

PD - AAR - 170

39645

FINAL REPORT

ANALYSIS OF THE DISTRIBUTION
OF FINDINGS DRAWN FROM THE
FY83 META-EVALUATION PROJECT

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

Conducted for PPC/E/PES

0001365
584

November 1984

TRITON

November 30, 1984

Ms. Nena Vreeland
Technical Officer
PPC/CDIE
Room 3659, New State
Agency for International
Development
Washington, D.C. 20523

Dear Ms. Vreeland:

TRITON Corporation is pleased to submit the report entitled, "Analysis of the Distribution of Findings Drawn from the FY83 Metaevaluation Project." This report contains detailed analyses of the distribution of findings, in particular with regards to three external variables: bureau, technical code and time of the evaluation. In addition, the issues of sustainability and coder bias are addressed, as well as a comparison (time series analysis) of the overall distribution of findings versus those from the FY82 Metaevaluation Project.

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

Sincerely,



Sonny S. Bloom
Vice President

Enclosure

SSB: 690

TRITON

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

TABLE OF CONTENTS

<u>Section</u>	<u>Page</u>
I. Introduction to the Findings Report.....	1
Definition of Terms.....	1
II. Analysis of the Distribution of Findings.....	3
Category Distribution.....	3
External Variable: Bureau.....	6
External Variable: Technical Activities.....	8
External Variable: Time of Evaluation.....	13
Coder Bias.....	18
Sustainability.....	22
Time Series Analysis.....	32
Summary of Findings.....	33
III. Recommendations.....	35
Evaluations.....	35
Analysis of Findings.....	38
IV. Methodology.....	40
Overview of the Project.....	40
Findings Categorization Scheme.....	42
Definitions and Examples.....	43
Data Procedures.....	61
Figure IV-1.....	Following Page 43

Appendices

Figure IV-1

DESIGN

Overly ambitious objectives
Conflicting objectives
Failed assumptions
Missing inputs and outputs
Scheduling and budget
Recommendations and planned changes.

IMPLEMENTATION

Problems finding U.S. contractors and personnel

Problems finding host country contractors and personnel

Commitment and performance of host country contractors, government and personnel

Commitment and performance of both U.S. and host country contractors and personnel

AID reporting requirements

Contracting and funding procedures

Coordination between AID and host countries

Procurement of commodities

Delay litanies

Coordination between AID and contractor.

INSTITUTION-BUILDING

Progress at the central level
Progress with decentralization
Progress at the community level
Progress with training
Problems with self-sufficiency and recurring costs
Problems with strategies and structures
Training problems.

DATA MANAGEMENT

Collection and analysis
Plans developed via that analysis
Disseminating information.

IMPACT

Production impact
Economic impact
Social impact
Spread/imitation effects.

I. INTRODUCTION

This report presents the results of the qualitative aspect of the FY83 "Metaevaluation Project," and describes and analyzes the findings produced during the application of TRITON's quantitative scoring tool. Analyses were performed on the distribution of these findings based on three key external factors and two conceptual issues. The external factors comprise bureau, technical activity and time of the evaluation. The two conceptual issues are: 1) an analysis of the occurrence of sustainability and its strength with the distribution of findings; and 2) a time-series analysis of the distributions found in the FY82 versus the FY83 metaevaluation projects.^{1/}

Definitions of Terms

Bureau is the term used to generically describe AID's missions, regional and central offices. It refers to either a geographic or content grouping. Missions and regional offices are grouped by their respective geographic bureau. Evaluation reports from central offices (i.e., Science and Technology and Food and Voluntary Assistance) are termed "bureaus" for the purposes of this analysis. Impact evaluations are treated as a separate bureau. There are four regional bureaus and three central bureaus. These are:

REGIONAL BUREAUS

Near East
Asia
Latin America and the Caribbean
Africa

CENTRAL BUREAUS

Impact
Science and Technology
Food and Voluntary
Assistance

^{1/} See: Section 3 "Methodology" of this report for an overview of the project.

"Latin America and the Caribbean" is abbreviated as "LatCarib."
"Science and Technology" is abbreviated as "SciTech", and "Food and Voluntary Assistance" as "FVA."

Technical Activity refers to AID's classification scheme for project activities. Every project is assigned a three-digit identification code. For the purposes of the computer analysis, these have been condensed to ten headings, based on the initial digit of that code and which also reflects AID's principal topics. These headings are (abbreviations used appear in parentheses):

Agriculture
Rural Non-Agricultural Activities (Rural Non-Agric.)
Rural Multi-Function (Rural-Multi)
Nutrition
Population
Health
Education
Human Resource Development (HRD)
Infrastructure, Industry and Housing (Infrastructure)
Other Activities (Other).

Time of the Evaluation provides an indication of what stage of the project cycle was being evaluated. "Interim evaluation" reports are those performed in the first years of the project and are used to assess progress. "Final evaluation" reports are those performed at the end of the project and are used to assess results. "Ex post evaluation" reports are those performed after the project has been completed and are used to assess impact. The category of "Other" includes primarily special studies, which are outside the project cycle and are used to perform further research on topics of interest.

Sustainability refers to the addressing of that topic in the evaluation report and to the degree of sustainability attributed to the project by the evaluation report. This is analyzed in this report by all three of the above-mentioned external factors, major headings, cost of the evaluation, and total cost of the project.

II. DISTRIBUTION OF FINDINGS BY CATEGORIES

Table 1 presents the overall distribution of findings by individual categories. These categories were derived inductively by a consultant^{1/} to the FY82 Metaevaluation project. This inductive process entailed three steps: 1) reading the findings reported by the coders for every evaluation report; 2) grouping the findings by like topics; and 3) developing generic names for those groupings. The categories reflect all of the five major stages of the project cycle and the findings are analyzed by both the thirty-one individual categories and by those project stages. The five stages, called "major headings" in the analysis, are: design, implementation, institution-building, data and impact.

