

PD - AAR - 159 39644

TRAINING MANUAL  
FOR THE  
META-EVALUATION PROJECT

Contract No. OTR-0000-C-00-3482-00

Conducted for PPC/E/PES

November 1984

**TRITON** =====

=====

November 9, 1984

Ms. Nena Vreeland  
Technical Officer  
PPC/CDIE  
Room 3 659, New State  
Agency for International  
Development  
Washington, D.C. 20523

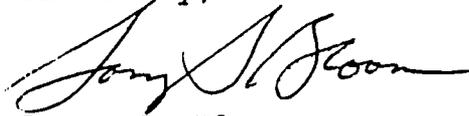
Dear Ms. Vreeland:

TRITON Corporation is pleased to submit the Training Manual of the Metaevaluation project conducted under contract OTR-0000-C-00-3482-00. This Training Manual represents a methodology for training reviewers in the applications of the scoring instrument, as well as detailed instructions concerning scoring procedures and utilization of scores.

Appendices to the Training Manual include copies of all forms used and a suggested reading list for training reviewers of evaluation reports.

Thank you again for your cooperation and assistance in developing this Training Manual.

Sincerely,



Sonny S. Bloom  
Vice President

**TRITON**

---

1255 Twenty-third Street, N.W. Suite 275, Washington, D.C. 20037 • (202) 296-9610

TRAINING MANUAL

for the

META-EVALUATION PROJECT

(Contract No. OTR-0000-C-00-3482-00)

TABLE OF CONTENTS

	<u>Page</u>
I. Introduction to the Project.....	I-1
II. The Logical Framework.....	II-1
III. Methodology.....	III-1
IV. How to Use the Methodology Scoring Instrument.....	VI-1
V. Scoring Process.....	V-1
VI. Utilization of Scores.....	VI-1

APPENDICES

2

## I. INTRODUCTION TO THE PROJECT

During FY 1982, TRITON conducted a "metaevaluation" project to assess the quality of Agency for International Development evaluation reports, under the auspices of AID's Program Evaluation Systems Division. A scoring instrument was designed to provide AID with a diagnostic tool to support its work in monitoring the Agency's evaluation system. The evaluation reports were read and then rated using the instrument developed by TRITON. The purpose of this document is to provide a training manual on how to score AID evaluations using that instrument.

The ultimate use of the instrument is to build up a data base derived from the routine review and scoring of AID evaluation reports which will aid in determining the strengths and weaknesses of the reports, based on sector, geographic focus, and other specific aspects of an evaluation. The instrument can be used on all types of AID evaluation reports, including mid-term evaluations, end-of-project evaluations and impact studies.

The instrument is based on a series of key issues concerned with quality and completeness for AID evaluation reports. These issues were identified through the following steps. First, TRITON staff developed a list of factors they believed should be found in evaluation reports. Concurrent with this effort, the Program Evaluation Systems Division (primarily Ms. Molly Hageboeck) prepared a list of characteristics found in "good" AID evaluations. Utilizing its own list and the PES list, TRITON integrated these inputs into a working list of attributes of "good" evaluations. The presence of these attributes was intended to indicate that a report was complete in its presentation and was of desirable "quality".

The next two stages of the project involved both a review and synthesis of evaluation literature from within AID and outside the Agency, and a series of interviews. These interviews were conducted by telephone and in-person with AID personnel associated with evaluations, individuals from other relevant agencies, and academicians. Interviews were conducted with personnel of institutions such as the World Bank, ACTION, the Inter-American Foundation, and the American Council on Volunteer Agencies for Foreign Services.

TRITON then prepared a report combining the results of these three stages in order to identify attributes of a "high quality" evaluation. This document, "Compilation of Attributes for Potential Use in Scoring Evaluation Reports," (submitted October 14, 1981), described in detail both the literature reviewed and interviews conducted. This compilation of attributes served as the basis for developing a scoring system for AID evaluation reports.

The Program Evaluation Systems Division performed a content analysis of the categories identified by TRITON. The purpose of this analysis was to identify the major quality and completeness factors found in the TRITON data and to segregate a number of subfactors found with each major category.

TRITON, in conjunction with PES, next used this master list to prepare a set of factors identified as relevant to determining the quality and completeness of an evaluation report, as distinct from the evaluation itself. Nine (9) factors which could be measured solely by reviewing the evaluation report were isolated. Three (3) additional factors from the master list were characterized as factors which could not be measured exclusively from the evaluation report itself, and were deleted during this part of the research effort.

Two concurrent activities were then undertaken. First, an iterative process was conducted between TRITON and PES staff to refine the factors and subfactor statements, to eliminate duplication and to coalesce all relevant attributes within the same factor. Next, all resulting statements were ranked in order of relative importance. To accomplish this objectively, TRITON, in conjunction with AID/PES, identified individuals both within AID and outside of the Agency who were then provided with a questionnaire for ranking the factors. Each participant was asked to rank order all statements contained on each form. Once these responses were obtained, TRITON utilized this data to develop the numerical weighting and scoring process.

The next step was to organize and structure the first draft of the scoring instrument itself. The draft of the scoring instrument was applied to five (5) USAID evaluation reports. The test results indicated that the nucleus of a useful, meaningful instrument had been developed, but that further refinement was necessary to clarify concepts, reduce application time, minimize differences in interpretation and eliminate any potential learning curve bias.

In order to improve the effectiveness of the instrument, USAID and TRITON staff met to ascertain the weak points of the first draft. The instrument was revised based on these meetings and a series of re-tests. USAID staff then selected 40 evaluation reports to be scored using this new version. A review of this round of scores indicated high rater consistency and inter-rater reliability, with a pattern of scores in a normal distribution.

In this final version, nine (9) characteristics--six (6) of which were further broken down into sub-characteristics--of a "good" evaluation were identified. Each evaluation report is rated for each characteristic/subcharacteristic on a scale of 0-4 (low-high), with a "not applicable" possibility for some

subcharacteristics. These scores are then summed, weighted and normalized on a 0-100 scale.

Revisions were made to this version of the instrument in 1984 to include more qualitative measures with regards to scoring the components of the logical framework. The newest version was then applied to the FY 83 evaluations, resulting in the conclusions and recommendations found in TRITON's 1984 metaevaluation report.

## II. THE LOGICAL FRAMEWORK

The "Logical Framework" or "logframe" is utilized in AID's design process as a tool for visualizing a project or program in terms of its inputs, outputs, purpose and goals(s). The logframes presented in Exhibit II-1 and II-2 are examples of formats used by AID and of the particular application developed by TRITON as part of the scoring process for this project.

The application of the logframe in the metaevaluation project is central to the scoring process. The logframe at the time of the planning of the project serves as an important mechanism for evaluation, since AID evaluations are concerned with the results produced by projects described via the logframe. The standard Project Evaluation Summary (PES) used by AID calls for assessment of the various elements of a project as per the logframe, and the logframe also serves as a useful frame of reference for evaluators not using that summary format. Persons scoring evaluations should, therefore, be familiar with the methodology, since its generation of a logframe assures that the coder/reviewer has an understanding of the project being evaluated. In addition, each element of the logframe is examined separately in the scoring instrument with regards to its description and analysis in the evaluation.

It is rare that a project's logframe can be found in its entirety in an evaluation report. Thus, for the purpose of scoring an evaluation report, the reviewer will usually have to write out a logframe for the project. The difficulty of this procedure varies according to the format and completeness of each evaluation report. Some will be primarily unilinear (that is, the progression from inputs to goals is composed of only one activity, such as education) while others may be multilinear and incorporate numerous diverse activities, as in an integrated approach containing health, agriculture and education components, which will each have distinct outputs and purposes, leading to a single goal.

A step by step explanation is presented below to provide guidelines for generating a logframe from a project evaluation:

1. Inputs are the resources made available for project implementation. Some examples are technical assistance, grants, loans, commodities and training.
2. Outputs are concrete accomplishments which are direct results of the transformation of the inputs. Some examples from a wide spectrum of technical activities are schools built, curriculum revised, participants trained, clinics staffed with trained personnel, written material produced and distributed, seeds planted, cooperative groups formed, credit extended to small farmers.

EXHIBIT II-1  
**PROJECT DESIGN SUMMARY**  
**LOGICAL FRAMEWORK**

Life of Project:  
 From FY \_\_\_\_\_ to FY \_\_\_\_\_  
 Total U.S. Funding \_\_\_\_\_  
 Date Prepared: \_\_\_\_\_

Project Title & Number: \_\_\_\_\_

NARRATIVE SUMMARY	OBJECTIVELY VERIFIABLE INDICATORS	MEANS OF VERIFICATION	IMPORTANT ASSUMPTIONS
Program or Sector Goal: The broader objective to which this project contributes:	Measures of Goal Achievement:		Assumptions for achieving goal targets:
Project Purpose:	Conditions that will indicate purpose has been achieved: End of project status.		Assumptions for achieving purpose:
Outputs:	Magnitude of Outputs:		Assumptions for achieving outputs:
Inputs:	Implementation Target (Type and Quantity)		Assumptions for providing inputs:

✓

EXHIBIT II-2  
ATTACHMENT 2  
LOGICAL FRAMEWORK

E-4

Hd

E-3

Hc

E-2

Hb

E-1

Ha

OUTPUTS

MANAGEMENT  
TRANSFORMATION

INPUTS

A-E3

A-E2

A-E1

A-0

3. Purpose is the mechanism which ties the concrete outputs into a directly linked, but more abstract, concept. Outputs are planned in order to accomplish a certain purpose. This often is not the ultimate goal of the project, but is the logical step to follow outputs as they are transformed from concrete things to actionable programs. There may be more than one purpose in a logframe, but they are usually complementary parts. Some examples of purposes are presented below:
  - a. Improve small holder access to agriculture institutions, services, and infrastructure, including inputs, credit, knowledge/extension to apply inputs, markets and/or storage facilities.
  - b. Expand the country's agricultural research capacity to alleviate the technological constraints faced by small traditional and agrarian reform farm facilities.
  - c. Provide a permanent comprehensive non-formal educational program for poor campesinos in the areas of agriculture, health, nutrition, family planning, literacy, appropriate technology, cooperatives and youth development.
  - d. Make foreign exchange capital available for use by the country's industrial sector.
  - e. Increase the capability of the National Marketing Board to store and market millet.
  - f. Reduce the incidence of communicable diseases among children below the age of six years.
4. The ultimate objective of the project is termed its goal. It may be necessary to include one or more subgoals before the goal can be reached. These are the least concrete, most abstract, and, in many cases, the most generic part of the logframe. Some examples are listed below:
  - a. Provision of an improved living environment for low-income families.
  - b. Reduced population growth rate.
  - c. Supply rural areas with adequate firewood and other forestry products.
  - d. Achieve self-sufficiency in food production.

