africa Bureaus respectively and are sending comments to Missions. Elliott keeps a tickler file on action dates in the PESs. A more important kind of review will occur in Africa Missions as a result of a new directive reemphasizing existing Agency policy about signatures on Project Evaluation Summaries; namely, that the Mission Director must sign indicating approval of methods and findings. Perhaps more important, the African Bureau now also requires a signature from a senior host government official.

The Near East Bureau requires that an evaluation team submit a draft report or a formal briefing to a Mission before it leaves the country. The Mission must have time to react. If the Mission disagrees, it cables Washington. No instances have occurred of a report not being issued.

Practical Concepts Incorporated tries to set up formal Evaluation Review meetings of all interested parties when it is in the field, although in some cases it can't. The practice of Development Alternatives Inc. to seek host government viewpoints and of the American Institutes of Research to seek advice at the beginning and end from interest groups was mentioned earlier.

#### Scope of Work

The Development Support Bureau requires that a scope of work for an in-depth evaluation be submitted to its Evaluation Officer two months before the evaluation is scheduled. Last year the Bureau started reviewing the scopes with a committee consisting of the Evaluation Officer, Senior Program Analyst, Project Manager and representatives of the Office of Research and Utilization Division (all parts of DSB). If a geographic

bureau is directly concerned, its comments are also requested.

The DSB Bureau has issued guidance on preparations of scopes of work (See Appendix F) which is quite detailed. It reminds that project results are to be measured by indicators in the logical framework matrix and urges review and up-dating of the matrix if necessary. The next section of the guidance says that the contents of the evaluation are the most important part of the scope of work and suggests about a dozen possible topics, including project management and strategy, some elements of the logical framework, and utilization of project results, which is of special concern for AID/W projects. This part of DSB guidance differs from the OMB advice cited in Chapter II that the number of questions in an evaluation should be limited. The guidance also refers to project documents and reports. It calls for nomination of team members. (Guidance #2 on team members is also reproduced in Appendix F). The Guidance contains almost nothing on methodology.

In the Latin American Bureau, Project Implementation Orders (PIO/Ts) containing scopes of work are usually prepared by Missions. The Office of Development Resources may get involved, either because of a visit to the field or because a Mission wants help. DR informs the Evaluation Officer, Bernice Goldstein, who sometimes advises, although most PIO/T's go directly to the Office of Contract Management without her review. Goldstein thought that better guidance on the qualities of a good scope of work is needed, since airgrams are scattered. She distributes the draft of Handbook Chapter 8 widely. She suggested distribution of some model scopes, along with cautions about common pitfalls.

Vance Elliott said many Asia teams have little in the way of a scope of work, especially when Missions telephone directly to an AID/W direct

hire employee and recruit him for a team or have him recruit some other named individual. If a contractor is used, there is a scope of work. Big Missions prepare them and small Missions cable ideas to AIE/W technical people. Elliott has to sign off on all scopes. He considered them quite good.

The African Bureau Indefinite Quantity Contractors have been under the supervision of the Office of Development Resources, which consulted with the Evaluation Officer. These contracts were scheduled to expire March 31. New contracts will be supervised by Paul Saenz, the Evaluation Officer, for evaluation work. Since November, Saenz has met with each evaluation team and talked about Agency evaluation requirements and the Project Evaluation Summary outline of topics. This has affected the work plans of teams and the questions they ask.

Joan Silver, the Near East Bureau Evaluation Officer, said the Bureau uses detailed scopes of work and terms of reference. She characterized preparation of a study design as amounting to about one-third of conducting an evaluation. Scopes for small evaluations are prepared in Missions; for large ones, in Washington. She reviews Mission scopes. In Washington, she prepares some and technicians prepare some. A scope has five elements:

1) Objectives of evaluation

What is information for?

2) Level of evaluation

Is it directed toward management of inputs, production .  
of outputs or impact on purpose and goal?

3) Focus of evaluation

What issues should be emphasized?

4) Methodology

Define the parameters, e.g. a household survey, statistical reliability.

5) Kind of report

One of the O/CM officers commented that Near East scopes of work may be so detailed that they inhibit initiative. Silver said the section on methodology is meant to be guidance and the Bureau may ask the contractor to help. However, the Bureau practice probably reflects the fact that it is more likely to contract with individuals than with firms and forms teams of people who may not have worked together before. Contractors are briefed by AID/W and also at the Mission. They are told to speak up if the Mission differs from AID/W.

Michael Snyder, who handles all Indefinite Quantity Contracts for evaluation, says that the Office of Contract Management reviews scopes of work only for contractability. That is, will it be possible to know whether the contractor has done what he was supposed to do? An evaluation which is methodologically poor may still be contractable. The guidance issued by O/CM on work orders (see Appendix E) calls for background on the reason for the order, a description of the project, the scope of the work order (as differentiated from that of the project), the available documentation, the extent of support for the contractor and the nature of the product desired. The PIO/T estimates level of effort and budget as guidance to O/CM but this is omitted from the Request for Proposals so that the contractor can submit his own estimates. In Snyder's opinion, the scopes of work for work orders under Indefinite Quantity Contracts are about the same quality as scopes for competitive RFP's. Likewise, he has not noticed any difference in quality between regions.

The Office of Food for Peace recently contracted with Robert R. Nathan Associates, Inc. to prepare a model scope of work to be used for

evaluating total country programs for Title II of P.L. 420. (See Appendix B - Annotated Bibliography for description). Title II programs may include supplemental foods for infants and lactating mothers, school lunches, and food-for-work projects. Connected with them may be family planning or nutrition education. Some of these activities are non-projects in that they do not aim at specific development changes. They may also lack baseline and progress data. For these reasons and because O/FFP wanted a broader approach than projects, they sought guidance on a uniform "systems" approach which could facilitate inter-country comparisons. The proposal was tested in Bolivia and Sri Lanka and has since been used in Tunis and Morocco. Edwin Fox, the Evaluation Officer, thought the new scope of work would result in more objectivity if all the questions were asked. Thus far, all the evaluations have suggested a sampling system to obtain better data.

Leon Rosenberg, of Practical Concepts, Inc. commented that A.I.D. does not write clear evaluative scopes of work to insure that it receives findings rather than opinions. For instance he would think that A.I.D. should tell which of the logical framework targets it wants measured and what additional items it also wants measured. It should request a statement of professional standards -- where the data were obtained and what validity they have. Scopes of work for separate evaluation contracts may be more detailed than for work orders under Indefinite Quantity Contracts but are not superior, according to Rosenberg.

Albert Brown of American Technical Assistance Corp. said he pays little attention to a scope of work, other than to assure that it receives technical compliance. He believes that scopes may come from many parts of the Agency and that an AID/W writer may not know the project well. In the field, ATAC teams see points which need emphasis and others which do

not. When ATAC hears of a work order, it talks to the technical officer to learn more than is in the scope about what is wanted and who to recruit. In any event, they know they must deal with the A.I.D. evaluation system, in which the project design sets forth indicators and quantitative measures.

Don Micklewaite of Development Alternatives, Inc. also has little interest in scopes of work, since they are usually written by the AID/W desk or Development Resources officer. More important than the contents of the work order is the time allowed -- six weeks gets more data and the result is better than three weeks. DAI teams start with the USAID to see what they want and then work with the host government. (Micklewaite says that truly collaborative evaluations are the exception). If possible, DAI will also try to ascertain the interests of the target population. As the firm name indicates, Development Alternatives has a philosophy about the nature of desirable development -- that it should improve well-being of people. Thus even if a scope of work emphasizes production, DAI will also try to look at distributional impact.

Don Weidman of the Office of Management and Budget (whose article on scopes of work was reviewed in Chapter II commented that junior staffers in the government (not necessarily in A.I.D.) are often asked to prepare RFPs and tend to be specific about the wrong things. Thus they may try to spell out procedural techniques for evaluation but be vague about the information the government wants. He would prefer training program officers about characteristics of good RFPs rather than having agencies imposing central review of scopes.

American Institutes of Research has also decided that it has a commitment to quality of life and often submits "irreverent responses to

to RFPs which don't know what they're talking about". This was discussed under the heading "Dissent" in this chapter.

#### Training of A.I.D. personnel

Several suggestions were made about formal and informal training which might improve scopes of work. Charles Molfetto has had two hour sessions with each staff group in the Development Support Bureau on evaluation procedures. He also volunteered that it is important for all project managers to attend the Program Design and Evaluation seminar. He had seen an improvement in the products of those who have attended. He suggested that a brief refresher for those trained 6 or 7 years ago would be helpful.

Bernice Goldstein urged prompt issuance of the draft handbook chapter on project evaluation.

Paul Saenz had recently sent several airgrams on evaluation, each of them signed by the Assistant Administrator. He also talks about methodologies with each Mission Director who visits Washington.

From time to time Joan Silver goes to a Mission on temporary duty to prepare a scope of work or to help conduct an evaluation. She uses these occasions to alert field personnel to qualities of a good evaluation and to possibilities for objectively verifiable indicators. Recently, for example, she prepared a scope of work for a complex project to help establish a research institute in Egypt. Both the Egyptians and their advisors from the Massachusetts Institute of Technology, all highly educated people, were sensitive about an outside team second-guessing their technical work. Silver's scope stressed institutional development and also utilization and relevance of research, rather than technical evaluation. The scope was, in effect, training on elements of

institutionalization. The approach was for each of 14 sub-projects to prepare their own report and then for a brief visit by a team to look at the recommendations.

In another case, a scope on a participant training project designed to promote Western management techniques developed indicators such as reorganization of work and reading of U.S. periodicals and then had a section on selection and training of interviewers. This was sent to other supporting assistance missions.

Russell Dilts called attention to the brief course on contracting for personnel outside O/CM. It stresses scopes of work.

#### Contract Monitoring

For Development Support Bureau team evaluations, the project manager always accompanies the team, although he may not be a team member. For institutional grants under Sec. 211d, a team must include an A.T.D/W Office Director or a Mission Director.

For the Regional Bureaus, as might be expected, Washington-based officers are not sure about the kind of monitoring given teams in the field. The Africa Bureau had recently directed that the second person in each Mission (Deputy or Assistant Director) should be the Evaluation Officer. Someone from the Development Resources Office usually goes out with a team. For the future, Saenz hoped to have a small staff and to send people on occasion. For LA Missions, Goldstein assumes Mission review of drafts. In the Near East, Silver says Mission orientation of teams and review of drafts is expected. She tells teams who to contact in Missions. Vance Elliott in the Asia Bureau expect that, since many evaluation teams spend most of the time interviewing rather than gathering hard data, project managers may often succeed in conning the teams.

Michael Snyder commented that, if anything, he thought the monitoring was sometimes overwhelming.

None of the contractors interviewed considering monitoring, or lack thereof, a problem. Each firm has experienced people who would, as a matter of course, keep people informed as work proceeded.

## VI. ANALYSIS OF ISSUES AND RECOMMENDATIONS

The case which provoked this study is an exception in several regards. First, almost no other instances of dissent were discovered. Second, the project was administered in AID/W and financed by a grant, rather than being a joint host-donor project. Third, the evaluation team lacked an appointed leader and a supervising contractor.

### Aspects of Concern

At second glance, the case contains certain aspects which may be fairly common to the Agency. For instance, the scope of work was poor. It gave the evaluation team nine tasks, a heavy workload to accomplish in three weeks. The first task was "to review grant documents to ascertain purposes and goals of the project, the budget plan...and time plan..." As the auditors pointed out, the grant document was vaguer than the project paper. Why didn't the commissioner of evaluation specify what objectives concerned him?