Findings categories which account for more than five percent of the total are listed below in Table 2 and provide a focus for the subsequent analysis:

Table 2: List of Categories Over Five Percent

<u>Categories</u>	<u>N</u>	<u>Pct.</u>
Recommendations and Planned Changes	165	13.6
Commitment and Performance of Host Country Government, Contractors and Personnel	87	7.1
Strategies and Structures	80	6.6
Progress at the Community Level	71	5.8
Failed Assumptions	68	5.6
Collection and Analysis of Data	67	5.5
Commitment and Performance of U.S. Contractors and Personnel	66	5.4
Scheduling and Budget	<u>65</u>	<u>5.3</u>
TOTAL	669	54.9

^{1/} Mr. Roger Popper.

Table 1: Findings Category Frequencies & Percentages

Categories		Frequencies	Percentages (rounded to nearest tenth)
DESIGN	Overly ambitious objectives	40	3.3
	Conflicting objectives	20	1.6
	Failed assumptions	68	5.6
	Missing inputs and outputs	28	2.3
	Scheduling & budget	65	5.3
	Recommendations & planned changes	165	13.6
IMPLEMENTATION	Problems finding US contractors & personnel	14	1.2
	Problems finding host country contractors & personnel	14	1.2
	Commitment & performance of US contractors & personnel	66	5.4
	Commitment & performance of host country gov't, contractors & personnel	87	7.1
	Commitment & performance of both US and host country contractors & personnel	56	4.6
	AID reporting requirements	5	0.4
	Contracting & funding procedures	16	1.3
	Coordination between AID & contractor	13	1.1
	Procurement of commodities	21	1.7
	Delay litanies	34	2.8
INSTITUTION BUILDING	Coordination between AID & contractor	23	1.9
	Progress at the central level	23	1.9
	Progress with decentralization	12	1.0
	Progress at the community level	71	5.8
	Progress with training	53	4.4
	Self-sufficiency & recurring costs	26	2.1
	Strategies & structures	80	6.6
Training problems	25	2.1	
DATA	Collection & analysis of data	67	5.5
	Plans developed via that analysis	15	1.2
	Disseminating information	6	0.5
IMPACT	Production impact	29	2.4
	Economic impact	31	2.5
	Social impact	37	3.0
	Spread/imitation effects	7	0.6
TOTAL		1217	100.0

"Recommendations and Planned Changes," "Failed Assumptions and Scheduling," and "Budget" all occur under the major heading of "Design." "Commitment and Performance of Host Country Government, Contractors and Personnel" and "Commitment and Performance of U.S. Contractors and Personnel" both occur under the major heading of Implementation. "Strategies and Structures" and "Progress at the Community Level" occur under the major heading of "Institution-Building." "Collection and Analysis of Data" occurs in the major heading of "Data." There is no category accounting for over 5% of the total for the major heading of "Impact."

These categories were then summed under their respective major headings to generate Table 3:

	<u>Sums from Table 2</u>	<u>Overall Total</u>	<u>Relative Pct.</u>
Design	24.5	31.7	77.3
Implementation	12.5	28.7	43.6
Institution-Building	12.4	23.9	51.9
Data	5.5	7.2	76.4
Impact	----	8.5	----
TOTAL	54.9	100.0	----

Table 3 clearly depicts that the categories which produced more than 5% of total findings dominate both major headings of "Design" and "Data," and account for just over half of the findings generated by "Institution-Building."

The conclusion drawn from the above frequencies is that the majority of the problems identified for the FY 83 evaluation reports occurred in the major heading of "Design." Over fifty (50) percent of these findings were revisions to the existing project. Since more than sixty percent (60%) of the evaluation

reports reviewed are interim reports, this indicates both significant problems with the task of designing projects and responsiveness to the timely revision of those designs.

One category -- "Collection and Analysis of Data" -- accounted for more than 75% of all findings for the major heading of "Data," but the overall frequency of this heading is relatively small so that disproportionality occurs with fairly low numbers. This may indicate that data are not collected by a significant number of the various bureaus in sufficient quantity, or perhaps that the various technical activities did not require a concentration of that type of activity.

Two categories in the major heading of "Institution Building" -- "Progress at the Community Level" and "Strategies Structures" -- accounted for over fifty percent (50%) of the total findings. There are slightly more problems than successes with regard to these two categories, but the overall success/problem percentage masks that individual difference, as there are 13.1% reflecting progress and 10.8% reflecting problems.

Distribution of Findings by Bureau

The graph of "Distribution of Findings by Bureaus by Percentages" depicted on the next page shows the percentages of findings identified broken out by bureau. Some key observations which can be made from this graph are:

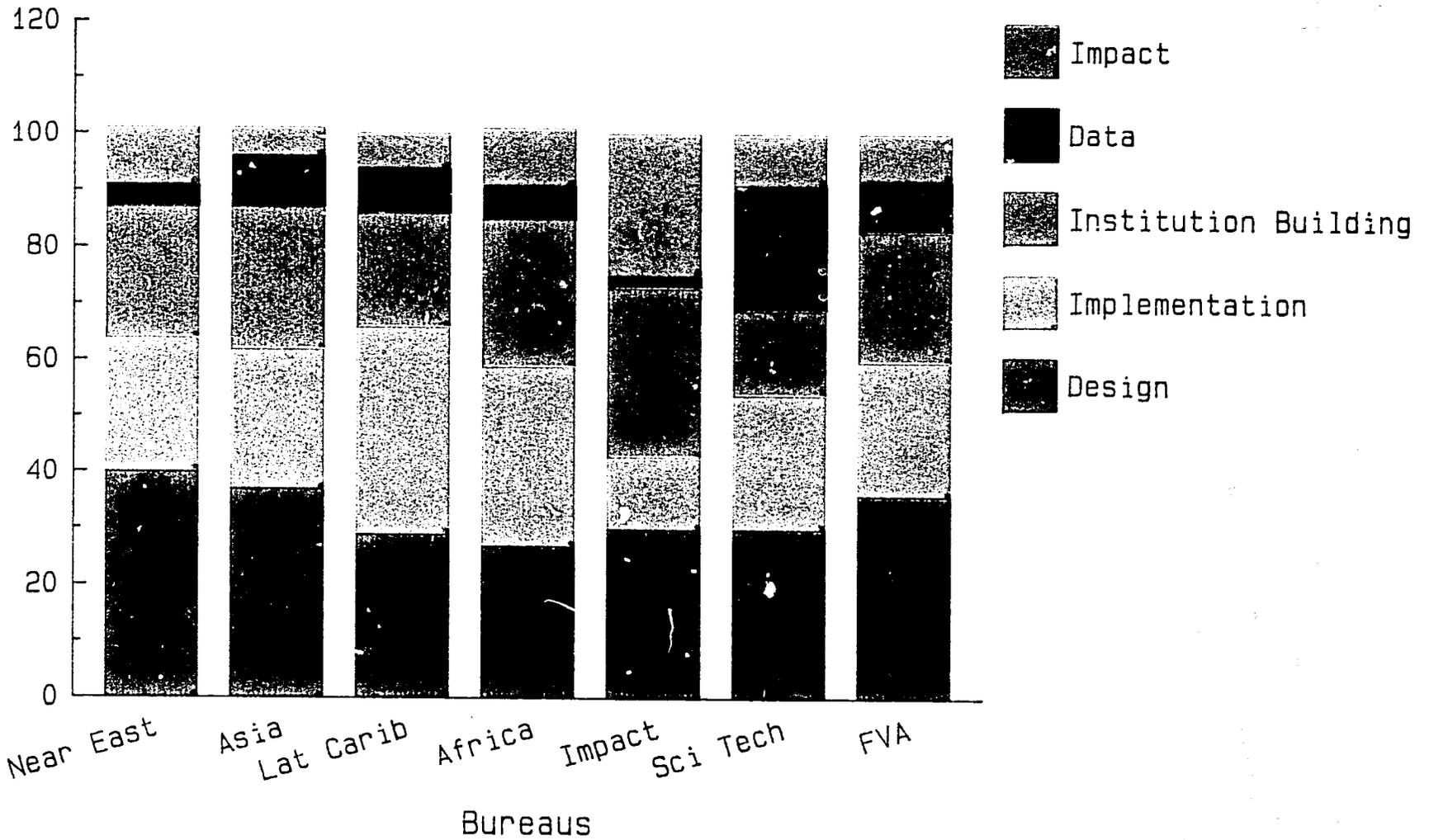
- o Impact has relatively very few "Data" findings and even fewer "Implementation" findings.
- o Impact has a relatively large percentage of "Impact" and "Institution-Building" findings.
- o Science & Technology (SciTech) has fewer Institution-Building findings than any other bureau.
- o Science & Technology has relatively more "Data" findings than any other bureau.

The values which produced the preceding graph are presented in Table 4, below.

Table 4: Bureau by Major Grouping						
Bureau	Major Grouping (NUMBER)					
	Design	Implemen- tation	Institu- tion- Building	Data	Impact	Total
Near East	67 39.9 17.4	41 24.4 11.8	38 22.6 13.1	6 3.6 6.8	16 9.5 15.4	168 13.8
Asia	94 36.6 24.4	65 25.3 18.6	64 24.9 22.1	22 8.6 25.0	12 4.7 11.5	257 21.1
Lat/Carib	46 29.1 11.9	59 37.3 16.9	31 19.6 10.7	13 8.2 14.8	9 5.7 8.7	158 13.0
Africa	120 26.6 31.1	146 32.3 41.8	116 25.7 40.0	27 6.0 30.7	43 9.5 41.4	452 37.1
Impact	16 30.2 4.2	7 13.2 2.0	15 30.2 5.5	1 1.9 1.1	13 24.5 12.5	53 4.4
SciTech	16 29.6 4.2	13 24.1 3.7	8 14.8 2.8	12 22.2 13.6	5 9.3 4.8	54 4.4
FVA	27 36.0 7.0	18 24.0 5.2	17 22.7 5.9	7 9.3 8.0	6 8.0 5.8	75 6.2
TOTAL	386 31.7	349 28.7	290 23.8	88 7.2	104 8.6	1217 100.0

Table 4 reveals that, in essence, there is little differentiation among the regional bureaus. The observations made about two

DISTRIBUTION OF FINDINGS BY BUREAUS BY PERCENTAGES



central bureau evaluation report types -- Impact and Science & Technology -- are strongly correlated to particular agendas of those bureaus: Impact evaluations are ex-post evaluations, examining the long-term effects of a project's implementation and are, therefore, strongest in the major headings which correspond to both that timeframe (i.e., "Design" and "Implementation" are completed at that stage of project cycle, while "Data" have already been processed and analyzed) and to that agenda (i.e., examining what effect the project had in that area). Science & Technology does not generally conduct projects, per se. That Office conducts research and development of particular techniques which are documented in a reasonably scientific manner. The only "Institution-Building" being performed would center around the establishment of research stations and the concurrent training of their personnel.