- e. Increase planning and management capabilities within the ministries to effectively carry out development.
5. Another element of the logframe comprises failed assumptions. These may refer to outputs, purpose and/or goals, and are factors which were not considered to be detrimental to the project during its design, or were not considered at all until it was too late to prevent them from causing problems or blocking the attainment of outputs. Some examples at the output level follow:
- a. Output Level
    - o Normal rainfall.
    - o Credit to small farmers will be used for farm improvements.
    - o Teachers will be willing to exert extra energy to implement curriculum reforms.
    - o No outbreak of disease.
    - o Construction design is appropriate for the local climate and conditions.
    - o Equipment can be maintained by the community once it has been installed.
6. The final element of the logframe used concerns unexpected effects. These are any significant results of project activities that were not anticipated; that is, they were not part of the project design. While many evaluations will not contain unexpected results, those that do may report either beneficial or harmful effects. Some examples are:
- o Small businesses have started up around the project site.
  - o Dissemination of family planning information has extended beyond the targeted villages through family members and friends who live outside the area.
  - o Soya production has been stimulated.
  - o Housing is now available to low-income groups through rentals.

- o Spending several hours each day working on the school farm caused many teachers to miss more classes than usual.

### III. METHODOLOGY

#### A. OVERVIEW

Development of the scoring and analysis procedure resulted in a six-part scoring instrument which consists of:

1. The "Facesheet:" which contains specific information about the project (the list of "findings" prepared by the reviewer is attached to this form);
2. "Findings:" short, concise sentences referring to conclusions and recommendations found in the evaluation;
3. The logical framework;
4. Attachment 1: a series of statements pertaining to evaluation quality, which the reviewer scores according to the extent to which they are true for the evaluation;
5. Attachment 3: which reflects how well a project's various components e.g., inputs, outputs, were assessed; and
6. Attachment 5: which records the quality of the evaluation of the management transformation and the hypotheses. It examines the degree to which the transformation from level to level were evaluated and how well.

Both internal and external factors are recorded on the scoring instrument. The internal variables (Attachments 1, 3 and 5) are utilized to score the evaluation for quality and completeness, while the external variables (Facesheet and part of Attachment 5) are used to analyze scoring trends and patterns. Internal variables assess completeness, clarity, appropriateness, validity, replicability, reliability, adequacy, and bias. External variables taken into account are: geographic bureau, type of evaluation, timing of evaluation, AID management unit, technical code, length of time for evaluation, host country participation, levels of logframe evaluated, evaluation cost, project cost and contractor/evaluation entity. In essence, the internal variables assess the evaluation report as a self-contained entity, while external variables serve to situate the evaluation report in a particular environmental, cultural and developmental contexts.

#### B. RECENT CHANGES TO INSTRUMENT

A number of changes in the instrument used for the 1983 metaevaluation project were made in the facesheet data and the scoring attachments for the 1984 metaevaluation. The number of facesheet questions was increased from seven (7) to eighteen (18) to include more specific information about the evaluation.

Slight changes were made to Attachment 1, consisting of the addition of an "N/A" (not applicable) option to four of the subfactors and disaggregating two of the statements.

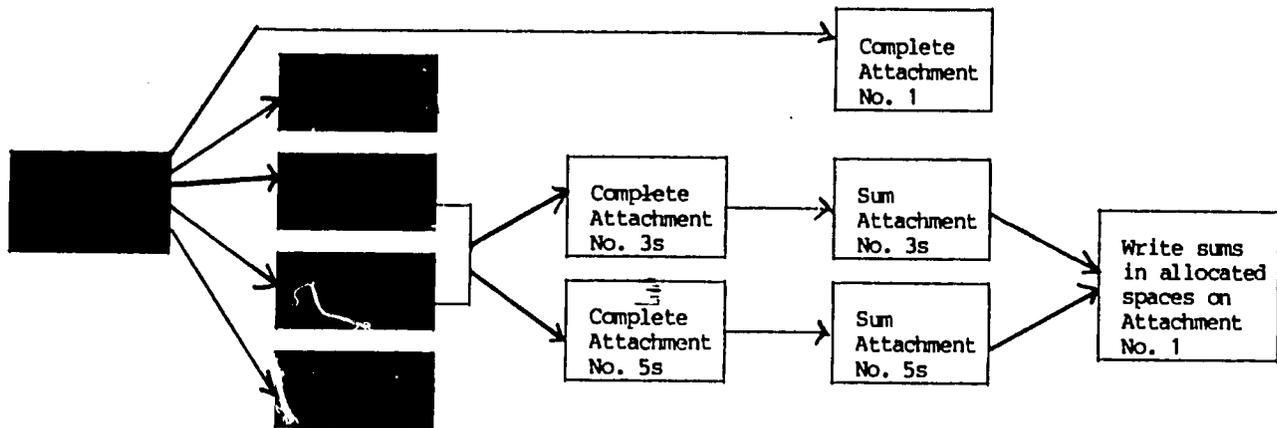
Attachment 3 was revised to accommodate six indicators, as opposed to four in the previous instrument. The category of "other" was deleted from the list containing "presence/absence" and "change in status."

On Attachment 5, the "other" alpha element category was eliminated, leaving only "Management Transformation" and "Hypotheses." Scoring aspects derived from psychological and experiential evaluation theory were also deleted as not being relevant to AID's evaluative processes. Questions on attribution, sustainability and external influences were added.

#### IV. HOW TO USE THE SCORING INSTRUMENT

There are a number of tasks which are required to complete the various attachments to the scoring instrument. These are summarized in the format of a flow chart in Exhibit IV-1, with each of the different steps highlighted in subsequent discussions for purposes of providing greater clarity and to illustrate the sequence of activities required in order to complete the attachments properly. Definitions and additional clarification pertinent to the various attachments follow each attachments' highlighted appearance on the flow chart. Filling out the attachments constitutes the first phase of using this instrument, while scoring (discussed in Section V) constitutes the second and analysis of results the third phase.

##### FILLING OUT FORMS



The first step in the application of the instrument is to read the evaluation report. While doing this, one should be able to generate the logical framework ("logframe") and to discern the specific indicators used to conduct the evaluation.

Concurrent with reading the report, the reviewer can also complete the Facesheet, which requires some additional explanation.

##### A. THE FACESHEET

While all the data elements on the Facesheet need not be completed at one time (the information needed can usually be found throughout an evaluation report), it is important that the reviewer make sure that all information is eventually filled in. Omitted information can cause delays later on in the scoring process.

Step-by-Step Instructions for Facesheet (see Appendix A.1 for copy of Facesheet)

1. Project title is generally found on the front page of the evaluation report.
2. Project number is usually a 7 or 9 digit number (e.g. 615-0003 or 520-0123.01) which is usually found on the front page. Some types of reports are not numbered, in particular those from the Office of Food and Voluntary Assistance (which are not, strictly speaking, projects).
3. Mission/AID/W/Office is the bureau which implemented the project.
4. Year is usually found on the facesheet on the title page.
5. Evaltype refers to the nature of the project or program, and should be numerically coded by consulting the list of appropriate codes (see Appendix B.1).
6. Evaltime is when the evaluation was conducted in terms of the project's life, and is numerically coded (see Appendix B.1).
7. Mangunit is the office that initiated the evaluation, and is numerically coded (see Appendix B.1).
8. Host country participation can usually be completed simultaneously with items number 9 and 10 (see information below).
9. Contractor(s) refers to the evaluation team. Appendix B.2 contains the appropriate numerical codes, according to the type of contractor. The first three contractors should be listed by name (or name of organization if individuals' names are not indicated) and by contractor code number (see Appendix B.2). Ascertaining the contractor(s) often takes a thorough review of the entire evaluation report. The introduction, evaluation methodology section, appendices or the PES facesheet should be initially examined for this data.
10. Authors are not often clearly indicated. If there is an evaluation team listed with no clear authors, the first name or team leader can be listed here together with "et al." to indicate others.
11. Time taken to do evaluation refers to how long it took the team to complete the evaluation and their report. This is not always indicated, but is sometimes

indirectly addressed in the introduction, evaluation methodology or the annexes. This data item is not to be confused with item number 12.

12. Time to score evaluation indicates the time spent by the reviewer to read the evaluation and complete the attachments.
13. No. levels examined refers to the logical framework generated by the reviewer. Inputs, outputs, purposes and goals each count as one level. If, for example, the reviewer completed attachments up to and including E-2, four levels were examined, that is, inputs, outputs, E-1 and E-2. "Assumptions," "MT," "H" and "U-E" do not count as levels.
14. General indicator should contain brief notes on the evaluation itself, especially anything that the reviewer found unusual. For example, "Attachments cited by report, but not contained in report." This section should also indicate the number of pages in the report, whether there were attachments or annexes, and whether a complete Project Evaluation Summary (PES) was included. This summary is a way to summarize the evaluation, which consists of a standard facesheet and 23 sections pertaining to evaluation findings.
15. Mission Comments are contained on the facesheet, if there is one. They may, on occasion, be otherwise noted in an evaluation report, but this is not common. The comments are usually numbered, and if they are too lengthy they can be summarized.
16. Scope of Work can usually be found in the introduction, evaluation methodology or in the annexes to the evaluation report. Many evaluations do not contain an explicit scope of work at all.
17. Reallocation of inputs, and
18. Reallocation of outputs, are usually difficult to find. A reviewer will usually become aware of "reallocations" while reading through the evaluation report in its entirety.