The next six tasks specified elements to be evaluated: financial management, administrative and teaching staff in three locations, curriculum, participant selection, overseas training, and participant follow-up (using reports for the last two). Nowhere were any criteria suggested. On the basis of looking at these process elements of inputs and outputs, the team was then to describe the extent to which purposes and goals had been fulfilled! Finally, it was to recommend improvements in operations or suggest new directions.

Considering the number of tasks, the short time, and the lack of criteria, it is obvious that A.I.D. was not buying an evaluation but merely some professional opinions. Several people interviewed feared that this

may often be the case in Bureaus other than the one where this case occurred. The presence of dissent would not necessarily indicate an effort to achieve objectivity but simply a difference of opinion between evaluators who lacked enough facts to make a valid finding.

Another aspect of the case which is of general concern is that the nature of evaluation may not always be understood. The use of outsiders even if they are experts, is not synonymous with evaluation, which is a systematic effort to get objective evidence about progress towards established targets. That last phrase -- established targets -- means that good evaluation rests on good project design. Besides targets, well-designed projects also contain built-in means for collecting progress data and take essential external factors into account. A project plan should also include an evaluation plan which designates critical points in the life of the project when evaluations should occur and specify issues that are expected to be important for each evaluation. With such a foundation, evaluations can be more objective, whether done by insiders or outsiders.

A third aspect, related to the extent of understanding about evaluation, was the emphasis put on processes rather than impact of the project. Several of the questions about management and staff probably should have been answered by a normal project monitoring. The evaluation then could have looked at the crucial issue of whether trainees were applying their new knowledge and whether that was leading to the desired results in their countries. Instead of a team of two experts in the United States, perhaps one good junior investigator should have visited several countries.

Several of the Evaluation Officers and contractors interviewed thought that A.I.D. did too little in the way of impact evaluation. This is

an especial problem for AID/W projects, many of which support other projects administered through Missions, so that their impact is one step removed. While host-donor projects are considered typical, AID/W projects now constitute nearly half the Agency's portfolio in numbers (not dollars). However, the lack of attention to impact is not confined to AID/W. Often people assume that impact will be delayed or that ascertaining facts about impact is difficult and so the effort is not made.

A fourth aspect illustrated by the problem case was the discrepancy between the size of the assignment and the resources devoted to it. Lack of resources was cited frequently by interviewees. It has several dimensions. One is that outside teams are hired because Mission staff lacks the time or because travel by AID/W staff is limited. But then the time for the teams is limited so that they do not obtain evidence.

Another resource problem is lack of data. This has two aspects. Better macro-statistics are needed for analyzing problems and selecting projects. But then evaluators sometimes try to use existing macro-data in lieu of project-specific information. Although the situation is improving, many projects still fail to collect selected baseline and progress information. Sometimes this requires part of the project budget; at other times it may simply require attention by advisors and their host counterparts.

Thus the case contained several aspects of general concern -- a scope of work which sought opinions on many subjects rather than evidence about a few, a lack of understanding of the nature of evaluation, emphasis on processes rather than impact, and insufficient resources for evaluation.

Although dissent has been an exception, the suggestion has been made that dissent might be one possible approach to a desirable characteristic of evaluations; namely, enough breadth of vision to reflect reality. Our

review of literature reveals that focusing too narrowly may be a difficulty, that evaluation tends to select a few variables and to accept existing structure. Moreover, evaluations tend to have a positive bias when they are done by staff and when they reflect views of "treated" clients. Therefore we need to bring other viewpoints to bear -- host officials, other donors, and untreated clients.

### Approach to Recommendations

The purpose of greater breadth and objectivity in evaluation would be to provide a better basis for informed management judgments. The hypothesis is that such judgments will contribute to the Agency's goal of economic development -- now defined as better living conditions for poor people.

As we consider possible ways to obtain breadth and objectivity we need a sense of proportion. First, evaluations are more likely to produce good evidence if they examine a limited number of issues, according to several authorities. Second, the researcher's ideas for good research design may conflict with the manager's desires for speed and economy. Finally, the professional researcher recognizes limitations to objectivity which are inherent in the questions asked, the methods used to find answers, and the values and cognitive biases through which data are screened and interpreted. Evaluation, like traditional management, is also a matter of judgment and cannot be completely "scientific".

The preceding chapters commented about various actions already adopted by A.I.D. for enhancing objectivity. These have included the logical framework approach to design and evaluation with its advance statement of targets and arrangements for data, the policy for collaborative design and evaluation, the formal review of evaluation findings, and the use of outside

consultants. The system is already complete enough that no startling changes in procedures can be recommended. Instead, the recommendations which follow deal with ways to re-emphasize existing procedures or to "fine-tune" them.

Then the recommendations move beyond procedures, which tend to become rote. As rote, procedures may be followed without really facing up to key issues. Therefore, it is strongly recommended that A.I.D. strengthen its evaluation by a substantive emphasis -- more attention to the human side of development.

#### Better Understanding of Evaluation

The recent delegation to the field of authority to approve all projects up to \$5 million makes a good understanding of the elements of design more necessary than ever. Missions may (and should) reduce paper work, relying more on verbal communication about the strategy of a project. But this simplification should not mean that they stop thinking in terms of a series of means/ends hypotheses or that they fail to arrange for progress data. If project design becomes careless, evaluation may play more of a role as Missions review and correct their projects after preliminary experience with them. Training or briefing Mission Directors can be important, since they set the tone for their Mission's approach. Some Directors have taken the seminar on project design and evaluation, either in Washington or at their Missions. The Africa Bureau Evaluation Officer is meeting with Directors as they visit Washington.

Apparently, maintaining a good design and evaluation system requires constant vigilance. Personnel turnover and reorganizations make it easy to ignore procedures. Maintaining a system has several elements. One is an up-to-date written manual or handbook. A chapter on project

evaluation was cleared by bureaus several years ago but has not been published because other chapters on program and policy evaluation are not ready. Another element is monitoring or review. The African Bureau recently revived its system after more than two years without an evaluation officer. Review of evaluation reports are also being re-instituted in the Asia Bureau. Circulation of reports or scopes of work which illustrate good methodology and give ideas on indicators and evidence can also be useful. Another kind of review, for scopes of work, will be suggested later in this chapter.

Because of turnover, training about design and evaluation continue to be necessary. It is also surprising that some veteran A.I.D. employees have yet to take such training. For example, in November 1977, the writer was one of a three-man team in Djarkarta giving four seminars on project design and evaluation to Indonesians. The Mission Director thought all Americans were familiar with the logical framework system and should not attend. Yet at the January 1979 session in Washington, two Americans on home leave from USAID/Indonesia attended the seminar at their own request and found it useful.

A group needing special attention are the International Development Interns, who will be given more responsibility as many retirements occur because of the Supreme Court decision upholding the mandatory age 60 for Foreign Service employees. Recent interns have not been given the full one-week seminar in project design and evaluation, since they react strongly against classrooms. They do get fairly good exposure to design with a field project and review. However, they are often thrown into evaluation as soon as they get on the job in either AID/W or the field.

As new Indefinite Quantity Contracts are awarded or new personnel join

experienced firms, these new personnel should also attend the seminar.

Understanding about design and evaluation should be fostered among host personnel as well as A.I.D. staff. The Professional Studies and Career Development Division of the Office of Personnel and Training has recently contracted with a firm which can give the Program Design and Evaluation Seminar overseas in English. It also needs to update existing materials in Spanish and French.

1. Recommendation - Revive efforts to enhance understanding of the Agency's project design and evaluation system. To this end, the Office of Program Evaluation in PPC, in cooperation with Bureau Evaluation Officers, should:

(a) Update the existing cleared draft chapter on project evaluation and publish it without waiting for other chapters in the proposed Evaluation Handbook.

(b) Brief new Mission Directors on design and evaluation or encourage them to take the seminar, either in Washington or overseas.

(c) Review with the Professional Studies and Career Development Division (PSCD) a practical and acceptable way to expose interns to more understanding about evaluation.

(d) Encourage PSCD to carry through on arrangements for more overseas training for American and host government officials about project design and evaluation.

(a) All geographic bureaus should review and comment on evaluation reports from Missions, at least on a sample basis.

(f) Distribute samples or excerpts of good evaluation reports and scopes of work.

## Collaborative Evaluations and Formal Reviews

A technique to achieve breadth of viewpoint is collaborative evaluation. This has been A.I.D. policy for years, but interviewees indicate that it is far from general practice. Missions probably do better on collaborative design but that could probably be practiced more widely too. Collaboration between host and donor is essential, not only for realistic and effective projects, but also for good relations.

One reason for collaborative design is that sovereign nations should set their own priorities on direction of development. The United States may not agree with all aspects, but should be able to find some common ground. Another reason is that most of the implementation will be the responsibility of the host government. Projects are likely to be more realistic if the host has decided about staffing and schedule. Progress will be better if the host understands and accepts the project targets and implementation plan.

Collaborative evaluation is another step in getting effective implementation. Officials who participate in reaching conclusions in an evaluation about necessary changes in a project are more likely to accept those findings and put the recommendations into effect.

Collaboration can also improve the quality of evaluations as well as their acceptance. Despite cooperation throughout a project, host officials usually have somewhat different views about an activity than do the Americans. This differing viewpoint is part of the reality which needs to be examined. Host officials can facilitate access to administrative information about a project. Although some officials in a capital city may be just as ignorant about their rural citizens as foreign advisers, most will probably know more and interpret evaluation findings more accurately.

In the beginning, collaborative evaluations require more time and effort. Staff and policy officers of the host government must be briefed on the purposes and procedures. Their time is limited and the Mission may have to wait until they are available. Language or cultural differences may slow communications. But eventually the investment of time will pay off in less trouble with implementation and less time to negotiate project and program changes, as people who tried in various parts of the world have testified.

The process will not seem so awkward or so forbidding when it is realized that collaboration need not prevent expedient division of labor between host and Mission officers on both the staff and policy levels. Thus staff might jointly agree on the issues needing answers for pending decisions and clear these evaluation proposals with policy officers. Then staff might divide the tasks of collecting and analyzing data or might use outside evaluation researchers, either from abroad or from a host university or market research firm. Then staff could analyze the findings and prepare recommendations for a joint policy review.

Even if the staff work for evaluation is not collaborative, the policy review can certainly be a joint one. A formal evaluation review with all interested parties is another technique to achieve a broad perspective. The attendance should include the host project director and A.I.D. team chief, host planning officials and A.I.D. program officers, representatives of other donors with related projects, etc. For effectiveness, a review should focus on a few key questions growing out of the evaluation findings of staff and should emphasize future adjustments rather than criticisms of the past. Sometimes such a review can be on a subsector and cover several projects.

Formal reviews, like collaborative evaluations, have also been urged by A.I.D. for years and are used from time to time in various Missions, depending on the project, the style of the Director, and the working relationships with officials of host agencies and other donors. They work best when participants have been briefed. If a host official were to call and preside over a review, a good atmosphere would be set.

One possible advantage of formal reviews is that suppression or disregard of evaluation findings by Mission Directors may be a little more difficult. This would be a kind of application of the "sunshine" principle of open hearings.

Apparently, formal review sessions are seldom held for AID/W projects. Considering that many of these projects lack the inherent checks and balances which come from host country acceptance and implementation, special care should be taken with evaluations in order to assure project relevance and utility for development. Review meetings might be attended by people from other institutions, by visiting foreign officials, and by representatives of geographic bureaus as well as the AID/W Bureau sponsoring the project. (The Development Services Bureau does have a requirement that an Office Director must justify to the Deputy Assistant Administrator his decision to disregard an evaluation team recommendation. This requirement meets one of the possible purposes of a Review meeting).