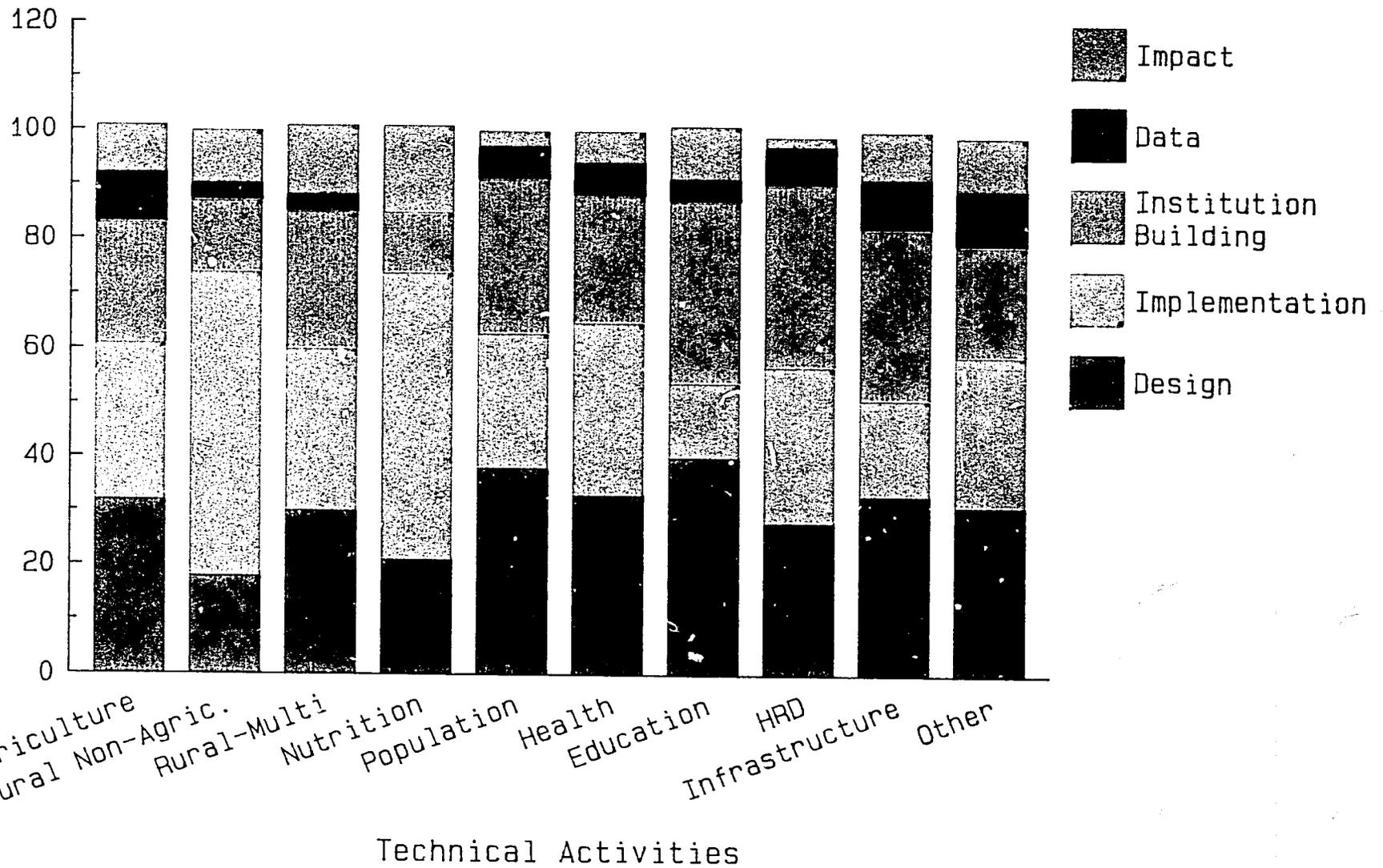
The relative percentages of findings produced by the bureaus correspond closely to the number of evaluation reports reviewed/scored for each bureau. This is presented in Table 5, and indicates that no bureau produced a disproportionate number of findings:

Table 5: Bureau Findings and Report Percentages		
Bureau	Findings Pct.	Report Pct.
Near East	13.8	13.1
Asia	21.1	21.2
Lat/Carib.	13.0	12.7
Africa	37.1	28.0
Impact	4.4	3.2
SciTech	4.4	6.5
FVA	6.2	5.3

Distribution of Findings by Technical Activities

The graph shown on the next page, entitled "Distribution of Findings in Technical Activities by Percentages," produces some interesting observations. These are:

DISTRIBUTION OF FINDINGS IN TECHNICAL ACTIVITIES BY PERCENTAGES



(1)

- o Rural Non-Agricultural Activities and Nutrition have relatively more Implementation findings than any other technical activity.
- o Education has relatively, the most Design and fewest Implementation findings of any technical activity.
- o Nutrition has the lowest frequency of Data and Institution-Building findings of any technical activity.
- o Population and HRD produce the lowest frequency of Impact findings of any technical activity.
- o All of the major headings of findings vary more by technical activities than they do by bureau. This is a key observation and one which will be elaborated in the summary of this section.

The observations which generated the above-referenced graph are represented numerically in Table 6, which demonstrates the actual frequencies of the findings used in the graph. The total frequency values produce the following rank order:

Design (highest frequency of findings)
 Implementation
 Institution-Building
 Impact
 Data (lowest frequency of findings).

This rank ordering is identical for the individual technical activities of "Agriculture," "Rural Multi-Function," "Health," and "Other." Reversals in rank order for "Implementation" and "Institution-Building" and/or "Data and Impact" are exhibited for "Population," "Education," "HRD," and "Infrastructure." The only exceptions to these two patterns are "Rural Non-Agriculture" and "Nutrition," which reverse the frequency of findings accounted for by "Design" and "Implementation."

This indicates that the technical activity apparently accounts for a stronger determination of the types of findings produced than does the bureau. It should, therefore, be possible to draw a profile of the relative strengths and weaknesses of the

Table 6: Technical Activities by Major Headings						
Technical Activity	Major Heading					
	Design	Implementation	Institution-Building	Data	Impact	Total
	(NUMBER)	(NUMBER)	(NUMBER)	(NUMBER)	(NUMBER)	(NUMBER)
	(ROW PCT)	(ROW PCT)	(ROW PCT)	(ROW PCT)	(ROW PCT)	(ROW PCT)
	(COL PCT)	(COL PCT)	(COL PCT)	(COL PCT)	(COL PCT)	(COL PCT)
Agriculture	137 32.2 35.5	123 28.9 35.2	93 21.9 32.1	36 8.5 40.9	36 8.5 34.6	425 34.9
Rural Non-Agric	11 18.0 2.9	34 55.7 9.7	8 13.1 2.8	2 3.3 2.3	6 9.8 5.8	61 5.0
Rural-Multi	19 29.7 4.9	19 29.7 5.4	16 25.0 5.5	2 3.1 2.3	8 12.5 7.7	64 5.3
Nutrition	4 21.1 1.0	10 52.6 2.9	2 10.5 0.7	0 0.0 0.0	3 15.8 2.9	19 1.6
Population	12 37.5 3.1	8 25.0 2.3	9 28.1 3.1	2 6.3 2.3	1 3.1 1.0	32 2.6
Health	42 33.3 10.9	40 31.8 11.5	29 23.0 10.0	7 5.6 8.0	8 6.4 7.4	126 10.4
Education	41 39.8 10.6	14 13.6 4.0	34 33.0 11.7	4 3.9 4.6	10 9.7 9.6	103 8.5
HRD	23 28.1 6.0	24 29.3 6.9	27 32.9 9.3	6 7.3 6.8	2 2.4 1.9	82 6.7
Infrastructure	30 33.3 7.8	16 17.8 4.6	28 31.1 9.7	8 8.9 9.1	8 8.9 7.7	90 7.4
Other	67 31.2 17.4	61 28.4 17.5	44 20.5 15.2	21 9.8 23.9	22 10.2 21.2	215 17.7
TOTAL	386 31.7	349 28.7	290 23.8	88 7.2	104 8.6	1217 100.0

technical activities by the types of findings they generate. This is presented in Table 7.