#### Findings (listed on Facesheet)

TRITON's reviewers were required to list the major findings for each evaluation as part of the qualitative analysis of content. Findings were defined as concise statements about the important conclusions or recommendations contained in the evaluation report. They generally numbered less than ten for each.

This list was attached to the Facesheet and was usually completed at the same time.

The findings were used for a qualitative analysis of the types of conclusions AID draws from evaluations. The analysis described the distribution of findings by reviewers, bureaus and technical activities, and focused on the dissimilarities in the patterns generated.

Examples of Findings:

- o High USAID staff turnover has led to poor communication with the various host country officials.
- o The use of a liaison office in the provincial capitol was an important factor in improving project management.
- o US contractor did not have a clear understanding of its responsibilities or of AID expectations.
- o Targeting towards pregnant mothers, children and ill patients has been particularly successful.
- o The project largely served as an auxiliary to a well-defined and strongly implemented national system.
- o Achievement of rural self-sufficiency was hindered by the inability of target groups to adjust to a cash economy.
- o Researchers can't see why all technologies are not adopted and farmers can't see how researchers can expect them to take so much risk.
- o A great deal more work needs to be directed towards the field testing of prototypes and general extension work.
- o Small private businesses have sprung up on the project areas.
- o The design of the project did not adequately define the role of key decision-makers and in fact contributed to role conflict.
- o The high level of outputs achieved is the direct result of more local participation than originally anticipated.

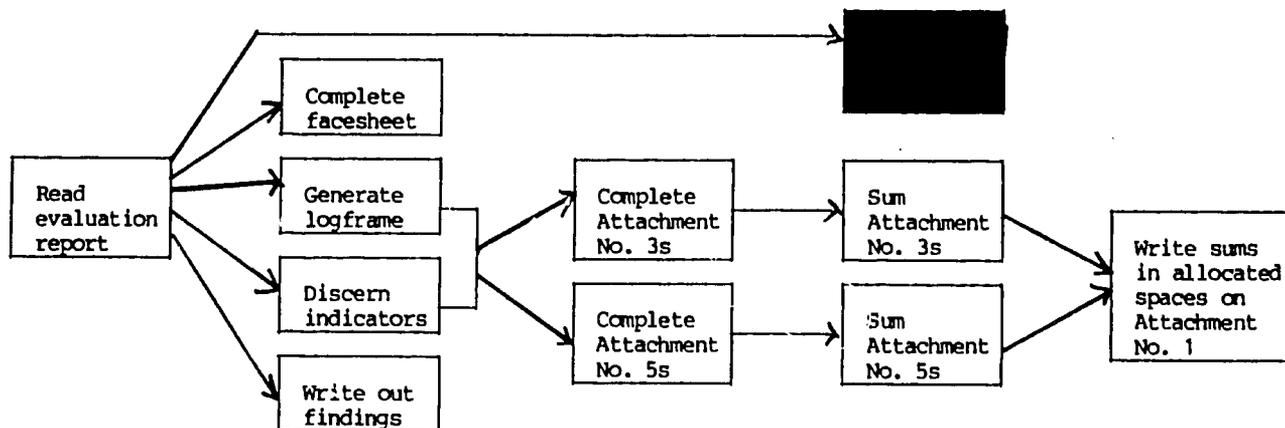
B. ATTACHMENTS 1, 3 and 5

A reviewer will typically have difficulty completing Attachments 1, 3, and 5 before having read a number of different evaluation reports. After exposure to several evaluation reports, the reviewer will then have a basis for ranking the various elements

contained in these Attachments. Thus, it is very critical that a reviewer be exposed early in his or her training to a wide quality range in evaluations reports. Additional information regarding specific scoring values is also provided in Exhibit C, and a suggested reading list on evaluation methodologies is found in Appendix F.

Attachment 1 (see Appendix A.2 for copy of Attachment)

FILLING OUT FORMS



Attachment No. 1 is completed in one of two fashions: after all other Attachments have been completed, or else just prior to completing Attachment Nos. 3 and 5. Coder consensus from the two metaevaluations suggests that Attachment No. 1 be completed before Attachment Nos. 3 and 5, although it is not mandatory.

Some of the subfactors against which the evaluation reports are scored employ specific terminology which requires explanation due to the context in which they are used in this instrument and/or because of potential ambiguity. These are presented in the following section.

Characteristic I

Subfactor 1. "The indicators are appropriate given the evaluation questions."

Indicators refer to the procedures used by the evaluation team to assess the project. These may include documentation reviews, surveys, interviews, site visits, etc.

Subfactor 4. "Assumptions made by the design are clearly and completely stated."

18

EXHIBIT C

**COMPLETENESS:** Select the response that best reflects your perception of how completely the particular factor/topic/issue is addressed by the report:

0-----1-----2-----3-----4

Not addressed.  
Factor/topic/issue  
is totally absent.

Minimally addressed and/or  
addressed in a very super-  
ficial manner.  
Several key aspects of factor/  
topic/issue are not dealt with.

Most key aspects  
are addressed and  
in adequate detail.

All aspects are  
addressed and are  
adequately explored.

**CLARITY:** Select the response that best reflects your perception of how clearly the particular factor/topic/issue is addressed by the report:

0-----1-----2-----3-----4

Not clear.  
Cannot understand point or  
concept that is being  
presented.  
Material not logically  
presented.

Can be understood, but reader has to "work"  
to determine point(s) being expressed. Not  
certain that understanding by reader corre-  
sponds to author's intent.  
Redundancy in presentation confusing.  
Presentation understandable but not logical.

Fully understandable.  
Expressed in very clear language.  
Reader is certain of author's points.  
Author fully conveys his/her thoughts.

**APPROPRIATENESS:** Select the response that best reflects your perception of how appropriately the particular factor/topic/issue is addressed by the report:

0-----1-----2-----3-----4

Totally inappropriate.  
Methods employed, analytical  
techniques, units of measure,  
statistical techniques, etc. are  
not appropriate for what is  
being analyzed, data being  
collected, and/or results  
being derived.

Generally addressed  
inappropriately, but  
selected aspects of the  
factor/topic/issue are  
appropriately analyzed,  
measured, etc.

Generally addressed in  
an appropriate manner  
but selected aspects  
(e.g., one of four units  
of measure) are not appro-  
priately addressed.

Totally appropriate.  
The methodology, analyses,  
measurement tools, etc. are  
fully consistent with  
generally accepted princi-  
ples and practices regarding  
evaluations and the particular  
factor/topic/issue being  
addressed.

This subfactor refers to the evaluation design, not the project's design.

#### Characteristic II

Subfactor 3. "The evaluation questions are clearly completely stated; priorities among questions are clear."

What exactly is the evaluation trying to find out about the project? What types of questions does it try to answer? Is there an implicit/explicit priority ordering of those questions?

#### Characteristic III

Subfactor 3. "Areas of "public interest"/broad concern covered by the evaluation are clearly identified."

This can sometimes be found under the "Lessons Learned" section in the evaluation report.

#### Characteristic IV

Subfactor 1. "Instruments/approaches for collecting data are valid and reliable."

For example, one would question the validity of making conclusions about benefits to community members based solely on discussions with mission personnel.

Subfactor 2. "Validity and reliability of any secondary data is checked and found acceptable."

Secondary data is data found in the evaluation report that has been taken from another source; for example, statistics from the Ministry of Health.

Characteristic V: "Findings, conclusions and recommendations are presented in a way that clearly separates facts from interpretations."

In general, the reviewer must be able to clearly recognize the differences between facts and interpretations, as well as between conclusions and recommendations, within the context of an evaluation report.

Characteristic VI: "The data analysis procedures are appropriate and adequate."

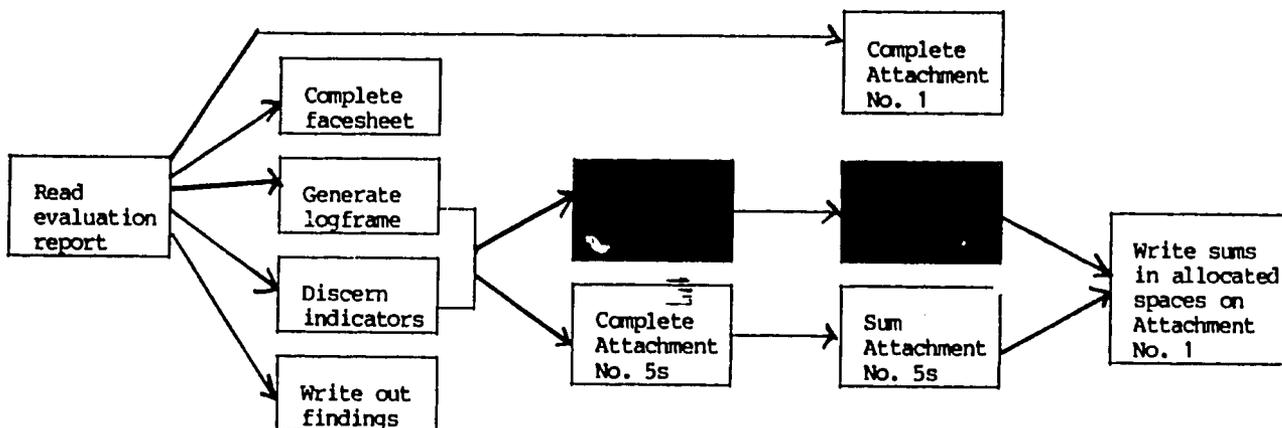
Analysis is the way in which the information gathered by the evaluators is converted into the form of conclusions and recommendations.