The question has been raised whether such a requirement should become general Agency practice, just as offices now report on each audit recommendation. The writer recommends against such a rigid requirement for several reasons. More precautions may be advisable in DSB because it lacks the interaction with a host government. Reports from ad hoc evaluation teams lack the stamp of authority represented by an official

audit report, which has been thoroughly reviewed before issuance. Missions are already overburdened with reporting requirements. A dialogue about how to respond to evaluation recommendations between a Mission and an AID/W Bureau could be much more drawn out and burdensome than between an AID/W Office Director and his Assistant Bureau Chief.

A possible procedure to try to assure collaborative evaluations with joint reviews would be to require the signature of a host official on the Project Evaluation Summary report. Several Missions proposed this when the draft PES was circulated for comments. The redesigned form has space which can be used for this purpose but does not require the signature because it was thought that some governments will not delegate power to sign documents or will authorize signatures only after cabinet meetings.

However, Paul Saenz, when he was Deputy Director and Evaluation Officer in Panama, found joint evaluations most beneficial and regularly obtained a signature of a Panamanian official. Now as Africa Bureau Evaluation Officer, he has persuaded his Assistant Administrator to require host signatures. At first, the Bureau specified the Deputy Minister of the Ministry with the project but has since agreed to accept the signature of a bureau chief.

2. Recommendation - Require collaborative evaluations and encourage formal review sessions. To this end, the Office of Evaluation in PPC, in cooperation with Bureau Evaluation Officers, should:

(a) Prepare a directive and airgram for issuance by the Administrator.

(b) Call attention to the advantages of formal Evaluation Review meetings for both field and AID/W projects. (The methodology

part of the Project Evaluation Summary can be used to report attendance at such Reviews).

(c) Consider whether host signatures on Project Evaluation Summaries should be required by other Bureaus besides Africa.

### Scopes of work

The contribution of a good scope of work toward achieving an objective, balanced evaluation report which supports useful action recommendations will vary with circumstances, as Chapter V Information from Interviews makes clear.

Sometimes, too precise a scope may interfere with good evaluation. Such a scope of work may inhibit useful methodological initiatives and innovations by able professional evaluators, who are apt to be more expert in methodology than most A.I.D. project or program officers who prepare scopes. Another possible shortcoming of too precise a scope of work is that it will direct attention, as do many projects, to a limited aspect of an inter-related system. If an evaluation adheres to the scope of work, it may omit the most important evaluation finding; namely, the development will not occur until other parts of the system are also modified.

All four of the reputable contract firms interviewed were frank to say that they paid only the necessary minimum of attention to scopes of work. They may be more important for teams formed of individuals who are not responsible to a reputable firm which is the case for several bureaus. Recognizing the limits to the influence of a scope of work on the quality of evaluation, some effort to improve scopes still appears advisable.

As the original case and as the DSB guidance on scopes of work

(Appendix F) illustrate, A.I.D. scopes of work are apt to ask for diffuse evaluation and put too little attention on evidence. The scope should tell which elements of the logical framework are most important at the time, should request the evaluator to describe his methodology for obtaining data and their validity, and tell which conclusions are based on evidence.

If A.I.D. wanted to institutionalize procedures for handling possible dissents in evaluations, the contract would be the applicable document. However, the occurrence of dissent has been rare enough that this hardly seems necessary. Generally contractors will continue to be responsible professionals. It is to be hoped that A.I.D. contract monitors will keep in close enough touch with teams that they can anticipate possible difference of opinion. These can be handled by arranging for more data, by discussion, or by inclusion of footnotes and minority views as part of a report. The awkwardness arises not from the existence of dissent but rather from efforts to suppress it so that it emerges after an evaluation instead of as part of it.

Beyond dissent, there is a question whether decision makers will be better served when different perspectives are encouraged. An example was cited of a feasibility study in which an economist and an engineer endorsed a proposed project while a social anthropologist expressed grave doubts. The economist who headed the team gave short shrift to the anthropologist's views. (In contrast is the approach mentioned to me by Dr. James Oxley, Chief of the Livestock Division, DSE. He says they investigate the social aspects first, before sending out technical specialists). While A.I.D., unlike the IBRD or some authors, does not include feasibility studies under the term "evaluation", evaluators do have to wrestle with problems created by poor feasibility studies. Also, feasibility studies for follow-

on projects are often conducted in conjunction with evaluation of terminating projects.

For multi-disciplinary teams, it may be desirable on occasion for a scope of work to request that the views of each team member be noted, along with the team recommendation. However, the weight of authority from both literature and interviews considered that this procedure might be awkward for management and could be artificial. Some suggest that the matter is usually covered in briefings of teams by AID/W and Mission officers.

3. Recommendation - Improve the quality of scopes of work for evaluation contracts and work orders. To this end, the Office of Evaluation in PPC, in cooperation with Bureau Evaluation Officers, should:

(a) Supplement the general instructions issued by O/CM for Indefinite Quantity Work Orders with guidance about scopes for evaluation which would:

(i) Urge that the number of issues be limited and that impact, even if only interim impact, be stressed.

(ii) Require a description of methodology -- where the data were obtained and what validity they have -- and a designation of which conclusions are based on evidence.

(iii) Require the evaluator to contact host officials for opinions.

(iv) Suggest consideration of whether multidisciplinary teams should be asked to present the views of each member as well as a team recommendation.

(b) Institute a review of scopes of work on a trial basis in the Bureaus where Evaluation Officers do not now review them, to see whether improvements in draft scopes justify the effort and possible delay.

(c) Continue to urge project managers to take the short course on contracting, which will help them for all types of contracts as well as those for evaluation.

#### Resources for Evaluation

Vance Elliott, the Asia Evaluation Officer, suggested spreading money for evaluation further by grouping related projects in a single evaluation. This would save travel money for AID/W staff or consultants. Unless the teams stay long enough in the field, however, grouping projects might again force them to skimp on evidence unless projects really have certain information aspects in common.

Elliott is also concerned that many projects are still evaluated annually, even though key decisions may not be pending and annual evaluations have not been required for several years.

Hopes about resources for evaluation lie in the requirement for inclusion of evaluation money in all project funding starting in the current fiscal year 1979. This also has the advantage that estimating funds needed will force formulation of a more definite evaluation plan for the life of the project. However, such funding will not help the majority of projects for some time, since projects approved in prior years may have a life of up to five years.

As for the problem of data, A.I.D. has recently sponsored projects using the Bureau of the Census, which can help get some basic national statistics. In addition, knowledgeable Census advisers in country could

also counsel project managers on how to get minimum evidence about project indicators. Instances cited in Chapter V illustrate that able professional contractors can obtain useful information, even retroactive baseline data, with modest sampling and interviews. The PPC Office of Evaluation has urged increased support for EuGen projects.

4. Recommendation - Increase the resources available for evaluation.

To this end, the Office of Evaluation in PPC, in cooperation with Bureau Evaluation Officers, should:

(a) Insure that the policy on inclusion of evaluation money in future project funding is adequately followed.

(b) Arrange for the allocation of more to Program Development and Support Funds for fiscal 1980.

(c) Ease restrictions on travel funds for evaluation, either by arranging a re-allocation of administrative appropriations for travel or by arranging to use Program Development and Support Funds for travel as well as for evaluation contracts.

(d) Bureau Evaluation Officers should review Office and Mission Evaluation Schedules to:

(i) Be sure evaluations are related to project phases and pending decisions rather than being annual.

(ii) See if evaluations of related projects can be combined or terminal evaluations done in conjunction with appraisals of new projects.

Evaluation Officers should enlist the help of Desk and Development Resource Officers in this review.

## Benefits for People

The four preceding recommendations -- to enhance understanding of evaluation, require collaborative evaluation and encourage formal reviews, improve scopes of work, and increase evaluation resources -- should do much to improve both the objectivity and breadth of evaluation. Objectivity will be helped when better designs establish criteria and arrange for data. Scopes of work can also emphasize criteria and evidence. More resources will give teams more time and help get more data. Breadth of perspective will come from collaborative evaluations, formal evaluation reviews with all interested parties, and from scopes of work that require the obtaining of host viewpoints and possibly the inclusion of views of all team members along with team recommendations.

One viewpoint may not be covered by the above procedures -- the project clients. A good evaluation must learn what is happening to the people whom the project is designed to benefit. One of the descriptions of the logical framework approach which is often cited is that it is ethically neutral. That is, it can be used to design logically consistent projects which do not necessarily follow A.I.D. policies to help poor people, preserve the environment, etc. These policy considerations may be included or excluded -- but the concern about them is not inherent in the procedure.

Because of this, USAID/Ecuador modified the logical framework worksheet or matrix several years ago. It divided the column on indicators into two parts, one for measures of physical or institutional changes and the other for measures of changes in well-being.

When the draft Project Evaluation Summary was circulated, USAID/Indonesia was also concerned that benefits for people might not be

covered by the narratives about project outputs, purpose, and goal. Therefore, the Mission proposed and AID/W accepted the addition of a separate narrative about benefits for people.

What we are talking about here is not just another procedure. It is a matter of substance. Generally, A.I.D.'s approach to evaluation has been that the criteria by which projects are to be judged are project-specific. The evaluation system does not tell what the targets should be, but simply helps people sort them out according to a mean-ends sequence of inputs-outputs-purpose-goal. The first exception to that approach was inclusion of the item on benefits to people in the Project Evaluation Summary.

The time has come to recognize that a neutral, procedural approach is inadequate. Even though this report makes four procedural recommendations, its main argument -- saved for the end of the report -- is that the best way to meet the concern of those who commissioned this study about how to ensure a greater breadth of vision is not procedural but substantive. The main problem is not whether dissent is being suppressed or whether viewpoints such as those of host officials are being disregarded. Presentation of different viewpoints -- whether different disciplines of evaluators, philosophies of host officials, perspectives of other donors, or concerns of A.I.D. technicians -- cannot assure that evaluations will pay adequate attention to what is actually happening to the human side of development, to poor people.

The development community, A.I.D., and the Congress were first alerted to this aspect of development by a World Bank study of 30 years or more of experience in Colombia. Assistance had begun with the InterAmerican Institute under Nelson Rockefeller in 1941. It had continued with Point 4, I.C.A., A.S.A., A.I.D. and with the CAS, IERD, IDA, and various UN agencies.

The record was good -- growing GNP, improving foreign exchange, rising per capita income, increasing agricultural production. Yet the IBRD found that real income of Colombian farmers was no higher after 30 years! The trickle down theory of economic development did not work. At that point, Administrator John Hannah proposed and the Congress imposed some policies or standards for A.I.D.

Thus A.I.D. is now concerned about improving the lot of poor people and is no longer so concerned about balance of payments, GNP, infrastructure, or production. At the same time, both A.I.D. and host governments are struggling because officials grudgingly change their ways of selecting and managing projects and resist the unknown. Sometimes their commitment to the new philosophy is more cosmetic than substantive.

In a time of experiment, evaluation should receive more emphasis. What works and what doesn't? If a project overlooks the real needs of the poor, an evaluation should reveal the slight. If some approaches are successful, the word should be spread to other Missions and other governments. If evaluations do not deal with this problem, they are too narrow.

Much of the literature and several of the contractors consulted felt strongly about the need for broader perspectives. Some of the writers put the idea in terms of techniques for understanding reality. Others went beyond considerations of the philosophy of knowledge or scientific method to ethics, arguing that the professional sociologist or anthropologist or program evaluator had an obligation to improve people's well-being and hence should consider broad criteria in judging social programs.

A somewhat similar range occurred among the evaluation firms. Perhaps the most direct approach was that of Leon Rosenberg of Practical Concepts, Inc. He urged that A.I.D. could and should do more impact

evaluations and said these would be more objective than examinations of project processes. Thus he would stress indicators for project purpose and goal. This approach would be apt to stress the effect of the project on people. He said PCI always tries to recommend ways to improve a project's impact. (At the same time, Rosenberg remarked that an important contribution of an expert was to test the validity of project assumptions about factors external to a project but critical to its success. Thus he would look beyond the logical sequence of planned targets.)