Table 7: Assessment of Strengths/Weakness of the Technical Activities					
Activity	Design	Implementation	Institution Building	Data	Impact
Agriculture	N/A	N/A	N/A	N/A	N/A
Rural Non-Agric.	+	-	-	-	N/A
Rural Mult.	N/A	N/A	N/A	-	+
Nutrition	+	-	-	-	+
Population	-	+	+	N/A	-
Health	N/A	-	N/A	N/A	-
Education	-	+	+	-	N/A
HRD	+	N/A	+	N/A	-
Infrastructure	N/A	+	+	N/A	N/A
Other	N/A	N/A	-	+	N/A

+ = strength
 - = weakness
 N/A = difference does not exceed 2%.

Evaluations are weakest, overall, in the major heading of Data and strongest in the major heading of Institution-Building.

The analytical framework for this table was developed as follows:

- Step 1. The percentage that each findings category was accounted for by a given technical activity was calculated (for example, 40%^{1/} of all Design findings accounted for by the technical activity Nutrition).
- Sept 2. This percentage from Step 1 was compared to the overall percentage of findings accounted for by that findings category (for example, the 40% from Step 1 compared to Design accounting for 60% of all findings).
- Sept 3. If the technical activity - percentage specific from Step 1 was greater than 2%, on an absolute basis, from the overall percentage from Step 2, then that cell of Table 7 was denoted with a "+", indicating that the findings profile in that technical activity for that findings category was relatively a "strength" for that technical activity, or a "-", indicating the profile to represent a relative "weakness." The denoting of a "+" versus a "-" depended on the findings category under analysis, as depicted below:

Overall Nature/ Tone of Findings	Design	Implemen- tation	Institu- tion- Building	Data	Impact
	Negative	Negative	Positive	Positive	Positive
If Tech. Act. Shows Findings Category % That Is 2% or More <u>Greater</u> Than Overall Findings Category %, Cell Noted As:	Weakness (-)	Weakness (-)	Strength (+)	Strength (+)	Strength (+)
If Tech. Act. Shows Findings Category % That Is 2% or More <u>Less</u> Than Overall Findings Category %, Cell Noted As:	Strength (+)	Strength (+)	Weakness (-)	Weakness (-)	Weakness (-)

(For example, because Design findings are generally negative as to project performance, and the Design - Nutrition percentage is 20% less than the overall percentage of Design findings (40% vs. 60%), then Nutrition's profile is relatively a strength as regards to extent of Design (negative) findings.)

^{1/} This number is used for illustrative purposes only.

When the relative percentage of findings produced by each technical activity and the number of evaluation reports scored is compared to determine if any disproportionate frequencies exist, the results indicate clearly that no technical activity produced a disproportionate amount of findings. These data are presented in Table 8:

Table 8: Technical Activity Findings and Report Percentages		
Technical Activity	Findings Pct.	Report Pct.
Agriculture	34.9	31.8
Rural Non-Agriculture	5.0	4.9
Rural Multi-Function	5.3	6.9
Nutrition	1.6	2.0
Population	2.6	3.3
Health	10.4	11.0
Education	8.5	7.8
HRD	6.7	6.1
Infrastructure	7.4	6.9
Other	17.7	19.2

Only three technical activities differ in findings frequency by more than one percent from the report frequency: "Rural Multi-Function" and "Other" tend to generate slightly fewer findings, while "Agriculture" tends to generate slightly more. These differences, however, are not significant.

Distribution of Findings by Time of Evaluation

Evaluation reports reviewed in the FY83 metaevaluation project comprised four types: interim, final, ex-post and other evaluations. "Interim" evaluations are those conducted in the early stages of project implementation, and are usually intended to assess the progress of the project and to propose any revisions which may be necessary. "Final" evaluations are those conducted

as the project is ending and are performed to assess achievement and, to a certain extent, the impact of the project. "Ex-post" evaluations are those conducted after a project has been completed and are designed to assess the long-term impacts of that project. The "Other" type is generally not related to a specific project cycle. The most common type of "other" time of evaluation report was a study testing a new methodology or conducting a specific analysis of certain aspects of some project/program.

The graph of the "Distribution of Findings by Time of Evaluation by Percentages" (shown on the next page) clearly demonstrates the differences in focus of both "Ex-post" and "Other" evaluation reports. Neither contain relatively as many Design or Implementation findings as "Interim" or "Final." "Ex-post" contained relatively more Institution-Building and Impact findings than any other, as well as far fewer Data. "Other" accounts for a larger percentage of all Data findings than report type do all the other categories combined.