Characteristic VII: "The evaluation report is a well-written, self-contained document."

Self-contained means that all the information that the user needs is contained within the evaluation report and attachments.

Attachment 3 (see Appendix C for example of Attachment)

FILLING OUT FORMS



One Attachment No. 3 is completed for: all inputs combined; each output: each "effect" level (that is, the purpose and goals of the program); and each assumption. At this point in the process, a reviewer should have completed the Facesheet, Findings, Logframe and Attachment No. 1.

Indicators

To complete this attachment, the reviewer must find out what types of indicators are used to assess the project's status. The most common of these are:

- o documentation review
- o interviews
- o questionnaires/surveys
- o site visits/observation.

21

These can usually be found in the Evaluation Methodology section of the report, if there is one. Sometimes they are not explicitly indicated, and the reviewer will need to "read into" the evaluation to find the indicators used. These are generally found while the reviewer is reading the evaluation the first time, and will not necessarily be the same for all elements examined. There is a space next to Section B on Attachment 3 where the indicators should be listed in the order in which they are scored on the Attachment.

Valid means that sufficient evidence was presented that the indicator was sound.

Replicable means that enough information was given about the measure (indicator) so that one could repeat it and obtain similar results.

Unbiased means that the information obtained was not skewed toward a particular viewpoint.

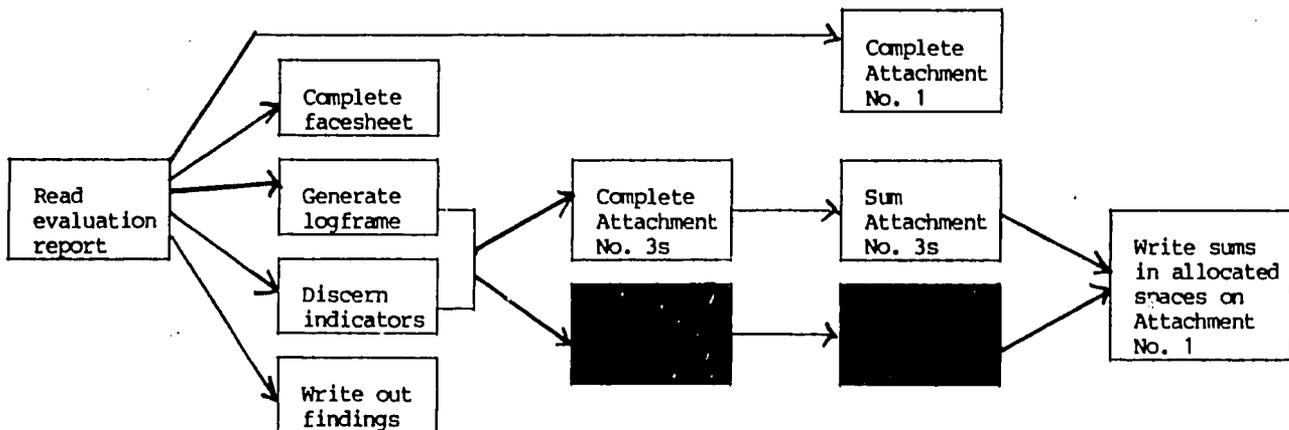
Objective means that a sufficient number of sources were used to ensure that a balanced and representative package of information was obtained.

Section D of Attachment 3 is usually completed for the following common types of data instruments for collecting indicators: questionnaires, surveys, interviews, site visits and observations. Other indicators may be found, but not significantly.

Section E should be completed by the reviewer. This is a preliminary scoring procedure, and consists of writing down the values for each indicator and then summarizing them.

Attachment 5 (see Appendix D for example of Attachment)

FILLING OUT FORMS



Once the Attachment No. 3s have been completed and summed, this leaves only three tasks to be performed. Two of these are the completion and summation of Attachment No. 5s. One Attachment No. 5 is completed for the Management Transformation, and one is completed for all 'Hypothesis' elements combined. These results are then summed.

Management transformation is the process by which the project implementors produce the outputs.

If the management transformation examined by the report only addresses outcomes of the project and not the process, by which the outputs were generated then the sum is a "missing" value (i.e., not applicable). If the process was examined in the management transformation, then the evaluation report discusses the project management involved in converting inputs to outputs and not merely their outcomes. "Process from another standpoint" means that the process by which outputs were generated was in the report from quality and efficiency perspectives. If this is the case, one of the following should be specified:

- o quality and time
- o quality and cost
- o quality, time and cost.

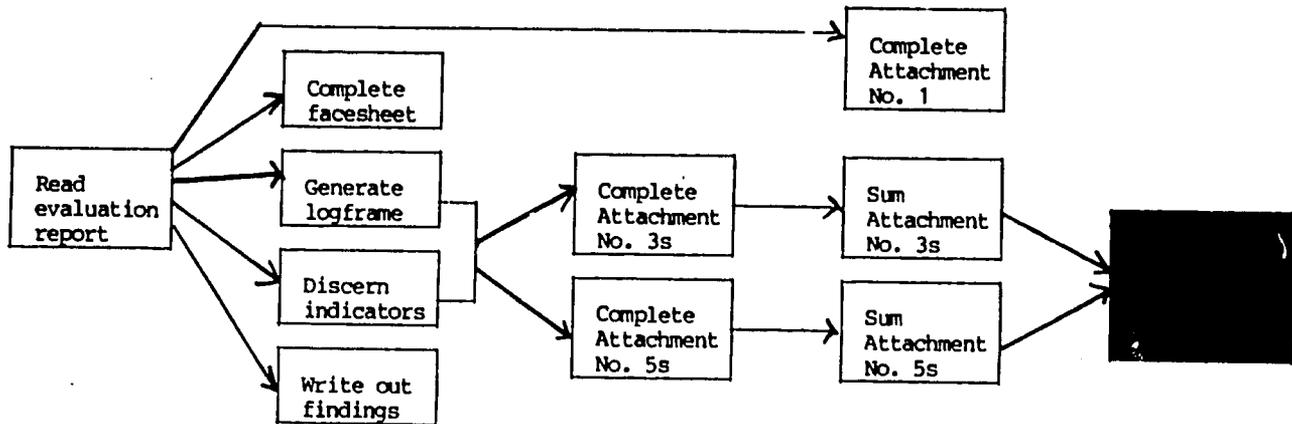
The use of the terms "valid," "reliable," "unbiased" and "objective" was explained under Attachment 3.

The "Hypotheses" refer to the <sup>sum</sup> transition from outputs to purpose, and from purpose to goal. If an evaluation report addresses the Hypotheses, it examines the logic of the connections from one level to another.

A given set of inputs will not automatically or by themselves lead to the desired outputs. The process (inputs generating outputs, etc.) cannot be assumed to have occurred in a causal relationship just because it was hypothesized in the project design. Therefore, it should also be addressed as part of the evaluative procedures include an analysis of that transition.

# Transfer Of Sums To Attachment No. 1

## FILLING OUT FORMS



The final step in completing the various attachments is to transfer the sums generated by score values for the types of attachments discussed above: For example, all Attachment No. 3's together concerning for outputs are transferred to Attachment No. 1, page one. The Attachment No. 5s are entered on page two of Attachment No. 1 and require one additional calculation: "Management Transformation" value is multiplied by 6.25 and the "Hypothesis" value is multiplied by 2.27 before completing Attachment No. 1.

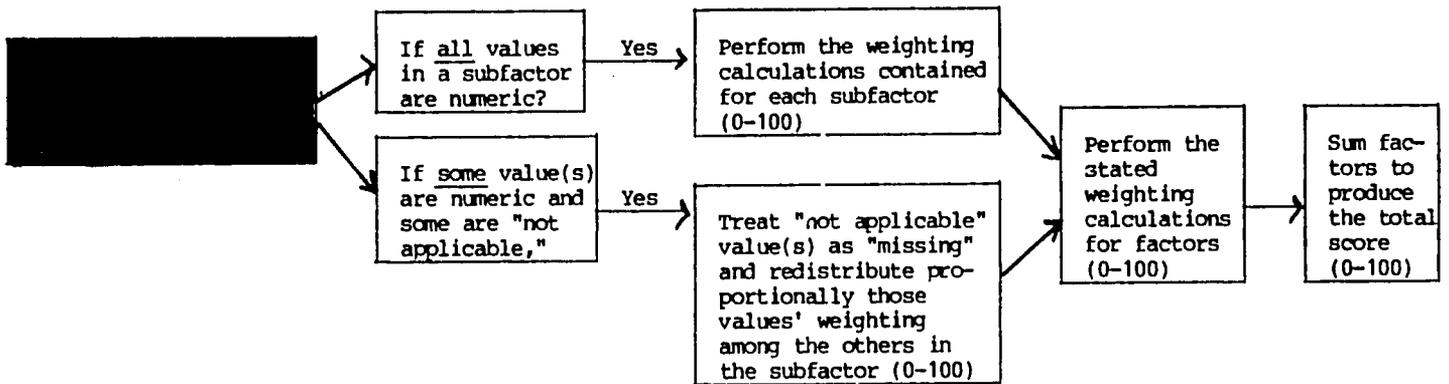
## V. SCORING PROCESS

### A. OVERVIEW

The scoring process itself is the second stage of scoring the evaluation reports. At this point, all of the steps for the first stage have been completed, and the reviewer should have both the scoring instrument and Attachment No. 1 completed and available for review.