The strongest views were expressed by Donald Micklewaite of Development Alternatives Inc. and Paul Schwartz of American Institutes of Research. Both are convinced that development consists of improving the well-being of people and cannot be evaluated by looking only at output or production targets or possibly limited purpose targets in a logical framework. Successful achievement of such targets may not indicate real development, on the one hand, and failure to achieve targets may reflect a slighting of critical social aspects, on the other hand. Micklewaite and Schwartz state that their firms will hire only those people who accept this philosophy and will accept only those evaluation assignments which permit this approach. It is a new definition of objectivity. Schwartz declared "Our concern is for the quality of life, 'science' is second."

This effort to see a broader and truer perspective does not necessarily imply dissent. All the contractors believe that they should give the decision maker the best recommendation possible, based on good evidence. But they want to get evidence about more of the system, and specifically about more of the parties at interest. Thus they seek to understand the viewpoints of host officials and project clients and to obtain

data about what happened to clients.

The methodology for obtaining such evidence can be varied. In some cases, direct information can be obtained as part of the process of administrative reporting needed for operating the activity, e.g. information on out-patients at a clinic or borrowers from an agricultural bank. In other cases, proxy indicators will show trends but not absolute amounts, e.g. consumer purchases or school attendance may indicate family well-being. Sometimes behavior can be observed, e.g. use of agricultural techniques.

But in the final analysis, one must go directly to a sample of the target population of the project. The Urban Institute has demonstrated conclusively that administrative data has its limitations. Even if it is honest and timely, it can only tell about the people who availed themselves of services and cannot explain about the potential clients who were not reached. For instance, a sample survey in Costa Rica about a successful rural clinic project revealed that, despite the excellent record on numbers treated, the project had reached only about half of the potential clients. Others did not come because of such reasons as a poor image caused by failure to charge fees, inaccessible clinics, and inconvenient hours of operation. Sometimes the administrative data are inaccurate, either because they were processed carelessly or even because of malfeasance of project staff. An evaluation of cattle inoculation in one country revealed that inoculators sold sera and used water in their syringes!

Although accepting the need for client reactions, this writer has previously argued that interviewing clients must necessarily be an occasional and exceptional procedure because it requires professionals to draw a sample, devise and test questions, and train interviewers. However,

instances cited by both Practical Concepts, Inc. and American Technical Assistance Corporation convince me that the process need not be as time consuming nor as expensive as previously thought. It should be done oftener and is becoming a standard of more projects, especially in Latin America.

A client survey should have a very limited number of questions about actions rather than opinions. This elicits better answers from respondents. It reduces errors by enumerators. It costs less and can be tabulated more easily. In many cases, the survey can find a control group among newly enrolled clients who have not received any benefits yet. Some interviewing of eligible people who are not enrolled may also be needed to reveal reasons why the project is failing to reach them.

As discussed in Chapter III on Codes of Ethics, A.I.D. should be concerned about precautions for preserving the confidentiality of respondents from any inquiries of their own governments for tax or other purposes. This problem should be considered by the Program Evaluation Committee to decide whether some special action is advisable to alert Missions to the need for care. For example, an airgram could go to the field. Or a reminder could go out each time a Mission reports in a project evaluation plan or an annual evaluation schedule that a survey is contemplated. At this stage, however, the importance of reaching a small sample for missing insights should not be made too difficult with ethical safeguards. If the sample is small, the host government will have little interest in misusing individual responses.

The recommended new evaluation perspective of substantive attention to the aspect most likely to be underevaluated and reported -- the human side of development -- will be consonant with A.I.D. policy. It is also

consistent with the idea surfaced in the review of literature that values are important, should be explicit, and that A.I.D. should specify some of the most important ones. Moreover, a substantive emphasis for evaluation will add interest for both staff and decision makers who may often see evaluation as a procedural formality which is a nuisance. It will help distinguish evaluation from monitoring and will raise sights from inputs and outputs to purpose and goal -- an objective for the past ten years.

The new emphasis should be applicable to all field evaluations. For AID/W projects, the question of utilization now receives attention, but now the question should be -- does this AID/W knowledge or service get used by Missions to help poor people? Attention can also be given to benefits for people in the increasing number of comparative evaluations being initiated by AID/W. These evaluations are being undertaken to answer questions raised by policy officials whose perspective is broader than that of individual projects. Such comparative evaluations also have the potential of greater objectivity, as the literature review pointed out.

Chapter V mentioned comparative studies planned by the Latin American Bureau and Asia Bureau. The House Appropriations Committee requested a study of rural electrification. PPC has set up a division for policy and comparative evaluations. Each of these should stress the benefits for people.

5. Recommendation - Promote a new emphasis in evaluation evidence about benefits for people. To this end, the Office of Evaluation in PPC, in cooperation with Bureau Evaluation Officers, should:

(a) Announce the new emphasis in a circular from the Administrator which mentions the following actions that will be undertaken.

(b) Review project Evaluation Plans received as part of Project Papers to be sure this element is covered.

(c) Review Project Evaluation Summaries and call attention to Missions and Offices which are giving only opinions rather than data about the item of Benefits for People.

(d) Stress benefits for people in the section of scopes of work calling for evidence about project impact.

(e) Recommend more frequent use of sample surveys of target populations. (Precautions may be needed to protect confidentiality for respondents).

(f) Stress benefits for people in comparative evaluations initiated and conducted by AID/W. bureaus.

## Appendix A - SCOPE OF WORK

The Contractor will provide a report analyzing the problem of objectivity and candor in evaluation and the ethics of dissent.

1. The Contractor will search the literature on the subject including the codes of ethics of consultant organizations and of other key professional societies; he will abstract or summarize the results of this search as appropriate.
2. The Contractor will interview leading contractors doing work with A.I.D. as well as A.I.D. officials concerned with evaluation, contracting and legal affairs, to define problem areas which exist currently or which are likely to arise given the increasing volume of evaluation activities and the increased attention given to evaluation results.
3. The Contractor will prepare a preliminary report on the above emphasizing the most important problem areas.
4. Based on this report and on additional technical guidance provided by PPC/E, the contractor will provide an analysis of the issues involved in promoting

objectivity and candor in evaluations, in defining the ethics of dissent, and in protecting the rights of individual researchers and firms or institutions participating in evaluations or evaluative research. The paper will also discuss measures needed to improve and protect the ability of the government to make effective use of the results of evaluation and/or evaluative research and the appropriate disposition of information gained on the basis of confidential relationships.

5. The Contractor will provide a draft discussion paper for review by PPC/E; the draft will include the contractor's recommendations for further action by A.I.D.

6. The Contractor will provide a final report by March 31, 1979.

## Appendix B - ANNOTATED BIBLIOGRAPHY

The scope of work called for a search of literature on the subject of objectivity and candor in evaluation and the ethics of dissent. The findings were to be abstracted or summarized as appropriate.

Chapter V of the foregoing report reviewed findings from the literature by topic. Here, summary and comments are arranged by author. To facilitate selective review, key words in the comments are underlined.

Anderson, Scarvia B, and Ball, Samuel The Profession and Practice or Program Evaluation. San Francisco: Jossey - Bass. Inc. 1978.

This was the most pertinent and useful discussion encountered. As a recent work, it has progressed well beyond the hang-up evident in some older articles whose authors were struggling to make evaluations conform to all the norms of scholarship. Anderson and Ball recognize that evaluation is not pure research but is management oriented. However, this recognition is still somewhat limited because they regard evaluation as ending with the preparation of a report rather than when decisions are made which use findings and recommendations in a report. Another minor limitation is that the book reflects its origins in a conference at the Educational Testing Service (the authors are both affiliated with ETS), and its examples deal mostly with education evaluations.

The authors touch upon many aspects which affect objectivity, candor, and ethics, including the purposes of evaluation, the evaluation strategies, relationships of the various parties concerned in evaluation, bias from values and ideologies, reporting style,

responsibility for dissemination, rights of communities (much more pertinent than rights of individuals). They include a draft statement of ethical responsibilities (reproduced as Appendix D of this report).

The book also discusses other aspects of evaluation as a profession, such as methods, evidence, utilization training and trends. It is concise and readable. I suggest that PPC/E offer to buy the book for any interested evaluation officer.

Baratz, Stephen S. and Marvin, Keith E. "Resolving Privacy, Access and Other Problems in the Audit and Reanalysis of Social Research for Policy" Evaluation and Change Special Issue 1978 of the magazine Evaluation Pages 31-35.

Baratz is a consultant for the General Accounting Office and Marvin is an Associate Director of GAO. They report on a problem which arose when GAO attempted to review evaluations in order to advise Congress on the probable reliability of information obtained from the executive branch. Both GAO and social scientists were concerned that confidentiality is essential for obtaining candid responses and that the presence of auditors while data is being analyzed or the later reanalysis by auditors might threaten that confidentiality. This problem is further complicated by a possible conflict between the Privacy Act of 1974 and the 1974 amendments to the Freedom of Information Act.

It is not believed that confidentiality is a major issue for A.I.D. since few projects deal with individuals in the way many domestic programs on delinquency, drug abuse, mental health, welfare, and education do. Hence, few evaluations rely on in-

dividual interviews. However, if A.I.D. does ever have occasion to be concerned, it may want to consult two reports reviewed in this article: National Academy of Science, "Protecting Individual Privacy in Evaluation Research" Washington D.C. 1975 and Social Science Research Council Privacy Protection Study Commission "Personal Privacy in an Information Society," Washington D.C. 1977.

Brandle, John E. "Evaluation and Politics" Evaluation and Change, special issue 1978 of the magazine "Evaluation", pages 6-7.

Brandle is a professor of public affairs at the University of Minnesota and a member of the Minnesota House of Representatives. He argues that politicians deal with issues for which there are accommodations, not solutions. They perceive evaluators as another interest group who seek truth with rigor and explicitness but who present only part of reality-- if evaluations were complete and definitive, politicians would be unnecessary. Nevertheless, he urges public dissemination of evaluations to help keep politicians honest and also urges evaluators to present their case with skill and force. Finally, he counsels that evaluations can have the most impact if they look at experiments and demonstrations before political commitment is made and if they deal with mechanisms, organization and management on which interest groups lack strong commitments.

Campbell, Donald T. "Reforms as Experiments" Handbook of Evaluation Research Vol 1, pages 71-100, Beverly Hills, California: Sage Publications Inc.

This is the classic paper prepared under a National Science Foundation grant and presented at various social psychology conferences in 1968. The Handbook reprints it, with some revisions by the author, from the April 1969 issue of the American Psychologist. The author argues for use of the quasi-experimental approach to obtain objective evaluations of social reforms and also advises that human courtesy means that the most dependable means of assuring a favorable evaluation is to use testimonials from those who have had the treatment.

Cox, Gary B. "Managerial Style: Implications for the Utilization of Program Evaluation Information" Evaluation Quarterly Vol. 1 Number 3 August 1977 pages 499-508.

This article is based on a paper given at the American Psychological Association Annual Convention in 1976. The author is from the University of Washington. It is another example of efforts to make evaluators understand the organizational and political realities of management. It argues that written reports may be necessary for documentation but that managers prefer verbal communication as quicker and involving more exchange. It urges that evaluators' oral reporting be frequent, simple, and succinct, stressing salient points such as a change in trends. Evaluators should report unexpected or unpleasant findings immediately.

Daines, Samuel R. Evaluation of the Haiti Small Farmer Improvement Loan, Practical Concepts, Inc. (draft), February 1979.