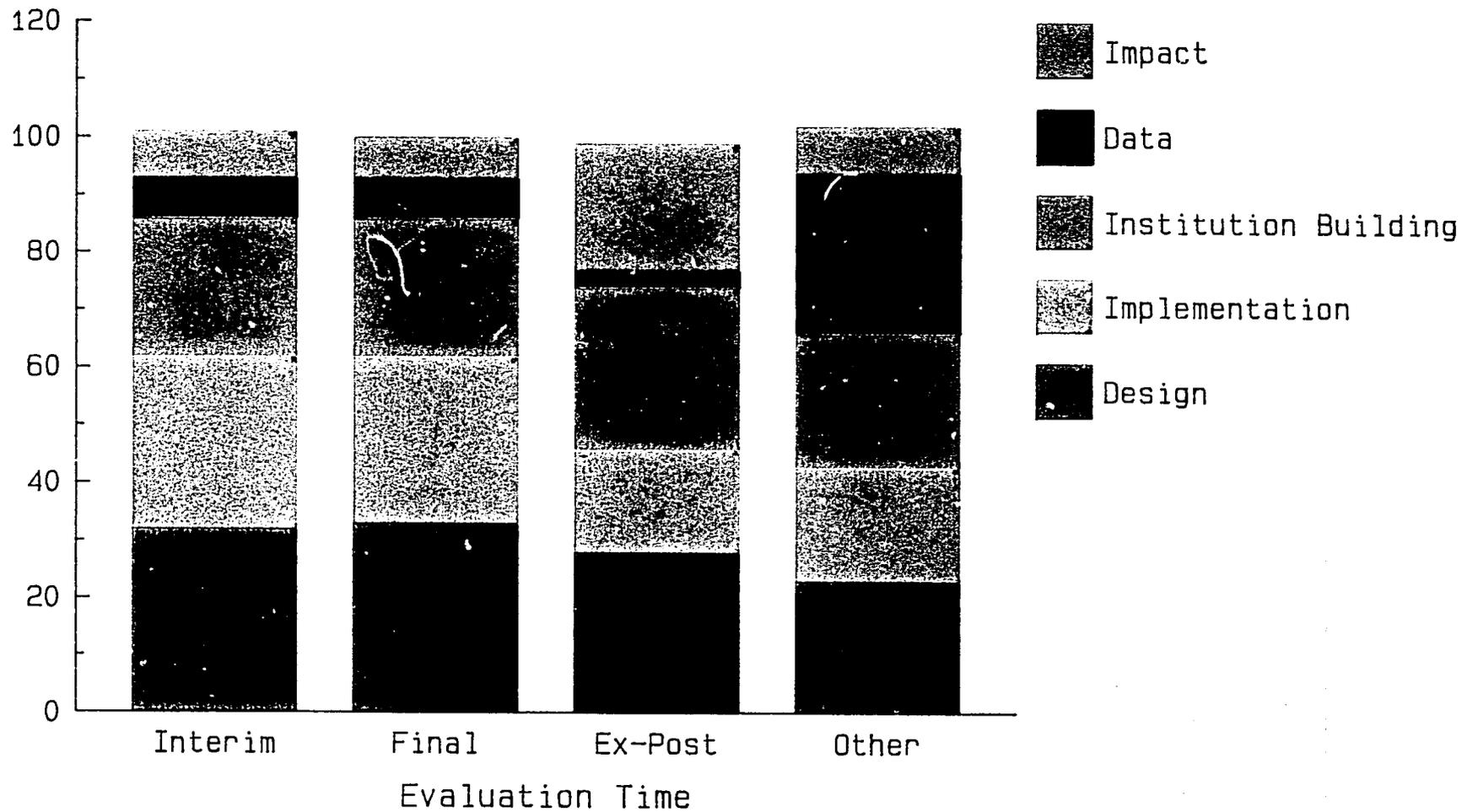
Table 9 presents this information numerically. It becomes evident that there are two relationships meriting further analysis. The first is the proportion of Institution-Building findings comparing Interim versus Final evaluations. The second is the extremely low percentage of Data findings for Ex-post evaluation reports.

Table 9: Table of Time of Evaluation by Major Grouping						
Time of Evaluation	Major Grouping (NUMBER)					
	Design	Implemen- tation	Institu- tion- Building	Data	Impact	Total
	(ROW PCT)	(ROW PCT)	(ROW PCT)	(COL PCT)	(COL PCT)	(COL PCT)
Interim	257 31.9 66.6	240 29.7 68.8	191 23.7 65.9	54 6.7 61.4	65 8.1 62.5	807 66.3
Final	103 33.2 26.7	90 29.0 25.8	73 23.6 25.2	21 6.8 23.9	23 7.4 22.1	310 25.5
Ex-post	17 28.3 4.4	11 18.3 3.2	17 28.3 5.9	2 3.3 2.3	13 21.7 12.5	60 4.9
Other	9 22.5 2.3	8 20.0 2.3	9 22.5 3.1	11 27.5 12.5	3 7.5 2.9	40 3.3
TOTAL	386 31.7	349 28.7	290 23.8	88 7.2	104 8.6	1217 100.0

The apparent anomaly concerning Institution-Building is that, Interim and Final evaluation reports should produce such similar proportions when theoretically, the expectation would be for Final evaluation reports, to generate a relatively higher frequency, since they are focused on the latter stages of project cycle.

There are several hypotheses for this result. The first hypothesis is that more Interim evaluation reports than might be expected are examining what effect even the earlier stages of the project cycle may have on a targeted population. The second hypothesis is that this observation is a result of the categorization

DISTRIBUTION OF FINDINGS IN EVALUATION TIME BY PERCENTAGES



15/11

scheme, as Institution-Building contains both positive and negative categories, that is, both progress and problems, so that an interim evaluation report might examine more problems, but that difference would be subsumed within the major heading. The third hypothesis is that this observation is a result of the data; i.e., there are more Interim findings of an Institution-Building nature than might be expected simply because of the fact that interim findings account for 66.3% of the overall total while final findings make up only 25.5%.

Table 10 demonstrates that there are no major differences in the overall percentage of findings produced by Interim evaluation reports as opposed to all others. There is a slight tendency for non-Interim evaluation reports to produce more Institution-Building, Data and Impact. This would support the subsumption of trends predicted by the third hypothesis presented above.

Time of Evaluation	(NUMBER)					
	Major Heading (ROW PCT) (COL PCT)					
	Design	Implemen- tation	Institu- tion- Building	Data	Impact	Total
Interim	257 31.9 66.6	240 29.7 68.8	191 23.7 65.9	54 6.7 61.4	65 8.1 62.5	807 66.3
All Others	129 31.5 33.4	109 26.6 31.2	99 24.2 34.1	34 8.3 38.6	39 9.5 37.5	410 33.7
Total	386 31.7	349 28.7	290 23.8	88 7.2	104 8.6	1217 100.0

The first and second hypotheses presented above were examined by developing a table of "evaluation" time by individual categories. The results indicated that while there were more Interim evaluation reports than all others combined, a larger percentage of Institution-Building findings were generated for Interim reports only for the category referring to "Progress with Training." When the categories for "Progress" versus "Problems" are compared, non-Interim evaluations account for a much larger proportion of findings under the "Problems" headings for the individual categories.