### B. CALCULATION STEPS

#### SCORING PROCEDURES



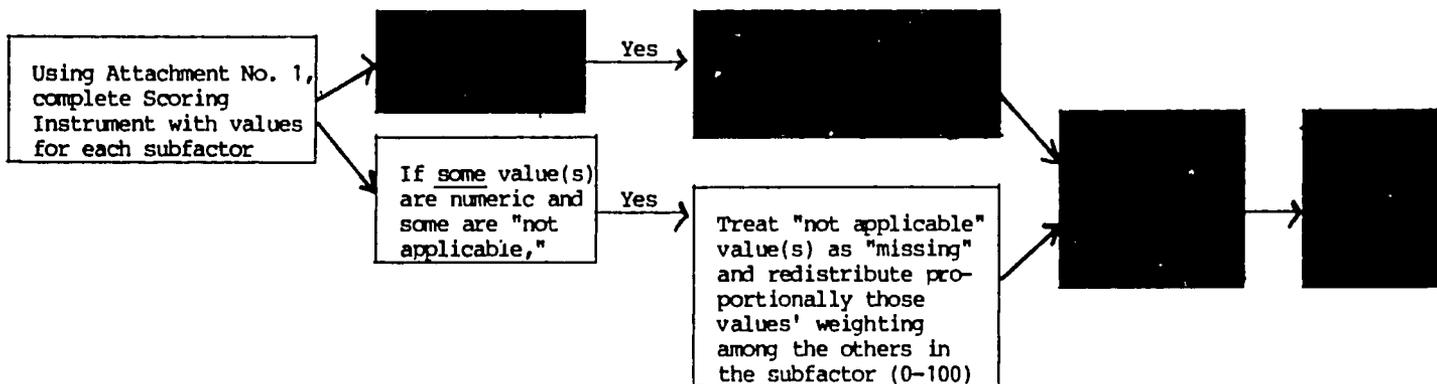
The process of computing the score from a completed scoring instrument is outlined below. This procedure is generally done by someone other than the reviewer who scored the report(s) to avoid bias, but can be done by the reviewer.

1. Write the project number on the upper right hand corner of the sheet titled Scoring Instrument (see Appendix E).
2. Tear off the first two sheets of Attachment No. 1 and place Logical Framework (sheet with rectangular boxes) in plain view. Use this as a guideline for the number of Attachments that were supposed to have been completed.
3. Transfer the sums from Attachment No. 1 to the scoring instrument, drawing a line through any space without a summary score. This is treated as a missing value (not applicable).
4. Add all the Elements together Characteristic I, in Subfactor 2, after the first equal sign.
5. If the sum for any of the summary scores in Step 4 is zero, write a zero in the space as that has value in the

scoring process and is differentiated from missing values in this manner.

6. Divide the results for Step 4 by the no. of elements given. Divide by 5 only if there are five elements. Cross out the 5 when there are less than that number of elements and place the actual number of elements underneath it. For example, if there were only "inputs," "outputs" and an assumption ("A") element, then divide the total by 3. Complete the computation.
7. If either "management transformation" or "hypothesis" is a missing value, then divide Subfactor 3 by 1, and not by 2.
8. Attachment 1 is the overall scoring instrument and is self-explanatory. Each "characteristic" corresponds to the scoring sheet.
9. Whenever "N/A" is circled on Attachment 1, place "M" after the "=" sign and DO NOT write any total for the characteristic. This is computed during the procedures subsequently used to perform the analysis.
10. Write down all the values for each factor and its subfactors from Attachment No. 1 on the scoring instrument.

SCORING PROCEDURES

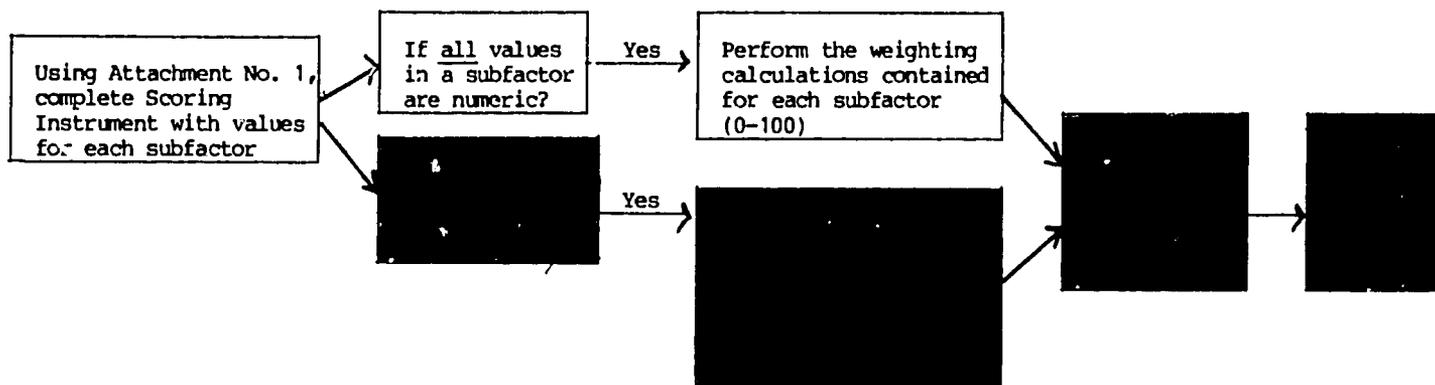


11. If all values in a subfactor are numeric; that is, there are no "missing" values within a subfactor, then perform the weighting calculations which will result in a value for that subfactor of from 0-100. If all subfactors for a given characteristic have numeric values, then the weighting calculations for that characteristic (factor)

can be performed. This should also result in a value from 0-100 for each characteristic.

12. Sum all the characteristics to produce the total score for the evaluation report.

SCORING PROCEDURES



13. If all values in a subfactor are not numeric, then one has to recalibrate the weighting system for each characteristic in a manner proportionate to the existing weights (which add up to 100).

For example, if Subfactor 7, Characteristic 1, is missing, then the weight of .10 for that subfactor must be divided among Subfactors 1-6 in proportion to their original relative weighting scheme. Subfactor 1's weighting would change from .13 to  $(.13 \div .90)$ , that is, .14. Subfactor 2's weighting would change from (.25 to  $.25 \div .90$ ), that is, .28, etc.

More than one subfactor may be missing, which results in more than one weight having to be redistributed.

14. Once a numeric value can be reached for each of the Characteristics, then these can be weighted, and the totals summed to produce a total score for the evaluation apart from 0-100.

## VI. UTILIZATION OF SCORES

There are three primary uses for scoring evaluation reports in the above-described manner. The first is in design of evaluations, the second is in evaluation review, and the third is in gaining a general knowledge and understanding of the evaluation reports. Each of these analytical uses of the metaevaluation methodology's results is discussed below.

### A. DESIGN

The compilation and analysis of scores provides a means to pinpoint failings in the design of both projects and evaluations. Evaluation design deficiencies can be addressed if a certain scoring trend is revealed; for example, persistent low scores for Characteristic IV (data collection). Remedial measures can then be undertaken to upgrade those scores. Project design can be addressed through an analysis of the scores by either technical activities or categories of findings. The concentration of low scores for one type of technical activity or the prevalence of negative findings are strong indicators of the need for remedial action with regard to project design.

### B. EVALUATION REVIEW

An efficient way to assess the historical performance of evaluation reports is through the use of the scoring instrument. The distribution of scores can be performed by a time-series analysis, which graphically illustrates scoring patterns and changes in these patterns over a specific time period. This provides an historical review of a mission's performance in evaluations, and can be used in conjunction with an annual application of the scoring instrument by sampling evaluations reports from previous years. This is especially useful if the instrument is used for at least two years, since remedial actions taken as the result of one year's quality review will require another year before being evaluated in the course of project cycles.

Another use associated with evaluation reviews is an analysis of the utilization of the recommendations contained in one year's evaluation report by reviewing subsequent evaluation reports. This provides a means of determining not only the information flow within an organization, but also the types of recommendations which actually acted on.

### C. GENERAL KNOWLEDGE

The more evaluation reports a reviewer reads, the better able that reviewer will be to: 1) organize his/her own evaluation report; 2) make suggestions for revising ongoing evaluation reports; and 3) propose revisions in the overall project cycle, with specific reference to design and evaluation. The application of the scoring instrument provides an unique opportunity for

acquiring a valuable background to project management, especially since the process requires the iterative use of AID's key project management tool, the logical framework.

The application of the scoring instrument, then, provides not only a broad base of information abstracted from the evaluation reports, but also the development of skills in generic project management and analysis.

APPENDIX A.1

FACE SHEET DATA

1. Project Title \_\_\_\_\_
2. Project Number \_\_\_\_\_
3. Mission/AID/W/Office \_\_\_\_\_
4. Year of Evaluation Review \_\_\_\_\_
5. Evaltype \_\_\_\_\_
6. Evaltime \_\_\_\_\_
7. Mangunit \_\_\_\_\_
8. Host Country Participation on the Evaluation Team?  
Yes \_\_\_\_\_ No \_\_\_\_\_ Can't Tell \_\_\_\_\_
9. Contractor(s): List principal one first  
\_\_\_\_\_
10. Author(s):
11. Time taken to do Evaluation \_\_\_\_\_
12. Time taken to Score Evaluation \_\_\_\_\_
13. No. levels examined \_\_\_\_\_
14. General Indicator of evaluation completeness/innovative techniques.
15. Mission comments:
16. Scope of Work Included in the Documents?  
Yes \_\_\_\_\_ No \_\_\_\_\_
17. Did the evaluators discover that resources needed to be reallocated among all the inputs to achieve outputs?  
Yes \_\_\_\_\_ No \_\_\_\_\_  
If yes, describe how:

FACE SHEET DATA  
(Continued)

18. Did the evaluators discover that resources needed to be reallocated among the outputs to achieve project purpose?

Yes \_\_\_\_\_ No \_\_\_\_\_

If yes, describe how:

FINDINGS

ATTACHMENT 2  
LOGICAL FRAMEWORK

E-4

Hd

E-3

Hc

E-2

Hb

E-1

Ha

OUTPUTS

MANAGEMENT  
TRANSFORMATION

INPUTS

A-E3

A-E2

A-E1

A-0

APPENDIX 4.2

ATTACHMENT 1

OVERALL SCORING INSTRUMENT

(with scales for Completeness, Clarity and Appropriateness)

CHARACTERISTIC I: The overall design of the evaluation is appropriate for answering the evaluation questions.