This report is being prepared for and will be issued by A.I.D. It is the third annual evaluation of the project, each one by a different consultant. The first two reports concentrated on management of inputs and outputs, (the institutional and delivery system objectives) and made useful suggestions for improvement. Neither looked at impact but both apparently expressed some doubt about progress. The scope of work for the third evaluation returned to the original evaluation plan in the Capital Assistance Paper which called for a farmer survey. Daines was able to use a sample of new loan applicants as a control group to compare with another sample of loan recipients. He also constructed a sketchy baseline from data in loan files. Thus he could compare present status of loan recipients with previous status and with non-recipients. He demonstrates a significant rise in coffee production by small farmers and in their income and living standard which is attributable to the project. The general significance of this report for evaluation procedures is that it shows that

objective data can be obtained without unreasonable expense, that objectivity and hence lack of dissent is easier for impact evaluation than for process evaluation, that a useful baseline can sometimes be constructed after several years of project activity, and that control groups are fairly easy to obtain from new clients for a project.

Diener, Edward and Crandall, Rick Ethics in Social and Behavioral Research, Chicago: University of Chicago Press, 1978.

This new book is disappointing. In contrast with the one by Anderson and Ball, this seems to be directed at students rather than professionals and is rather simplistic. Moreover, despite a broad definition of social research as dealing with human functioning outside the biomedical area, the authors, who are psychologists, seldom get beyond a view of research as dealing with individuals rather than with social programs or trends. Hence the problems they cite do not often pertain program evaluation of interest to A.I.D. even when they speak of cross-cultural research or evaluation of "experimental interventions." They discuss three areas: 1) treatment of research participants or subjects--not harming them, obtaining consent, preserving privacy 2) professional issues such as honesty, accuracy, use of grants, and giving credit in publication, and 3) relationships among science, values, and society. In an appendix are selected portions of codes of ethics of three associations--anthropological, psychological, and sociological.

Dwellely, Hugh L., Uniform Procedures for Issuance of Work Orders Under Agency-Wide Indefinite Quantity Contracts, memorandum issued June 1, 1978 by SER/CM

This memorandum assigns responsibility for business management of all indefinite quantity contracts (IQCs) to two contracting officers. It also establishes coordinating technical offices responsible for helping ensure proper distribution of work orders among contractors and for identifying contractors suitable for the required services. There is a specific policy against A.I.D. suggesting use of a specific individual. The memorandum contains an outline for a project implementation order (PIO/T) which will be the basis for a request for a proposal (RFP) and a subsequent work order.

Evans, John "Summary of Testimony to Senate Committee on Human Resources" Evaluation and Change, special issue 1978 of Evaluation magazine. Page 17.

Evans is Assistant Commissioner in the Office of Education for Planning, Budgeting and Evaluation. He says that his office prepares designs of program evaluations studies in-house before contracting out the studies in order to assure that evaluations are methodologically sound and are relevant to policy needs. They also monitor evaluation projects closely and prepare short plain-language summaries for executives and Congress.

Freeman, H.E. The Present Status of Evaluation Research Paris: UNESCO (mimeo) 1976.

This is a thoughtful and practical monograph by a Professor of sociology and Director of the Institute of Social Science Research at UCLA. Unfortunately, it is for limited distribution and labeled not to be quoted without permission. Freeman argues strongly that evaluation is a political decision making tool and not ordinary research. Hence, the leadership comes from policy makers and the results are not judged by peer assessment of other scientists. Correspondingly, the major problems for utilization are not methodological (except as studies may take too long or cost too much) but organizational. Interestingly, he urges a biasing "of evaluation for results", meaning that evaluation should concentrate on targets most apt to change and stress strong rather than minor program interventions. Such evaluations would still follow the usual social science rules on data. He also advocates that the evaluator become involved with action to provide quicker and more effective feedback even though such participation may threaten independent objectivity. Finally, Freeman suggests that evaluations should pay

attention to the processes used to reach targets and not just progress toward targets, since he thinks a major explanation for lack of impact is incomplete implementation. Process evaluation can be objective by defining expected action elements and using samples, records and observations about such actions. But it can threaten privacy of both program staff and clients. Thus it could pose an ethical problem.

Gurel, Lee "The Human Side of Evaluating Human Services Programs: Problems and Prospects Handbook of Evaluation Research, Vol. 2 pages 11-28 Beverly Hills, California Sage 1975

Gurel is a psychiatrist. Parts of this paper were presented in his 1972 presidential address to the Division of Psychologists in Public Service of the American Psychological Association.

The author says that the evaluator often perceives what he was hired to do as research and loses sight of a critical distinction between research and evaluation, which must possess policy and decision-making implications. He goes on "those of us who are evaluators often act as though we had a moral imperative to search out truth and we not infrequently regard those who hinder our quest as malevolent or worse". Evaluators must explore the background and reasons for the evaluation and limit their inquiry to information needed. He considers but does not give answers to the differing views about publication--the manager may be satisfied with progress reports and a verbal report and resent the time to prepare for publication while the evaluator seeks recognition from a professional group. Moreover, the manager fears whistle blowing.

Guttentag, Marcia and Struening, Elmer L, editors  
Handbook of Evaluation Research, Vols. I and II,  
Beverly Hills, California. Sage Publications 1975.

This two volume work was sponsored by the Society for the Psychological Study of Social Issues, although some of the contributors have a background other than psychology. Guttentag was a Professor at the Harvard Graduate School of Education and first President of the Evaluation Research Society. Struening is a professor at Columbia and affiliated with the New York State Department of Mental Hygiene. Individual articles pertinent to this report have been listed separately in this bibliography. The volumes contain many other articles dealing with such topics as methodology and evaluation examples.

Guttentag, Marcia "Summary of Testimony to Senate Committee on Human Resources" Evaluation and Change, special issue 1978 of magazine Evaluation page 18.

After reviewing inadequacies of traditional research methods for evaluation, the author lists six responsibilities of evaluators. Two are pertinent for this report: Explicit statements on the certainty of information provided and a routine use of multiple sources of information to increase the certainty

Havens, H.S. "Summary of Testimony to Senate Committee on Human Resource", Evaluation and Change, special issue 1978 of magazine Evaluation, pages 11-12

The Director of the Program Analysis Division of the United States General Accounting Office speaks of problems of timeliness, relevance and credibility of evaluations. In relation to the latter he mentions GAO reanalyses discussed more fully by Baratz (see above).

Levine, Adeline and Levine, Murray "The Social Context of Evaluative Research: A Case Study" Evaluation Quarterly November 1977, pages 515-542.

This is a case study about evaluations of the Gary Indiana school system in 1914 and 1915 in connection with a proposal to adopt the Gary approach in New York City. The matter became a major political issue. The authors argue that the social context influences evaluative research considerably, that the difficulties are intricate. Evaluation research must be understood on its own terms, rather than as an extension of the research approach to an applied problem.

They dismiss as inadequate various suggestions by Campbell (1969, 1970, 1971) for achieving objectivity such as replicated experiments with the research divided between two contractors, critical reanalysis of data with basic data available to interested persons, minority reports and alternative analysis. These technical controls do not deal directly with political pressures.

They suggest prior identification of various interests so that an evaluation design may include various measures deriving from different perspectives.

They also suggest use of adversary methods, with public hearings and cross examinations.

Among the special issues identified in the case study were: financial and political independence of evaluators, professional ethics, evaluator's responsibility for other side of story and for consequences of a published report.

An Evaluation Scope of Work for P.L. 480 Title II  
Country Programs, Robert R. Nathan Associates, Inc.,  
1978.

This was prepared for A.I.D.'s Office of Food for Peace. It outlines a systems approach in which a Title II program will be related not only to its own targets but also to other programs and to the country nutrition, health, and economic situation. It anticipates unilateral evaluations by a three-person team over a period of three weeks plus report writing time. The team would make some site visits, conduct extensive interviews and use existing records. The report lists 10 possible kinds of impacts and suggests indicators for them, most of which would require tests or surveys not possible in a regular evaluation. It outlines a seven-chapter study and report and contains proposed interview questionnaires to use at MCH Centers, Schools, and Food-for Work projects.

Nichols, Donald, "Summary of Testimony to Senate Committee on Human Resources", Evaluation and Change, special issue 1978 of magazine Evaluation, pages 14-15.

The Deputy Assistant Secretary of Labor for Program Analysis and Policy Development describes the monitoring of evaluation contracts. Evaluation plans of departmental agencies are reviewed in a Research Clearinghouse. After evaluation contracts are let, Nichols' staff participates in site visits, seminars and meetings to monitor and give additional directions to the contractors. After an evaluation report is completed, Nichols' staff writes an assessment of its quality and appropriateness. It also publishes summaries for the department and Congress.

Office of Management and Budget, Evaluation Management, A Background Paper, May 1975.

This paper prepared primarily by James W. Morrison, Jr. of the Evaluation and Program Implementation Division, OMB argues that evaluation can be a useful management tool but seldom is, primarily because evaluation itself is not well managed. It urges that evaluation studies be decision-driven, that contracts for studies do a better job of specifying the expected use of findings and that they be monitored more carefully. The paper distinguishes between program evaluation ( a systematic analysis of progress toward objectives) and evaluation research (adherence to experimental design, not oriented to specific program decisions).

Rossi, Peter J. and Wright, Sonia R., "Evaluation Research: An Assessment of Theory, Practice and Politics" Evaluation Quarterly, Feb. 1977 Pages 5-52.

This is a revision of a paper presented at a UNESCO Conference on Evaluation Research at Washington in September 1976. The authors are members of the University of Massachusetts at Amherst. The authors point out that all research requires precise definition but that the precision for evaluation comes from the program managers who state the criteria for success and the target clients. Moreover, political or moral considerations may preclude random experiments or control groups. Nevertheless, they conclude that there are no differences, for the most part, in the logic, methods and techniques used in basic and evaluation research. The most important methodological difference is that evaluation should be able to detect small effects, since the effects of a program will usually be weaker than the proponents hoped. The paper then reviews various major evaluations.

Salasin, Susan, "Investment in People as a National Priority: An Interview with the Honorable Harrison A Williams, Jr." Evaluation and Change, special issue 1978 of magazine Evaluation, pages 19-23.

The chairman of the Senate Committee on Human Resources talks about his conclusions after hearings October 6 and 27, 1977 on evaluation:

"...I am not comfortable with legislation that directly regulates professional evaluators and their associations... Code of procedures and ethics is needed to guide the development and activities of evaluation personnel.... profession itself should move ahead.... so that its strength and growth are derived from its own experts and their universities and research organizations".

Scheiser, Mary Ann, "Program Participants' Positive Perceptions: Psychological Conflict of Interest in Social Program Evaluation", Evaluation Quarterly, Feb. 1978, pages 53-70.

This is a thorough article which shows that participants (both recipients and staff) think social programs are beneficial despite lack of evidence of progress toward stated program goals. The author cites social psychology theory and the theory of cognitive consistency to support her conclusion.

The positive bias goes far beyond Campbell's idea of courtesy (cited earlier). For instance, Gordon and Morse reviewed 93 published evaluations in 1975 (Annual Review of Sociology: Palo Alto, California: Annual Reviews, Inc.) and found that 52% of evaluations with poor methodology reported success while only 22% with adequate methodology found successful outcomes. Further, evaluators affiliated with the program were much more likely to report success (58%) than nonaffiliated researchers (14%).

The author points out that these findings cast doubt on various evaluation methodologies, such as

Campbell in 1974 saying that reports of staff, clients and families can add important qualitative information to more rigorous methods or Guttentag in 1975 and 1976 seeking a Bayesian combination of information bearing on all participants' values rather than behavioral changes or the Stanford Evaluation Consortium (1976) advocacy of the evaluator influencing the program as it proceeds.

The concludes with a dilemma--evaluation must be independent of the program agency but such independence can bring non-cooperation which precludes good research.