Thus, it appears that interim evaluation reports are not only examining the negative aspects of Institution-Building, but are incorporating all aspects into the evaluation report, but at a relatively earlier stage in the project cycle.

Table 11				
Categories (Problems)	INTERIM	FINAL	EX-POST	OTHER*
Strategies and self-sufficiency	4.3	4.8	1.7	5.0
Recurring Costs	1.4	3.2	5.0	5.0
Problems with Training	5.8	7.1	16.7	2.5

* Percent of total findings.

It would seem that the relative proportionality of Interim versus non-Interim findings is an artifact of the categorization scheme and of the preponderance of Interim evaluation reports. There were 169 Interim evaluation reports (69% of all reports), which accounted for 70% or more of all evaluation reports read for every single Bureau except "Impact" and "Food and Voluntary Assistance,"

Interim reports accounted for 9% and 30.8% of these two bureaus, respectively.

Coder Bias In Evaluation Reports Scoring Results

Coder bias was assessed by conducting an analysis of inter-rater reliability. This serves as a means to verify that the coders/scorers are examining the same aspects of the evaluation reports and to provide corrective measures (such as clarifying the definitions of terms) in cases where similar examinations did not occur. Analysis of that inter-rater reliability is presented in another report prepared as a result of this project, which examines evaluation scores.

Table 12 presents the distribution of major headings by the coders, while the total percentage of findings reported by each coder, which closely reflects the percentage of evaluation reports read is presented in Table 13. The observations drawn from Table 12 emphasize certain differences which need to be analyzed further. These observations are: 1) Coder 1 produced the fewest Data findings, especially in proportion to the other coders, and particularly given that the overall number of findings produced by Coder 1 was 25% greater than Coder 2; and 2) Coders 2 and 3 produced proportionately more Data findings than their number of reports reviewed would indicate. This disproportion was especially strong for Coder 2.

Table 12: Table of Coder by Major Heading						
Coder	Major Heading (NUMBER)					
	Design	Implementation	Institution-Building	Data	Impact	Total
1	157 33.6 40.7	130 27.8 37.3	114 24.4 39.3	18 3.9 20.4	48 10.3 46.2	467 38.4
2	100 28.0 25.9	103 28.9 29.5	92 25.8 31.7	37 10.3 42.1	25 7.0 24.0	357 29.3
3	41 38.3 10.6	30 28.0 8.6	20 18.7 6.9	12 11.2 13.6	4 3.7 3.9	107 8.8
4	88 30.8 22.8	86 30.1 24.6	64 22.4 22.1	21 7.3 23.9	27 9.4 26.0	286 23.5
TOTAL	386 31.7	349 28.7	290 23.8	88 7.2	104 8.6	1217 100.0

Table 13: Percentages of Findings Produced and Reports Scored by Coders		
CODER	PRODUCED PCT	SCORED PCT
1	38.4	37.1
2	29.3	28.2
3	8.8	13.9
4	23.5	20.8

It has been presented earlier in this report that SciTech projects concentrate more heavily on Data findings. It is, therefore, possible that some particular external factor (such as bureau, technical activity or time of evaluations) could produce the types of observations noted above for Coders 1-3. Coders 1 and 2 each read five (5) SciTech evaluation reports, while Coders 3 and 4 read three such reports each, for a total of 16 evaluation reports. For Coders 2 and 3, however, these represented greater percentages of the total evaluations they read than for Coders 1 and 4. This is clearly demonstrated on Table 14, and apparently accounts for the disproportionality observed above.

Table 14: Coder Distribution					
	Coder (NUMBER)				
	Coder (ROW PCT)				
	Coder (COL PCT)				
EVALTIME	1	2	3	4	TOTAL
	15	7	0	10	32
Near East	46.9	21.9	----	31.3	100.0
	16.5	10.1	----	19.6	13.1
	14	18	9	11	52
Far East	26.9	34.6	17.3	21.2	100.0
	15.4	26.1	26.5	21.6	21.2
	14	3	6	8	31
LAC	45.2	9.7	19.4	25.8	100.00
	15.4	4.3	17.6	15.7	12.7
	35	33	12	13	93
Africa	37.6	35.5	12.9	14.0	100.0
	38.5	47.8	35.3	25.5	38.0
	3	2	3	0	8
Impact	37.5	25.0	37.5	----	100.0
	3.3	2.9	8.8	----	3.3
	5	5	3	3	16
SciTech	31.3	31.3	18.8	18.8	100.0
	5.5	7.2	8.8	5.9	6.5
	5	1	1	6	13
FVA	38.5	7.7	7.7	6.2	100.0
	5.5	1.4	2.9	11.8	5.3
	91	69	34	51	245
TOTAL	37.1	28.2	13.9	20.8	100.00
	100.0	100.0	100.0	100.0	100.0

Neither technical activities nor time of evaluation indicate any contradictory evidence to the Table 14 analysis.^{1/} An examination of the coder distribution by individual category produces the following percentages shown in Table 15:

Table 15: Coder by Individual Data Categories					
Data Categories	Coder 1	Coder 2	Coder 3	Coder 4	TOTAL
Collection & Analysis	20.9	44.8	11.9	22.4	100
Plans	20.0	20.0	26.7	33.3	100
Dissemination	16.7	66.7	0.0	16.7	100

This clearly demonstrates that Coder 2 produced the greatest percent of Data findings, while Coder 3 produced a disproportionate percentage relative to that coder's volume of reports reviewed.

The operative hypothesis to explain the above data is that the nature of evaluation findings is determined to a relatively extent when a bureau's evaluative focus is sharply defined, as in the central bureaus, than by technical activity or time of evaluation, and that the metaevaluation coders do not significantly "distort" or "reorient" that focus, even when there is a pattern of one coder to producing slightly more or less findings per evaluation report than other coders.