SUB-FACTORS TO BE ADDRESSED FOR THIS CHARACTERISTIC

1. The indicators are appropriate given the evaluation questions.

Appropriateness	0	1	2	3	4
-----------------	---	---	---	---	---

2. As appropriate, given the stage of the evaluation, the evaluation design contains procedures for measuring project efficiency, effectiveness (e.g., the provision of goods/services to intended beneficiaries of the goods/services provided by a project or program). All measurement approaches in the design are conceptually valid. To the degree appropriate, the measurement approaches consider such factors as the timeliness with which goods/services are delivered, the duration of services, etc.

Enter values from Worksheet:

Summary Score for U elements: \_\_\_\_\_

Summary Score for E elements: \_\_\_\_\_

Summary Score for A elements: \_\_\_\_\_

Summary Score for Output elements: \_\_\_\_\_

Summary Score for Input elements: \_\_\_\_\_

3. As appropriate, given the stage of the evaluation, the evaluation design contains procedures for examining the strength and validity of hypothesized cause and effect linkages. These procedures are appropriate for making determinations concerning the probability that a particular cause or means (provided by the project or program) explains

the effects/outcomes/impacts (of the project or program). The procedures for examining cause and effect relationships are strong enough to give reasonable assurance that major "rival" explanations will be considered and eliminated before claims of a relationship between a project or program and a set of effects/outcomes/impacts are made.

Enter values from Worksheet:

Summary Score for MT elements: \_\_\_\_\_

Summary Score for H elements: \_\_\_\_\_

4. Assumptions made by the design are clearly and completely stated.

Completeness:	0	1	2	3	4
Clarity:	0	1	2	3	4

5. If the design is adapted from another evaluation or research study, it is customized for the situation in which it is to be used, if required.

Completeness:	0	1	2	3	4	N/A
Clarity:	0	1	2	3	4	N/A
Appropriateness:	0	1	2	3	4	N/A

6. The evaluation design is fully and clearly described by the evaluation report.

Completeness:	0	1	2	3	4
Clarity:	0	1	2	3	4

7. The design includes procedures for recording any changes in the methodology made during the course of the evaluation and where such changes occur, the evaluation report discusses them.

Completeness:	0	1	2	3	4	N/A
Clarity:	0	1	2	3	4	N/A
Appropriateness:	0	1	2	3	4	N/A

CHARACTERISTIC II: The evaluation clearly and completely identifies the objectives of the project or program which is being evaluated as well as the evaluation objectives and questions.

SUBFACTORS TO BE ASSESSED FOR THIS CHARACTERISTIC

1. Project or program objectives are clearly and completely stated.

Completeness: 0 1 2 3 4  
Clarity: 0 1 2 3 4

2. The objectives of the evaluation are clearly and completely stated; priorities among objectives and reasons for some are clear.

Completeness: 0 1 2 3 4  
Clarity: 0 1 2 3 4

3. The evaluation questions are clearly and completely stated; priorities among questions are clear.

Completeness: 0 1 2 3 4  
Clarity: 0 1 2 3 4

CHARACTERISTIC III: The evaluation focuses on the evaluation users and their needs/questions.

SUB-FACTORS TO BE ASSESSED FOR THIS CHARACTERISTIC

1. Evaluation clients/users are clearly and completely identified.

Completeness:	0	1	2	3	4
Clarity:	0	1	2	3	4

2. User needs/expectations are clearly and completely identified.

Completeness:	0	1	2	3	4
Clarity:	0	1	2	3	4

3. Areas of "public interest"/broad concern covered by the evaluation are clearly identified.

Completeness:	0	1	2	3	4
Clarity:	0	1	2	3	4

CHARACTERISTIC IV: The data collection procedures/secondary data are appropriate and adequate, not excessive or inadequate.

SUB-FACTORS TO BE ADDRESSED FOR THIS CHARACTERISTIC

1. Instruments/approaches for collecting data are valid and reliable;

Completeness:	0	1	2	3	4
Clarity:	0	1	2	3	4
Appropriateness:	0	1	2	3	4

2. Validity and reliability of any secondary data is checked and found acceptable.

Completeness:	0	1	2	3	4	N/A
Clarity:	0	1	2	3	4	N/A
Appropriateness:	0	1	2	3	4	N/A

3. Sources of error/biases in the instruments or data collection procedures are described as fully as possible.

Completeness:	0	1	2	3	4
Clarity:	0	1	2	3	4

4. Where there is a need to generalize from the data to a larger population, either sampling procedures which allow such generalization are properly used or the limits on generalizing from the data are fully stated.

Completeness:	0	1	2	3	4	N/A
Clarity:	0	1	2	3	4	N/A
Appropriateness:	0	1	2	3	4	N/A

5. Neither too much or too little data is secured.

Appropriateness: 0 1 2 3 4

6. Where cross-cultural sensitivity, language, etc. are potential issues, they are properly handled (e.g. local data collectors used, female data collectors, etc.)

Completeness: 0 1 2 3 4 N/A

Clarity: 0 1 2 3 4 N/A

Appropriateness: 0 1 2 3 4 N/A

7. Where data must be collected and it is important to do this in a non-disruptive manner, the data collection procedures are as non-disruptive as possible.

Completeness: 0 1 2 3 4 N/A

Clarity: 0 1 2 3 4 N/A

Appropriateness: 0 1 2 3 4 N/A

8. Instruments used to collect raw data, such as questionnaires, are included as exhibits to evaluation reports.

Completeness: 0 1 2 3 4 N/A

CHARACTERISTIC V: Findings, conclusions and recommendations are presented in a way that clearly separates facts from interpretations.

SUB-FACTORS TO BE ADDRESSED FOR THIS CHARACTERISTICS

1. Facts are separated from interpretations.

Completeness:	0	1	2	3	4
Clarity:	0	1	2	3	4

2. Alternative interpretations are discussed.

Completeness:	0	1	2	3	4
Clarity:	0	1	2	3	4

3. The reason for selecting a specific interpretation or conclusion is made clear.

Completeness:	0	1	2	3	4
Clarity:	0	1	2	3	4

4. Conclusions are separated from recommendations.

Completeness:	0	1	2	3	4
Clarity:	0	1	2	3	4

5. Alternative recommendations are discussed and the reason for selecting a specific recommendation is made clear.

Completeness:	0	1	2	3	4
Clarity:	0	1	2	3	4

6. The reasons for selecting a specific recommendation are made clear

Completeness:	0	1	2	3	4
Clarity:	0	1	2	3	4

7. The study findings, conclusions and recommendations are well organized and presented in a fashion that is understandable to a busy reader/decision-maker who may not be familiar with how studies are conducted.

Clarity:                    0            1            2            3            4

8. The material on findings, conclusions and recommendations is presented clearly and objectively, in the sense that it neither "hides" data nor makes assertions without adequate facts.

Clarity:                    0            1            2            3            4

Appropriateness:        0            1            2            3            4

9. The evaluators come a "bottom line" where the evaluation questions and purposes require that some firm conclusions be drawn in the course of the evaluation; i.e., did the project succeed in achieving its objectives or not?

Completeness:            0            1            2            3            4

Clarity:                    0            1            2            3            4

CHARACTERISTIC VI: The data analysis procedures are appropriate and adequate.

SUB-FACTORS TO BE ADDRESSED FOR THIS CHARACTERISTIC

1. The analysis procedures are clearly presented, match the purposes of the evaluation and fit the evaluation questions and data collected to answer those questions.

Completeness:	0	1	2	3	4
Clarity:	0	1	2	3	4
Appropriateness:	0	1	2	3	4

2. The analysis procedures are appropriate; they are neither weak nor excessive.

Appropriateness:	0	1	2	3	4
------------------	---	---	---	---	---

3. Where appropriate, the confidence level of findings is given; e.g., statistical significances of comparisons of quantitative data on two groups, descriptive statements about the confidence that should be placed in answers arrived at through non-quantitative data and analysis.

Completeness:	0	1	2	3	4	N/A
Clarity:	0	1	2	3	4	N/A
Appropriateness:	0	1	2	3	4	N/A

4. Both quantitative and qualitative data are analyzed if both were secured.

Completeness:	0	1	2	3	4
Clarity:	0	1	2	3	4

5. Where possible, the evaluation examines how realistic were the project's original estimates of cost, economic return, etc., as well as data on project/program effectiveness and impact.

Completeness:	0	1	2	3	4	N/A
Clarity:	0	1	2	3	4	N/A
Appropriateness:	0	1	2	3	4	N/A

6. The strength and weaknesses of the data analysis aspects of the evaluation are clearly and completely stated.

Completeness:	0	1	2	3	4
Clarity:	0	1	2	3	4

7. Where appropriate, the raw data from the study are included, or their availability made known, should it be necessary/appropriate to re-analyze all or part of the study data.

Completeness:	0	1	2	3	4	N/A
Clarity:	0	1	2	3	4	N/A

CHARACTERISTIC VII: The evaluation report is a well-written, self contained document.

Completeness:	0	1	2	3	4
Clarity:	0	1	2	3	4

CHARACTERISTIC VIII: The evaluation produces the types of information it was expected to produce; i.e., insofar as possible, the full set of evaluation questions are answered.

Completeness:	0	1	2	3	4
Clarity:	0	1	2	3	4

CHARACTERISTIC IX: Action implications of the evaluation are clearly stated and are annotated to indicate who or what unit should act.