Sjoberg, Gideon, "Politics, Ethics and Evaluation Research", Handbook of Evaluation Research, Vol. 2, pages 29-54, Beverly Hills: California 1975.

The author argues that the ethics of evaluation are too narrow, that emphasis on the scientific method leaves out ethical questions arising from the fact research is a variable and that researchers tend to formulate issues and select indicators compatible with the established social order. Somewhat inconsistently, he says that the testing of hypotheses (the logico-deductive method) is akin to political liberalism. Another way in which goal-oriented evaluation is too narrow is that client-centered organizations require multiple and ambitious goals. An alternative approach which might be more "ethical" or unbiased would be the logic of dialectic in which the evaluator would use "countersystem" with different values. For example, a school might be judged by whether it helped students as well as by what they learned--the evaluator should look at the failures as well as the successes.

Weidman, Donald R., "Writing a Better RFP: Ten Hints for Obtaining More Successful Evaluation Studies", Public Administration Review, November/December 1977, pages 714-717.

The author is a program analyst in the Management Improvement and Evaluation Division of the Office of Management and Budget, with previous experience doing evaluation in the Urban Institute. His "do's" and "don'ts" are summarized in Chapter V.

Weiss, Carol H., "Evaluation Research in the Political Context", Handbook of Evaluation Research Vol. 1 pages 13-25, Beverly Hills, California: Sage Publication 1975.

The author heads the Bureau of Applied Social Research at Columbia University. This paper was presented at the 1973 annual meeting of the American Psychological Association. She pointed out that politics is implicit in the evaluation process, thus affecting its complete objectivity. Evaluation accepts the desirability and possibility of reaching stated goals. When evaluation accepts experimental variables, it conveys that other things in the situation of the target population are unimportant--looking at services, not social structure. Evaluation has a reformist bias, wanting to modify without drastic change.

She argues that both inside and independent evaluations have a role. The former is consonant with decision-makers goals and values but may stretch their sights a bit. The latter, especially when sponsored by higher echelons, has wider perspectives with more concern for ensuring that programs are worth their money.

A responsible evaluation puts goals in proper perspective, compares strong intensive programs with ordinary program levels, locates positive programs so they can be expanded, and keeps uncovering short comings.

Weiss, Carol H., Evaluation Research: Methods for Assessing Program Effectiveness, Englewood Cliffs, New Jersey: Prentice-Hall, Inc. 1972.

This brief and readable textbook is part of a series on methods of social science. Weiss defines evaluation research more broadly than OMB (see this bibliography). Her basic theme is that evaluation applies the many methods of social research in an action context that is intrinsically inhospitable to them. Thus much of the book is pertinent to a general exploration of how to achieve objectivity. While much of her advice can be classified under the heading of ethics, it is not so presented. Thus when speaking of evaluations commissioned for covert purposes other than information for better decisions, she says "the evaluator might well ponder whether he wishes to get involved in the situation or whether he can find more productive uses for his talents elsewhere." In another instance she remarks that when an evaluator makes recommendations, it "behooves" him to indicate which ones are supported by study data and which reflect his own knowledge or values. She does speak of dual obligations

to the sponsoring organization and to society and the profession, but implies that they have little conflict because the evaluator owes them all objectivity.

## Appendix C - CODES OF ETHICS

The scope of work called for a review of the codes of ethics of consultant organizations and of other key professional societies, to be summarized as appropriate.

In their chapter on "Ethical Responsibilities in Evaluation" Anderson and Ball note that useful advice can be found in standards of various professional groups and that it is to be expected that relevant standards from these groups will be met. But they believe that certain ethical questions are unique to program evaluation. With this caveat, codes of several associations are reviewed below.

### American Psychological Association

This organization has three sets of standards- for psychologists, for providers of psychological services, and for educational and psychological tests. Parts of the first have some pertinence for evaluation. They were revised on January 30, 1977. One of the recurring phrases is that psychologists "maintain the standards of their profession", which does not add clarity.

### Ethical Standards of Psychologists

Preamble- Psychologists respect the dignity and worth of the individual and honor the preserva-

tion and protection of fundamental human rights. They are committed to increasing knowledge of human behavior and of people's understanding of themselves and others and to the utilization of such knowledge and for the promotion of human welfare.... they make every effort to protect the welfare of.... any human being.... that may be the object of study. They use their skills only for purposes consistent with these values.... While demanding for themselves freedom of inquiry and communication, psychologists accept the responsibility this freedom requires: competence, objectivity....

#### Principle 1. Responsibility

....Psychologists value objectivity and integrity....

a.) As scientists, psychologists.... select, appropriate areas and methods.... They plan their research in ways to minimize the possibility that their findings will be misleading. They provide thorough discussions of the limitations of their data and alternative hypotheses, especially where their work touches on social policy.... In publishing reports.... they never suppress disconfirming data.

Psychologists clarify in advance with all appropriate persons or agencies the expectations for

sharing and utilizing research data. They avoid dual relationships which may limit objectivity, whether political or monetary....

b.) As employees of an institution.... psychologists attempt.... to moderate institutional pressures that may distort reports.... or impede their proper use.

c.) As members of governmental.... bodies, psychologists remain accountable as individuals to the highest standards of their profession.

d.).... e.).... f.)....

#### Principle 2. Competence

The maintenance of high standards of professional competence is a responsibility shared by all psychologists in the interest of the public and the profession....

a.).... b.).... c.).... d.).... e.)....

#### Principle 3. Moral and Legal Standards

a.).... b.).... c.):....

d.) As researchers, psychologists remain abreast of relevant federal.... regulations concerning.... research with human participants....

#### Principle 4. Public Statements

(Preamble and nine points about announcing availability of services).

### Principle 5. Confidentiality

(In the 1977 revision, the APA Council could not agree on a proposed draft and retained a 1975 version until a revision has been adopted).

Safeguarding information about an individual that has been obtained by the psychologist.... is a primary obligation....

a.) Information received in confidence is revealed only after most careful deliberation and when there is clear and imminent danger to an individual or to society, and then only to appropriate professional workers or public authorities.

b.).... c.).... d.)....

e.) Only after explicit permission has been granted is the identity of research subjects published. When data have been published without permission for identification, the psychologist .... adequately disguises sources.

### Principle 6. Welfare of the Consumer

Preamble and 5 points on clients.

### Principle 7. Professional Relationship

....Psychologists respect the prerogatives and obligations of the institutions or organizations with which they are associated.

a.).... b.).... c.).... d.)....

e.) In the pursuit of research, psychologists give sponsoring agencies, host institutions and publication channels the same respect and opportunity for giving informed consent that they accord to individual research participants....

f.).... g.)....

h.) Members of the Association cooperate with.... the Committee on Scientific and Professional Ethics and Conduct and the Committee on Professional Standards Review....

Principle 8. Utilization of Assessment Techniques

(Preamble and 5 points on testing).

Principle 9. Pursuit of Research Activities

.... Psychologists carry out their investigations with respect for the people who participate....

a.) In planning a study the investigator has the responsibility to make a careful evaluation of its ethical acceptability....

b.)....

c.) Ethical practice requires the investigator to inform the participant of all features of the research that might reasonably be expected to influence willingness to participate....

d.)....

e.) ....respect the individual's freedom to decline

to participant in or withdraw from research.

f.).... g.).... h.).... i.)....

j.) Information obtained about the individual research participants.... is confidential unless otherwise agreed in advance. When the possibility exists that others may obtain access to such information, this possibility, together with the plans for protecting confidentiality, should be explained to the participants as part of the procedure for obtaining informed consent.

k.).... l.)....

American Anthropological Association

Principles of Professional Responsibility

(adopted May, 1971)

Preamble

Anthropologists work in many parts of the world in close personal association with the peoples and situations they study. Their professional situation is, therefore, uniquely varied and complex. They are involved with their discipline.... their sponsors, their subjects, their own and host governments.... Misunderstandings, conflicts, and the necessity to make choices among conflicting values are bound to arise and to generate ethical dilemmas. It is a prime responsibility of anthropologists to anticipate

these and to plan to resolve them in such a way as to damage neither to those whom they study nor.... to their scholarly community....

1. Relations with those studied

In research, an anthropologist's paramount responsibility is to those he studies. Where there is a conflict of interest, these individuals must come first. The anthropologist must do everything in his power to protect their physical, social, and psychological welfare and to honor their dignity and privacy.

a.)....

b.) The aims of the investigation should be communicated as well as possible to the informant.

c.) Informants have a right to remain anonymous. This should be respected both where it has been promised explicitly and where no clear understanding to the contrary has been reached. These strictures apply to the collection of data by means of camera, tape recorders, and other data-gathering devices, as well as to data collected in face-to-face interviews, or in participant observation....

d.)....

e.) There is an obligation to reflect on the fore-

seeable repercussions of research and publication on the general population being studied.

f.) The anticipated consequences of research should be communicated as fully as possible to the individual and groups likely to be affected.

g.) In accordance with the Association's general position on clandestine and secret research, no reports should be provided to sponsors that are not also available to the general public, and, where practicable, to the population studied.

h.) Every effort should be exerted to cooperate with members of the host society in the planning and execution of research project.

i.) All of the above points should be acted upon in full recognition of the social and cultural pluralism of host societies and the consequent plurality of values, interests and demands.... This complicates choice-making in research, but ignoring it leads to irresponsible decisions.

## 2. Responsibility to the public

The anthropologist is also responsible to the public-- all presumed consumers of his professional efforts. To them he owes.... candor and truth....

a.)....

- b.) He should not knowingly falsify or color his findings.
- c.) In providing professional opinions, he is responsible not only for their content but also for .... explaining.... their bases.
- d.) .... anthropologists bear a positive responsibility to speak out publicly.... on what they know and believe as a result of their professional expertise.... to contribute to an "adequate definition of reality" upon which.... public policy may be based.
- e.)....

### 3. Responsibility to the Discipline

An anthropologist bears responsibility for the good reputation of his discipline and its practitioners.

- a.) He should undertake no secret research, or any research whose results cannot be fully derived and publicly reported.
- b.)....
- c.) He should attempt to maintain a level of integrity and rapport in the field such that.... he will not jeopardize future research there. The responsibility is not to analyze and report so as to offend no one, but to conduct research in a way consistent with

.... honesty, open inquiry, clear communication of sponsorship and research aims, and concern for the welfare and privacy of informants.

4. Responsibility to students

5. Responsibility to sponsors

..... reflect seriously upon the purposes of his sponsors.... require of the sponsor full disclosure of the.... aims of the institution and the research project, disposition of the research results. He must retain the right to make all ethical decisions in his research....

6. Responsibilities to one's own government and to host governments

.... be honest and candid. He should demand assurance that he will not be required to compromise his professional responsibilities and ethics....

American Sociological Association

Preamble

Sociological inquiry is often disturbing to many persons and groups. Its results may challenge long-established beliefs and lead to change in old taboos. In consequence such findings may create demands for the suppression or control of this inquiry or for a dilution of the findings. Similarly, the results

of sociological investigation may be of significant use to individuals in power....because such findings, suitably manipulated, may facilitate the misuse of power....

For these reasons, we affirm the autonomy of sociological inquiry. The sociologist must be responsive, first and foremost, to the truth of his investigation. Sociology must not be an instrument of any person or group who seeks to suppress or misuse knowledge....

At the same time this search for social truths must itself operate within constraints. Its limits arise when inquiry infringes on the rights of individuals to be treated as.... ends and not means....

Code of Ethics (adopted September 1, 1971)

1. Objectivity in Research

In his research, the sociologist must maintain scientific objectivity.

2. The sociologist should recognize his own limitations and, when appropriate, seek more expert assistance or decline to undertake research beyond his competence....

3. Respect of the Research Subject's Rights to Privacy and Dignity

Every person is entitled to the right of privacy and dignity of treatment....

4. Protection of Subjects from Personal Harm

5. Preservation of Confidentiality of Research Data

Confidential information provided by a research subject must be treated as such.... Even though research information is not a privileged communication under the law, the sociologist must, as far as possible, protect subjects and informants.... However, provided he respects the assurance he has given his subjects, the sociologist has no obligation to withhold information of misconduct of individuals or organizations.

If any informant or other subject should wish, however, he can formally release the researcher of a promise of confidentiality.

6. Presentation of Research Findings

The sociologist must present his findings honestly and without distortion. There should be no omission of data from a research report which might significantly modify the interpretation of findings.

7. Misuse of Research Role

The sociologist must not use his role as a cover

to obtain information for other than professional purposes.

8. Acknowledgment of Research Collaboration and Assistance

9. Disclosure of the Sources of Financial Support

The sociologist must report fully all sources of financial support in his research publications and any special relations to the sponsor that might affect the interpretation of the findings.

10. Distortion of Findings by Sponsor

The sociologist is obliged to clarify publicly any distortion by a sponsor or client of the findings of a research project in which he has participated.

11. Disassociation from Unethical Research Arrangements

The sociologist must not accept such grants, contracts, or research assignments as appear likely to require violation of the principles above, and must publicly terminate the work or formally disassociate himself from the research if he discovers such a violation and is unable to achieve its correction.

12. Interpretation of Ethical Principles

When the meaning and application of these principles are unclear, the sociologist should seek the

judgment of the relevant agency or committee designated by the American Sociological Association....

13. Applicability of Principles

In the conduct of research the principles enunciated above should apply to research in any area either within or outside the United States of America.

14. Interpretation and Enforcement of Ethical Principles

The Standing Committee on Professional Ethics, appointed by the Council of the Association, shall have primary responsibility for the enforcement and interpretation of the Ethical Code.

Appendix D - A DRAFT STATEMENT OF ETHICAL RESPONSIBILITIES

Anderson and Ball (see bibliography) in their excellent chapter 8 on "Ethical Responsibilities in Evaluation" include a draft statement which they hope will serve as a guide until a statement comes from the appropriate professional organizations.

Ethical Responsibilities in Program Evaluation

- | Evaluator to Commissioner of Evaluation (COE), Participants, Public, and Profession   | Commissioner of the Evaluation (COE), Participants, and Secondary Evaluator to Evaluator  |
|---|---|
| 1. To acquaint the potential COE with those values and orientations of the evaluator that may bear on the proposed evaluation effort. | COE: To provide the potential evaluator with as complete information as possible about the program (or proposed program), the COE's expectations for the evaluation, and the proposed conditions and resources for carrying it out. |
| 2. To work toward a contract or "agreement" with the COE that is ethically sound.   | COE: To work toward a contract or "agreement" with the evaluator that is ethically, legally, and professionally sound.  |
| 3. To refuse to perform work until such a contract or "agreement" is reached.   | COE: To refrain from insisting that work be performed before such an "agreement" is reached.  |
| 4. To fulfill the terms of the contract or "agreement" to the best of the evaluator's ability.  | COE: To cooperate with the evaluator and to fulfill to the best of the COE's ability any commitments or obligations called for in the contract or "agreement."  |

5. To acquaint the COE promptly with problems arising in fulfilling such terms and attempt to work out a solution.
 

COE: To acquaint the evaluator promptly with problems associated with the program that may affect the evaluation effort; to work with the evaluator in attempting to solve any mutual problems that arise
6. To adhere to relevant professional and legal standards and ethics in the conduct of the evaluation, including appropriate provisions for privacy and informed consent of participants and confidentiality of data.
 

COE: To support the evaluator's adherence to relevant professional and legal standards and ethics in the conduct of the evaluation.
7. To carry out data collection and other evaluation activities with as little interference as practicable with the operation of the program.
 

COE: To encourage full and honest cooperation by program participants in supplying data needed for the evaluation effort.  
Participants: To cooperate in the data-collection effort associated with the evaluation and to provide accurate information in response to legitimate requests.
8. To acquaint the COE with any aspects of program philosophy or operation that do not appear to be ethically sound or physically safe but are observed by the evaluator, even if such observation is not part of the evaluator's specific charge; in addition, to inform the appropriate authority if the evaluator obtains evidence of legal misconduct by the COE.
 

COE: To recognize the evaluator's "amicus" role in noting ethical, legal, safety, or professional problems associated with the program; to seriously consider the evaluator's observations in this area.

9. To acquaint the COE, in advance of any response, with requests received by the evaluator from superordinate agencies for information (testimony, and so forth) about the program or evaluation; to ascertain with the COE whether such requests are valid; if so, to acquaint the COE fully with the nature of the response.
- COE: To advise the evaluator on the validity of requests for information from superordinate agencies.
10. To present a "balanced" report of results to the COE in timely fashion and in a form usable to the COE; to spell out limitations of the investigation, along with the evaluator's values and orientations, that may bear on the conclusions.
- COE: To discourage misinterpretation and misuse of the evaluation results.
11. To reserve the right to publish rejoinders to any misinterpretation or misuse of the evaluation results by the COE.
12. To identify other groups with a legitimate concern for the results of the evaluation and to make the results available to them.
- COE: To advise the evaluator about groups that, to the COE's knowledge, have a legitimate interest in the results of the evaluation; to encourage dissemination of results to such groups.
13. To allow interested professionals to examine the data produced by the evaluation, within the limitations of accepted standards for privacy, confidentiality and informed consent related to the purposes for which the data were collected.
- Secondary evaluator: To specify, at the time when permission is sought to review the evaluation data, the purposes of the secondary evaluation effort; to maintain professional and ethical standards in conducting the secondary evaluation, including honoring any relevant commitments

to those who supplied the original data; to report in a professionally sound manner on the results of the secondary evaluation.

14. To reserve the right to publish rejoinders to any misinterpretations or misuse by the secondary evaluator of the original evaluation data or results.
15. To share with professional colleagues and relevant agencies and institutions knowledge about basic processes (educational, psychological, social, and so forth) derived from evaluation studies.

Definitions used in the presentation:

Program--institution, organization, activities, or materials with an interventionist function in improving human welfare.

Evaluator--person(s) or agency with major responsibility for planning, carrying out, and reporting evaluation activities. May be independent or dependent.

Commissioner of the Evaluation (COE)--person(s) or agency with major responsibility for securing the services of an evaluator.

Participants--administrators, staff, program recipients, and other persons with a role in the program being evaluated.

Secondary evaluator--person(s) or agency engaging in critical review of evaluation activities. May include reanalysis of previously collected data.

## Appendix E - INSTRUCTIONS ON SCOPE OF WORK

Most A.I.D. evaluations by contractors are performed under the device known as an Indefinite Quantity Contract (IQC). A work order is issued to a contractor who had already qualified in a general competition, thus saving the time for a new competition and negotiation before each evaluation. The procedures for such work orders are set forth in a June 1, 1978 memorandum by Hugh L. Dwelley, Director of the Office of Contract Management. Below is part of Attachment A to that memorandum giving guidance to requesting offices about information to be included in the Project Implementation Order (PIO/T) which will be used both for a Request for Proposal (RFP) to the contractor and then for the Work Order.

**Background:** (This paragraph should contain enough information to enable the proposed contractor to understand the context in which the work is required and will be undertaken. This paragraph will appear in the Request for Proposal, but not in the Work Order).

- I. **TITLE** (The title is used for identification purposes only. It will appear in the request for proposal and the work order)
- II. **OBJECTIVE** (It is important that the objective be for the work order, not a larger project of which the work order is a part. When the contractor is to supply one member of a team, the objective of the work order would be something like-- "To provide technical guidance in the field of \_\_\_\_\_ to a team which shall \_\_\_\_\_".)

- III. STATEMENT OF WORK (Often it is useful to use "general" and "specific" headings. Do not include in the statement of work:  
--background  
--what support, assistance etc. is to be furnished to the contractor  
--the submission of reports and other written items prepared by the contractor.)

The statement of work will appear in the Request for Proposal and in the Work Order.)

- IV. REPORTS (The report section should cover all items to be submitted by the contractor. It should indicate if A.I.D. requires submission of a draft. It should indicate for finals (and drafts if appropriate) when, to whom, and in how many copies each item should be submitted.)

- V. RELATIONSHIPS AND RESPONSIBILITIES (You should be careful not to split the technical directions function. The contractor can not comply with two, often conflicting sets of directions.)

- VI. TERM OF PERFORMANCE (Straightforward. Will appear in the request for proposal and the work order.

- VII. LEVEL OF EFFORT (The technical office must project the number of days services required and the fixed daily rate. This item will not appear in the request for proposal. It will be used to judge the reasonableness of the level of effort proposed by the contractor. The work order will contain the level of effort negotiated.)

- VIII. AID ILLUSTRATIVE BUDGET (This is the Government's independent estimate. It must contain a detailed

breakdown of the budget. It will not be included in the request for proposals but will again be utilized to judge the reasonableness of the level of effort proposed by the contractor. Work order will contain the negotiated budget.)

IX. etc. Include the following, as appropriate.

DUTY POST

LANGUAGE REQUIREMENTS (Will assume none if not stated)

ACCESS TO CLASSIFIED INFORMATION (Will assume none if not stated differently)

LOGISTICS SUPPORT Should detail all support to be given to the contractor, including material to review, in kind services to be provided by Missions and host governments, etc.

If the host government is to provide services or goods, the language should be precise. For example--differentiate between "will provide" and "will provide, if available".

## Appendix F - DEVELOPMENT SUPPORT BUREAU EVALUATION GUIDANCE

### #1: Preparation of Scope-of-Work for Team Project Evaluation

The scope-of-work prepared for a Team evaluation effort is the basis on which the DSB agrees to expend funds and its resources. It is essential that these documents be of high quality reflecting the high professional standard expected of this Bureau. A good scope-of-work will normally result in sound recommendations emanating from the evaluation effort and provide guidance for the Bureau to make management decisions on a project.

In developing the scope-of-work, the project manager may refer to the Cluster Senior Program Officer and the Bureau Evaluation Officer for assistance. Scope-of-work for Team evaluation of research projects should be developed in conjunction with DS/RES.

A scope-of-work is also required for a team evaluation to be conducted under an IQC arrangement (PIO/T) or under other contractual mechanism (letter to contractor).

The Office Directors are advised that a draft scope-of-work must be submitted to DS/PO and to Regional Bureau if project is country specific for clearance at least 60 days prior to the scheduled date of the evaluation. This will permit other interested offices to make inputs and allow participants in the review including the contractor/grantee to make orderly arrangements.

The cleared scope-of-work is forwarded to the perspective cluster DAA for approval with signed copies to the AA/DA and DS/PO.

To assist the project manager in the preparation of the scope-of-work, the following guidelines are provided:

#### Outline for scope-of-work

- A. Title and number of Project to be reviewed.

- B. Name of contractor/grantee or PASA implementing organizations.
- C. Purpose and rationale for having team evaluation.
- D. Composition of team - provide names, organizational affiliations, brief curriculum vita, rationale for their selection, and expected inputs to the evaluation. Identify the team leader. If a team member has not been identified at the time of the submission of the scope-of-work, indicate the qualifications of the individual and the Bureau from which such services may be required. (see PG Evaluation #2 for criteria in selecting team members.)
- E. Other participants/observers expected at the evaluation- provide names or organizational affiliations, purpose for attending.
- F. Dates and places of evaluation.
- G. Cost Analysis for the evaluation - analysis should provide breakdown of travel, per diem, consultant services costs and any other costs incidental to the evaluation e.g. clerical service, printing, etc.
- H. Project Background - A brief paragraph on the history of the project indicating starting date, and dates of previous regular and team evaluations conducted on the project.
- I. Measurement of Progress to Date - One of the purposes of an evaluation by a team of experts is to determine progress achieved by a contractor or grantee toward stated purpose and goal. The basic document to measure such progress is the Project Paper or grant statement when appropriate and in particular the Log Frame Matrix (LFM). Before setting forth the methodology for the team to address progress to date, the project manager should review the LFM to ensure that no changes have occurred since the project was initiated. If changes have occurred the LFM should be changed to reflect current project situation.