Completeness:	0	1	2	3	4
Clarity:	0	1	2	3	4
Appropriateness:	0	1	2	3	4

CODES

APPENDIX B.1

BUREAU  
(BUREAU)

1. Near East
2. Asia
3. LAC
4. frica
5. Impact
6. S & T
7. FVA

EVALUATION TYPE  
(EVALTYPE)

1. Country project
2. Country multi-project
3. Country program
4. Regional
5. Other

MANAGEMENT UNIT  
(MANGUNIT)

1. Mission
2. Sub-regional
3. Regional
4. AID/W Central Bureau
5. AID bureau level
6. PPC Impact
7. Other

TECHNICAL\*  
(RETEC)

0. Agric single function
1. Rural non-agric activities
2. Rural multi function
3. Nutrition
4. Population
5. Health
6. Education
7. Human resources
8. Infrastructure and housing
9. Other

EVALUATION TIME  
(MANGTIME)

1. Interim
2. Final
3. Ex/post
4. Other

HOST PARTICIPATION  
(HOST COUN)

1. Yes
2. No
3. Unknown

LEVELS EXAMINED  
(LEVEL)

0. Process only
1. 1 effect only
2. 2 effect levels
3. 3 effect levels
4. 4 effect levels

SCOPE OF WORK

1. Yes
2. No

RESOURCE REALLOCATION  
(INPREAL) (OUTREAL)

- |        |        |
|--------|--------|
| 1. Yes | 1. Yes |
| 2. No  | 2. No  |

<u>CODE</u>	<u>DESCRIPTION</u>
01	AID Mission staff as implementors/evaluators
02	AID Mission staff as external evaluators
03	AID (TDY) as external evaluators
10	US university staff as implementors/evaluators
11	US university staff as external evaluators
12	Host country university staff as implementors/evaluators
13	Host country university staff as external evaluators
14	Third country university staff as external evaluators
20	US consulting firm/private research organization as implementors/evaluators
21	US consulting firm/private research organization as external evaluators
22	Host country consulting firm/private research organization as implementors/evaluators
23	Host country consulting firm/private research organization as external evaluators
30	Free lance US consultant as implementor/evaluator
31	Free lance US consultant as external evaluator
32	Free lance host country consultant as implementor/evaluator
33	Free lance host country consultant as external evaluator
34	Free lance third country consultant as external evaluator
40	PASA/RSSA personnel (e.g. USDA) as implementors/evaluators
41	PASA/RSSA personnel as external evaluators
50	Peace Corps staff or volunteers as external evaluators
51	Peace Corps staff or volunteers as implementors/evaluators
60	Int'l agencies (bilateral or multinational) as external evaluators
70	LDC government staff as external evaluators
71	LDC government staff as implementor/evaluator
80	US based PVO as implementors/evaluators
81	US based PVO as external evaluators
82	Host country PVO as implementors/evaluators
83	Host country PVO as external evaluators
90	IMPACT: AID personnel
91	IMPACT: Other than AID

APPENDIX C

ATTACHMENT 3

RATING FORM FOR SCORING INPUTS, OUTPUTS,  
DEPENDENT VARIABLES ASSUMPTIONS,  
AND UNPLANNED RESULTS

Note: Complete 1 copy of Form to address all INPUTS together.  
Complete 1 copy of Form for each OUTPUT.  
Complete 1 copy of Form for each DEPENDENT VARIABLE  
Complete 1 copy of Form for each set of ASSUMPTIONS

Element being scored: \_\_\_\_\_

(For example, Inputs, Output1, E1, A-E1, U-E1)

A. Type of variable addressed by this project element being evaluated:

- \_\_\_\_\_ Independent variable (for this project/program/policy)
- \_\_\_\_\_ Dependent variable (for this project/program/policy)
- \_\_\_\_\_ Other. Specify type of variable/element and describe:

B. Number of indicators used in evaluation report to measure status of variable: \_\_\_\_\_

C. Answer for each indicator measured for this element:

(1) Check which of these is applicable:

Ind	Ind	Ind	Ind	Ind	Ind	
1	2	3	4	5	6	
_____	_____	_____	_____	_____	_____	a. Presence/absence (i.e., indicator was <u>not</u> present "before" activity being evaluated began).
_____	_____	_____	_____	_____	_____	b. Change in status (i.e., indicator <u>was</u> present "before" activity being evaluated began; measure focuses on change)

(2) Complete only if C (1) response = presence/absence (response a). Score 0 = No, 2 = Somewhat, 4 = Yes:

Ind	Ind	Ind	Ind	Ind	Ind	
1	2	3	4	5	6	
_____	_____	_____	_____	_____	_____	(a) Measure was <u>valid</u> measure of presence/absence for the indicator
_____	_____	_____	_____	_____	_____	(b) Measure was <u>replicable</u>
_____	_____	_____	_____	_____	_____	(c) Measure was <u>unbiased</u>
_____	_____	_____	_____	_____	_____	(d) Measure was <u>objective</u>

(3) Complete only if C (1) response = change in status (response b). Score 0 = No, 2 = Somewhat, 4 = Yes

Ind Ind Ind Ind Ind Ind  
 1 2 3 4 5 6

- |     |     |     |     |     |     |   |
|-----|-----|-----|-----|-----|-----|---|
| ___ | ___ | ___ | ___ | ___ | ___ | (a) Measure was <u>valid</u> measure of indicator which was to have changed |
| ___ | ___ | ___ | ___ | ___ | ___ | (b) Measures at all points were made in <u>consistent</u> manner            |
| ___ | ___ | ___ | ___ | ___ | ___ | (c) Measures of indicator was <u>unbiased</u>                               |
| ___ | ___ | ___ | ___ | ___ | ___ | (d) Measure was <u>adequate</u> , given inherent variability in indicator   |
| ___ | ___ | ___ | ___ | ___ | ___ | (e) Measures at all points were <u>objective</u>                            |

D. Generalization: Complete only if evaluation sought/attempted to generalize for a universe based on measures made of indicator for a subset of that relevant universe. Enter one value for each indicator form which a generalization was made:

Ind Ind Ind Ind Ind Ind  
 1 2 3 4 5 6

- |     |     |     |     |     |     |  |
|-----|-----|-----|-----|-----|-----|--|
| ___ | ___ | ___ | ___ | ___ | ___ | Statistically sound/representative sample = 4        |
| ___ | ___ | ___ | ___ | ___ | ___ | Random selection procedure/universe size unknown = 3 |
| ___ | ___ | ___ | ___ | ___ | ___ | Criteria or other purposive sample = 2               |
| ___ | ___ | ___ | ___ | ___ | ___ | Convenience or volunteer sample = 1                  |
| ___ | ___ | ___ | ___ | ___ | ___ | Single case (of larger universe) = 1                 |
| ___ | ___ | ___ | ___ | ___ | ___ | Only case (automatic census)/all cases = 4           |
| ___ | ___ | ___ | ___ | ___ | ___ | Can't tell from evaluation report = 0                |

**E. Summary score on the finding/measure made:**

Ind 1	Ind 2	Ind 3	Ind 4	Ind 5	Ind 6	Total for All <u>Indicators</u>	
---	---	---	---	---	---	---	Validity: Score from C(2)(a) <u>or</u> C(3)(a)
---	---	---	---	---	---	---	Replicability/consistency: Scor from C(2)(b) or C(3)(b)
---	---	---	---	---	---	---	Bias: Score from C(2)(c) <u>or</u> C(3)
---	---	---	---	---	---	---	Representativeness/Adequacy: Sc from C(3)(d)
---	---	---	---	---	---	---	Objectivity: Score from C(2)(d) C(3)(e)
---	---	---	---	---	---	---	Generalization: Score from Item
---	---	---	---	---	---	---	Grand Total

APPENDIX D

ATTACHMENT 5

RATING FORM FOR SCORING THE MANAGEMENT  
TRANSFORMATION AND HYPOTHESES (Ha, Hb, Hc....)

Note: Complete 1 copy of Form for the MANAGEMENT TRANSFORMATION  
Complete 1 copy of Form for all HYPOTHESES (Ha, Hb, etc.)

50

Element being scored: \_\_\_\_\_  
(MT or H)

---

Type of alpha element (check one):

\_\_\_\_\_ Management transformation (no hypothesis presented;  
i.e., "effective management" is the primary process  
needed to generate desired effects).

\_\_\_\_\_ Hypothesis (from independent to dependent variable,  
planned or unplanned, etc.)

---

A. Answer if element = Management Transformation:

(1) What was examined to determine whether transformation  
occurred:

\_\_\_\_\_ (a) Outcome only (specify which outcomes, as per  
diagram in Attachment 2: Output # \_\_\_)

\_\_\_\_\_ (b) Process, from a quality standpoint

\_\_\_\_\_ (c) Process, from an efficiency standpoint (specify  
from from which perspective(s): \_\_\_\_\_ time, \_\_\_\_\_  
cost, \_\_\_\_\_ time and cost)

\_\_\_\_\_ (d) Process, from another standpoint. Specify:

---

---

- (2) Complete only if answer to A(1) = process in any form (response b, c or d); Score 0 = No, 2 = Somewhat, 4 = Yes:

\_\_\_\_\_ Process measure was valid for situation.

\_\_\_\_\_ Process measure was reliable.

\_\_\_\_\_ Process measure was unbiased.

\_\_\_\_\_ Process measure was objective.