Note: Using the LFM (or revised LFM), the project manager will reflect in this section where the means of verification from input to goal levels are and how they can be made available.

J. Problems and Issues to be addressed by the Team - This is the most important element in the scope-of-work. It should reflect problems and issues which are project specific as they relate to:

- (1) adequacy of project design (matrix),
- (2) staffing,
- (3) management,
- (4) reporting,
- (5) performance of contractor/grantee toward achieving the target of the project, at all levels - output, purpose, goal,
- (6) adequacy of project strategy,
- (7) resource inputs,
- (8) implementation plan and implementation management,
- (9) statue of critical assumptions,
- (10) impact of unplanned events,
- (11) current or projected utilization of results and impact on A.I.D. target group,
- (12) recommendations for changes in project design or implementation.

In developing the problems and issues, the project manager should review the following documents but not be limited to such documents.

- a. Administrative history of project,
- b. Project Paper including matrix,
- c. PAF,

- d. R & DC and RAC comments,
  - e. Contract, Grant and/or PASA document,
  - f. Annual and other progress reports,
  - g. Reports on site visits,
  - h. Previous evaluation reports with particular reference to recommendations made and action taken,
  - i. Publications resulting from the project,
  - j. Related documentation such as correspondence, new items etc.
- K. List documents to be attached to the scope-of-work for team members e.g. project papers, previous evaluation reports, annual reports, correspondence, field reports and others as listed in Par. J above.
- L. Agenda for the review - Give dates and time for actions to occur starting with distribution of documents to the team members and terminate when the evaluation report is submitted to the Director of the sponsoring technical office.

**#2: Criteria for the Selection of Evaluation Team Members**

The size of the team should be dictated by the complexity of the project, problems and issues identified, and magnitude of the activity. The work scopes should spell out what special skills that are required for execution of a successful evaluation, and the curriculum vitae of the proposed consultants should verify that they have these skills. It is conceivable that an in-depth evaluation of one project may require only one person, and yet, on the other hand, it is very possible that several persons will be required to address the questions outlined in the work scope for another project.

The project manager, with the assistance and/or clearance of DS/PO, has the primary responsibility for

nominating members to the evaluation team. The selection and approval of the team rests with the DAA of the Cluster. However, to encourage impartial reviews, it is suggested that the initiating Technical Office submit a request to the interested Bureau offices or DSB offices providing credentials and qualifications required of an individual to serve on the evaluation team. The respective Office Directors will select and nominate a qualified member of their staff to serve as team member.

Although the project manager does not necessarily have to serve as member of the team, he must participate in all review sessions.

Since evaluations are concerned with problems of project management as well as technical matters, where it is appropriate, the team should include a person with program and management skills. The project manager is also encouraged to use whenever possible the services of outside experts, regional bureau technical specialists, and country desk personnel as members of the team. It is the Agency policy that members of the evaluation team be familiar with A.I.D. project design and evaluation methods. In selecting team members to evaluate research projects, it is urged that at least one member represent RAC and/or DS/PO/RES to be selected by DS/RES which is the Agency sole contact point with RAC.

If the team members have not been approved at the time the draft scope-of-work for the evaluation is prepared (See PG Evaluation #3), indicate the Office or Bureau to participate and credentials and qualification of type of individual desired to serve on the team.

## Appendix G - LEGISLATION ON EVALUATION

The Foreign Assistance Act had a new Section 125 on evaluation added last fall (P.L. 95-424, Oct. 6, 1978):

"Sec. 125 PROJECT AND PROGRAM EVALUATION--  
(a) The Administrator of the agency primarily responsible for administering this part is directed to improve the assessment and evaluation of the programs and projects carried out by that agency under this chapter. The Administrator shall consult with the appropriate committees of the Congress in establishing standards for such evaluations."

\* \* \* \*

The House Appropriations Committee had the following to say about evaluation (page 17 of Report 1195-1250 "Foreign Assistance and Related Programs Appropriation Bill 1979):

### EVALUATION

The Committee emphasizes the importance of a strong A.I.D. evaluation program which:

- provides the A.I.D. Administrator with the information he needs to continue to improve the quality and effectiveness of A.I.D. programs;
- insures that field Missions and central bureaus learn the lessons of A.I.D.'s successes and failures;
- analyzes the impact of technologies used in A.I.D. projects on the poor;
- identifies the developing countries that evidence the greatest commitment to helping their poor people;
- is viewed as a prestige assignment in the agency;
- recommends improvements in A.I.D.'s project design and management system (such as the gathering of baseline data) to insure the effective evaluation of programs; and
- recommends to the Administrator any revisions in

programs and policies indicated by the results of evaluations.

Through fiscal year 1978, A.I.D. has obligated over \$260 million for programs in rural electrification with \$78 million in loans and grants planned for fiscal year 1979. Given this sizable investment, it is recommended that A.I.D.'s evaluation office undertake an assessment of rural electrification efforts to determine their effect on the poor.

\* \* \* \*

The Privacy Act of 1974 (P.L. 93-579, 5USC 552a) has requirements about handling data on individuals which can affect raw files on evaluation interviews and observations. It has also served to reinforce the ethical standards about obtaining informed consent. Perhaps its major impact is to make it more difficult for GAO or scholars to re-analyze evaluative data. Key provisions are summarized below:

- Public disclosure that an agency maintains a system of records about individuals
- Strictly enforceable procedures for ensuring that individuals have access to their records and opportunity to correct them.
- Controls on inter-agency transfer of individual's identifiable data.
- Safeguards to prevent unauthorized access to data and penalties for violation.

\* \* \* \*

The Freedom of Information Act (P.L. 93-502) requires adequate procedures to ensure that the disclosure of evaluation findings is responsibly accomplished. To ensure the availability of evaluation reports, there are implications concerning indexing, storage, retrieval, and access control. Another implication is that a reporter could request and receive unpublished dissents to evaluation reports.

\* \* \* \*

The Office of Management and Budget circulated a draft circular last summer entitled "Management Improvement and the Use of Evaluation in the Executive Branch". Comments have been received and issuance is pending. This circular would supersede one issued in 1972 on the management review and improvement program and another issued in 1976 on preparation of management plans. It sets a policy that all agencies will assess the effectiveness and efficiency of their programs and seek improvement on a continuing basis. It defines program evaluation as a formal assessment, through objective measurements and systematic analyses, of the manner and extent to which programs achieve their objectives or produce other significant effects, used to assist management and policy

decision making. Evaluations should focus on substantive program operations and results. The circular will require an annual report on management improvement activities, including a listing of officials responsible for evaluation and resources devoted to evaluation.

Appendix H - PEOPLE INTERVIEWED

Agency for International Development

Office of Auditor General  
Linck, Eugene G.  
Stevens, Benjarnin

Office of Contract Management  
Dilts, Russell  
Snyder, Michael

Evaluation Officers  
Elliott, Vance L. (Asia)  
Fox, Edwin K (FFP)  
Goldstein, Bernice (LA)  
Molfetto, Charles (DSB)  
Saenz, Paul (AFRICA)  
Silver, Joan (NE)

Office of Management and Budget

Division of Management Improvement and Evaluation  
Taylor, Oliver.  
Weidman, Donald R.

Consulting Firms

American Institutes of Research  
Schwartz, Paul

American Technical Assistance Corp.  
Brown, Albert E.

Development Alternatives, Inc.  
Micklewaits, Donald

Practical Concepts, Inc.  
Rosenberg, Leon

Appendix I - RECOMMENDATIONS

1. Revive efforts to enhance understanding of the Agency's project design and evaluation system. To this end, the Office of Evaluation in PPC, in cooperation with Bureau Evaluation Officers, should:

(a) Update the existing cleared draft chapter on project evaluation and publish it without waiting for other chapters in the proposed Evaluation Handbook.

(b) Brief new Mission Directors on design and evaluation or encourage them to take the seminar, either in Washington or overseas.

(c) Review with the Professional Studies and Career Development Division (PSCD) a practical and acceptable way to expose interns to more understanding about evaluation.

(d) Encourage PSCD to carry through on arrangements for more overseas training for American and host government officials about project design and evaluation.

(e) All geographic bureaus should review and comment on evaluation reports from Missions, at least on a sample basis.

(f) Distribute samples or excerpts of good evaluation reports and scopes of work.

2. Require collaborative evaluations and encourage formal review sessions.

To this end, the Office of Evaluation in PPC, in cooperation with Bureau Evaluation Officers, should:

(a) Prepare a directive and airgram for issuance by the Administrator.

(b) Call attention to the advantages of formal Evaluation Review meetings for both field and AID/W projects. (The methodology part of

the Project Evaluation Summary can be used to report attendance at such Reviews).

(c) Consider whether host signatures on Project Evaluation Summaries should be required by other Bureaus besides Africa.

3. Improve the quality of scopes of work for evaluation contracts and work orders. To the end, the Office of Evaluation in PPC, in cooperation with Bureau Evaluation Officers, should:

(a) Supplement the general instructions issued by O/CM for Indefinite Quantity Work Orders with guidance about scopes for evaluation which would:

(i) Urge that the number of issues be limited and that impact, even if only interim impact, be stressed.

(ii) Require a description of methodology -- where the data were obtained and what validity they have -- and a designation of which conclusions are based on evidence.

(iii) Require the evaluator to contact host officials for opinions.

(iv) Suggest consideration of whether multi-disciplinary teams should be asked to present the views of each member as well as a team recommendation.

(b) Institute a review of scopes of work on a trial basis in the Bureaus where Evaluation Officers do not now review them, to see whether improvements in draft scopes justify the effort and possible delay.

(c) Continue to urge project managers to take the short course on contracting, which will help them for all types of contracts as well as those for evaluation.

4. Increase the resources available for evaluation. To this end, the Office of Evaluation in PPC, in cooperation with Bureau Evaluation Officers, should:

(a) Insure that the policy on inclusion of evaluation money in future project funding is adequately followed.

(b) Arrange for the allocation of more to Program Development and Support Funds for fiscal 1980.

(c) Ease restrictions on travel funds for evaluation, either by arranging a re-allocation of administrative appropriations for travel or by arranging to use Program Development and Support Funds for travel as well as for evaluation contracts.

(d) Bureau Evaluation Officers should review Office and Mission Evaluation Schedules to:

(i) Be sure evaluations are related to project phases and pending decisions rather than being annual.

(ii) See if evaluations of related projects can be combined or terminal evaluations done in conjunction with appraisals of new projects.

Evaluation Officers should enlist the help of Desk and Development Resource Officers in this review.

5. Promote a new emphasis in evaluation evidence about benefits for people To this end, the Office of Evaluation in PPC, in cooperation with Bureau Evaluation Officers, should:

(a) Announce the new emphasis in a circular from the Administrator which mentions the following actions that will be undertaken.

v-12-11-11

(b) Review project Evaluation Plans received as part of Project Papers to be sure this element is covered.

(c) Review Project Evaluation Summaries and call attention to Missions and Offices which are giving only opinions rather than data about the item of Benefits for People.

(d) Stress benefits for people in the section of scopes of work calling for evidence about project impact.

(e) Recommend more frequent use of sample surveys of target populations. (Precautions may be needed to protect confidentiality for respondents).

(f) Stress benefits for people in comparative evaluations initiated and conducted by AID/W bureaus.