B. Complete only if element = hypothesis:

(1) Was the logic requirement that the hypothesized cause preceded the effect met: \_\_\_\_\_ Yes \_\_\_\_\_ No \_\_\_\_\_ Can't Tell

(2) Was the logic requirement that the hypothesized cause and effect covaried (both changed in status) met: \_\_\_\_\_ Yes \_\_\_\_\_ No \_\_\_\_\_ Can't Tell

C. Attribution

1. Did the evaluation attribute some result to some aspect of the project?

Yes \_\_\_\_\_ No \_\_\_\_\_

2. If the evaluation made such a statement, was the proof:

Adequate:	0	1	2	3	4
Unbiased:	0	1	2	3	4
Valid:	0	1	2	3	4

3. To what extent were exogenous variables (price, self-selection, initial economic order) examined?

0            1            2            3            4

4. To what extent were exogenous variables responsible for project achievements/failures?

0            1            2            3            4

5. Were exogenous variables examined in the evaluation?

Yes \_\_\_\_\_ No \_\_\_\_\_

If yes, list: 1. \_\_\_\_\_  
2. \_\_\_\_\_  
3. \_\_\_\_\_

6. Did the evaluators come to a conclusion about the project's sustainability?

Yes \_\_\_\_\_ No \_\_\_\_\_

7. If the evaluators came to a conclusion, was the project considered sustainable?

0            1            2            3            4

Summary score on element:

6.25 x A(2) Score \_\_\_\_\_ or 2.27 x (B(1) + B(2) + C Score)

APPENDIX E

SCORING INSTRUMENT

Attachment 1

CHARACTERISTIC I:

Subfactor 1: Ap \_\_\_\_\_ x 25.0 = \_\_\_\_\_ x .13 = \_\_\_\_\_

Subfactor 2: Summary Score for U Elements \_\_\_\_\_  
+ Summary Score for E Elements \_\_\_\_\_  
+ Summary Score for A Elements \_\_\_\_\_  
+ Summary Score for Output Elements \_\_\_\_\_  
+ Score for Input Elements \_\_\_\_\_  
= \_\_\_\_\_ + 5.0\* x .25 = \_\_\_\_\_

Subfactor 3: Score for MT element \_\_\_\_\_  
+ Score for H element \_\_\_\_\_  
= \_\_\_\_\_ + 2.0\* x .15 = \_\_\_\_\_

Subfactor 4: Co \_\_\_\_\_ + Cl \_\_\_\_\_ = \_\_\_\_\_ x 12.5 = \_\_\_\_\_ x .15 = \_\_\_\_\_

Subfactor 5: Co \_\_\_\_\_ + Cl \_\_\_\_\_ + Ap \_\_\_\_\_ = \_\_\_\_\_ x 8.33 = \_\_\_\_\_  
x .10 = \_\_\_\_\_

Subfactor 6: Co \_\_\_\_\_ + Cl \_\_\_\_\_ = \_\_\_\_\_ x 12.5 = \_\_\_\_\_ x .12 = \_\_\_\_\_

Subfactor 7: Co \_\_\_\_\_ + Cl \_\_\_\_\_ + Ap \_\_\_\_\_ = \_\_\_\_\_ x 8.33 = \_\_\_\_\_  
x .10 = \_\_\_\_\_

Total for Characteristic = \_\_\_\_\_  
x .11 =

\* Precisely, by the number of elements present, which varies.

56

CHARACTERISTIC II:

Subfactor 1: Co \_\_\_ + Cl \_\_\_ = \_\_\_ x 12.5 = \_\_\_ x .43 = \_\_\_  
Subfactor 2: Co \_\_\_ + Cl \_\_\_ = \_\_\_ x 12.5 = \_\_\_ x .32 = \_\_\_  
Subfactor 3: Co \_\_\_ + Cl \_\_\_ = \_\_\_ x 12.5 = \_\_\_ x .25 = \_\_\_

Total for Characteristic = \_\_\_\_\_

x .15 =

CHARACTERISTIC III:

Subfactor 1: Co \_\_\_ + Cl \_\_\_ = \_\_\_ x 12.5 = \_\_\_ x .39 = \_\_\_  
Subfactor 2: Co \_\_\_ + Cl \_\_\_ = \_\_\_ x 12.5 = \_\_\_ x .39 = \_\_\_  
Subfactor 3: Co \_\_\_ + Cl \_\_\_ = \_\_\_ x 12.5 = \_\_\_ x .22 = \_\_\_

Total for Characteristic = \_\_\_\_\_

x .15 =

CHARACTERISTIC IV:

Subfactor 1: Co \_\_\_ + Cl \_\_\_ + Ap \_\_\_ = \_\_\_ x 8.33 = \_\_\_  
x .105 = \_\_\_\_\_  
Subfactor 2: Co \_\_\_ + Cl \_\_\_ + Ap \_\_\_ = \_\_\_ x 8.33 = \_\_\_  
x .105 = \_\_\_\_\_  
Subfactor 3: Co \_\_\_ + Cl \_\_\_ = \_\_\_ x 12.5 = \_\_\_ x .19 = \_\_\_  
Subfactor 4: Co \_\_\_ + Cl \_\_\_ + Ap \_\_\_ = \_\_\_ x 8.33 = \_\_\_  
x .19 = \_\_\_\_\_  
Subfactor 5: Ap \_\_\_ x 25.0 = \_\_\_ x .15 = \_\_\_\_\_  
Subfactor 6: Co \_\_\_ + Cl \_\_\_ + Ap \_\_\_ = \_\_\_ x 8.33 = \_\_\_  
x .10 = \_\_\_\_\_  
Subfactor 7: Co \_\_\_ + Cl \_\_\_ + Ap \_\_\_ = \_\_\_ x 8.33 = \_\_\_  
x .06 = \_\_\_\_\_  
Subfactor 8: Co \_\_\_ x 25.0 = \_\_\_ x .10 = \_\_\_\_\_

Total for Characteristic = \_\_\_\_\_

x .09 =

CHARACTERISTIC V:

Subfactor 1: Co \_\_\_ + C1 \_\_\_ = \_\_\_ x 12.5 = \_\_\_ x .16 = \_\_\_  
 Subfactor 2: Co \_\_\_ + C1 \_\_\_ = \_\_\_ x 12.5 = \_\_\_ x .08 = \_\_\_  
 Subfactor 3: Co \_\_\_ + C1 \_\_\_ = \_\_\_ x 12.5 = \_\_\_ x .08 = \_\_\_  
 Subfactor 4: Co \_\_\_ + C1 \_\_\_ = \_\_\_ x 12.5 = \_\_\_ x .10 = \_\_\_  
 Subfactor 5: Co \_\_\_ + C1 \_\_\_ = \_\_\_ x 12.5 = \_\_\_ x .05 = \_\_\_  
 Subfactor 6: Co \_\_\_ + C1 \_\_\_ = \_\_\_ x 12.5 = \_\_\_ x .05 = \_\_\_  
 Subfactor 7: C1 \_\_\_ x 25.0 = \_\_\_ x .16 = \_\_\_  
 Subfactor 8: C1 \_\_\_ + Ap \_\_\_ = \_\_\_ x 12.5 = \_\_\_ x .16 = \_\_\_  
 Subfactor 9: Co \_\_\_ + C1 \_\_\_ = \_\_\_ x 12.5 = \_\_\_ x .16 = \_\_\_

Total for Characteristics = \_\_\_\_\_

x .11 =

CHARACTERISTIC VI:

Subfactor 1: Co \_\_\_ + C1 \_\_\_ + Ap \_\_\_ = \_\_\_ x 8.33 = \_\_\_  
 x .23 = \_\_\_\_\_  
 Subfactor 2: Ap \_\_\_ x 25.0 = \_\_\_ x .13 = \_\_\_\_\_  
 Subfactor 3: Co \_\_\_ + C1 \_\_\_ + Ap \_\_\_ = \_\_\_ x 8.33 = \_\_\_\_\_  
 x .13 = \_\_\_\_\_  
 Subfactor 4: Co \_\_\_ + C1 \_\_\_ = \_\_\_ x 12.5 = \_\_\_ x .13 = \_\_\_\_\_  
 Subfactor 5: Co \_\_\_ + C1 \_\_\_ + Ap \_\_\_ = \_\_\_ x 8.33 = \_\_\_\_\_  
 x .16 = \_\_\_\_\_  
 Subfactor 6: Co \_\_\_ + C1 \_\_\_ = \_\_\_ x 12.5 = \_\_\_ x .16 = \_\_\_\_\_  
 Subfactor 7: Co \_\_\_ + C1 \_\_\_ = \_\_\_ x 12.5 = \_\_\_ x .06 = \_\_\_\_\_

Total for Characteristic = \_\_\_\_\_

x .10 =

CHARACTERISTIC VII: Co \_\_\_\_\_ + C1 \_\_\_\_\_ = \_\_\_\_\_ x 12.5

Total for Characteristic = \_\_\_\_\_

x .10 =

CHARACTERISTIC VIII: Co \_\_\_\_\_ + Cl \_\_\_\_\_ = \_\_\_\_\_ x 12.5

Total for Characteristic = \_\_\_\_\_

x .10 =

CHARACTERISTIC IX: Co \_\_\_\_\_ + Cl \_\_\_\_\_ + Ap \_\_\_\_\_ = \_\_\_\_\_ x 8.33

Total for Characteristic = \_\_\_\_\_

x .09 =

SUMMARY (OVERVIEW) SCORE FOR REPORT

	<u>Weighted Score</u> <input type="text"/>
Characteristic I	_____
Characteristic II	_____
Characteristic III	_____
Characteristic IV	_____
Characteristic V	_____
Characteristic VI	_____
Characteristic VII	_____
Characteristic VIII	_____
Characteristic IX	_____
Total Score =	<input type="text"/>

#535

59

## APPENDIX F

### List of Suggested Readings

- Bryant, Coralie and Louise G. White, Managing Development in the Third World.  
1982: Westview Press, Boulder, Colorado.
- TRITON Corporation, Analysis of the Quality of FY80-FY82 AID Evaluation Reports.  
1984: unpublished, Washington, D.C.
- TRITON Corporation, Development of a Quality/Completeness Scoring Instrument for USAID Evaluations. 1982: unpublished, Washington, D.C.
- TRITON Corporation, Findings Compendium and Analysis of FY82 AID Evaluation Reports. 1984: unpublished, Washington, D.C.
- TRITON Corporation, Findings Compendium of the FY83 Metaevaluation.  
1984: unpublished, Washington, D.C.
- USAID, Effective Institution Building: A Guide for Project Managers Based on Lessons Learned from the AID Portfolio. Discussion Paper No. 11,  
1982: PPC/E, Washington, D.C.
- USAID, Manager's Guide to Data Collection.  
1979: PPC/E, Washington, D.C.
- USAID, Design and Evaluation of AID-Assisted Projects.  
1974: Office of Evaluation, Washington, D.C.