Sustainability as an Additional External Factor

One of the major differences between the FY82 and FY83 metaevaluation projects conducted by TRITON was the deletion of

^{1/} See Appendix 1 for the analytical tables by technical activities and time of evaluation.

experimentally-oriented hypothesis-testing in the FY83 metaevaluation project and its replacement with a more qualitative analysis of "sustainability" and "resource reallocation" at various points in the project cycle.^{1/}

The issue of sustainability revolves around a particular concern with regards to the project: Will the project activities continue once external support has ceased? In essence, this focuses on "development" as an entity, as distinct from "development assistance."

The metaevaluation project for FY 83 required the evaluation report scorers to assess whether or not sustainability had been a concern of the evaluation team and, if it was, to rate how sustainable (on a 0-4 scale) those projects were considered to be.

Sustainability was only rated for 89 (36%) of the evaluation reports; i.e., only those reports which addressed this issue. Those were, in turn, analyzed by the three external factors discussed previously, as well as major headings, cost of the evaluation, and total project cost.

Tables 16-18 present the sustainability distributions by the individual external factors of bureau, technical activity and time of evaluation. Overall, the sustainability scores distribution is fairly "normal" for the total, although slightly skewed to the left (i.e., towards less sustainability). The following observations can be made about the range and patterns for each of those external factors.

^{1/} See: Analysis of the Distribution of Scores (forthcoming); Contract No. OTR-0000-C-00-3482-00.

Bureau

- o Near East, LatCarib and Africa are slightly skewed towards less sustainability.
- o Asia and the three Central Bureaus are slightly skewed towards greater sustainability.
- o SciTech and Africa are the only bureaus that generated only scores of "4" (very sustainable).
- o Africa is the only bureau that generated scores of "0" (not sustainable) and these represent more than 20% of its total.

Technical Activity

- o Agriculture, Rural Non-Agric., Population, Education, HRD, Infrastructure and Other are slightly skewed towards less sustainability.
- o Rural Multi-Function and Health are slightly skewed towards greater sustainability.

Table 16: Bureau By Sustainability

BUREAU	SUSTAINABILITY (N) (ROW PCT) (COL PCT)					TOTAL
	0	1	2	3	4	
NEAR EAST	0 0.0 0.0	6 66.7 21.4	3 33.3 10.3	0 0.0 0.0	0 0.0 0.0	9 10.1
ASIA	0 0.0 0.0	4 25.0 14.3	7 43.8 24.1	5 31.25 23.81	0 0.0 0.0	16 18.0
LATCARIB	0 0.0 0.0	6 46.2 21.4	6 46.2 20.7	1 7.69 4.76	0 0.0 0.0	13 14.6
AFRICA	9 22.0 100.0	11 26.8 39.3	9 22.0 31.0	11 26.83 52.38	1 2.4 50.0	41 46.1
IMPACT	0 0.0 0.0	0 0.0 0.0	2 66.7 6.9	1 33.33 4.76	0 0.0 0.0	3 3.41
SCITECH	0 0.0 0.0	0 0.0 0.0	1 33.3 3.5	1 33.33 4.76	1 33.3 50.0	3 3.41
FVA	0 0.0 0.0	1 25.0 3.6	1 25.0 3.5	2 50.0 9.5	0 0.0 0.0	4 4.51
TOTAL	9 10.1	28 31.5	29 32.6	21 23.6	2 2.3	89 100.0

Table 17: Technical Activities by Sustainability

TECHNICAL ACTIVITY	SUSTAINABILITY (N) (ROW PCT) (COL PCT)					TOTAL
	0	1	2	3	4	
		4 13.8 44.4	10 34.9 35.7	7 24.1 24.1	8 27.6 38.1	
AGRICULTURE						
	1 14.3 11.1	3 42.9 10.7	2 28.6 6.9	1 14.3 4.8	0 0.0 0.0	7 7.91
RURAL NON-AGRIC						
	0 0.0 0.0	0 0.0 0.0	3 60.0 10.3	2 40.0 9.5	0 0.0 0.0	5 5.6
RURAL MULTI						
	0 0.0 0.0	0 0.0 0.0	1 100.0 3.5	0 0.0 0.0	0 0.0 0.0	1 1.2
NUTRITION						
	0 0.0 0.0	1 33.3 3.6	2 66.7 6.9	0 0.0 0.0	0 0.0 0.0	3 3.41
POPULATION						
	0 0.0 0.0	2 16.7 7.1	6 50.0 20.7	4 33.3 19.1	0 0.0 0.0	12 13.5
HEALTH						
	1 16.7 11.1	2 33.3 7.1	2 33.3 6.9	1 16.7 4.8	0 0.0 0.0	6 6.7
EDUCATION						
	0 0.0	2 7.1	3 10.3	0 0.0	1 50.0	6 6.7
HRD						
	1 11.1 11.1	4 44.4 14.3	3 33.3 10.3	1 11.1 4.8	0 0.0 0.0	9 10.1
INFRA- STRUCTURE						
	2 18.2 22.2	4 36.4 14.3	0 0.0 0.0	4 36.4 19.1	1 9.1 50.0	11 12.4
OTHER						
TOTAL	9 10.1	28 31.5	29 32.6	21 23.6	2 2.3	89 100.0

Table 18: Time Of Evaluation By Sustainability						
TIME OF EVALUATION	SUSTAINABILITY (N)					TOTAL
	(ROW PCT)					
	0	1	2	3	4	
INTERIM	3	20	17	9	1	50
	6.0	40.0	34.0	18.0	2.0	56.2
	33.3	71.4	58.6	42.9	50.0	
FINAL	4	8	10	11	1	34
	11.8	23.5	29.4	32.4	2.9	38.2
	44.4	28.6	34.5	52.4	50.0	
EX-POST	0	0	2	1	0	3
	0.0	0.0	66.7	33.3	0.0	3.41
	0.0	0.0	6.9	4.8	0.0	
OTHER	2	0	0	0	0	2
	100.0	0.0	0.0	0.0	0.0	2.31
	22.2	0.0	0.0	0.0	0.0	
TOTAL	9	28	29	21	2	89
	10.1	31.5	32.6	23.6	2.3	100.0

EvalTime

- o Interim evaluations rated sustainability lower than did Final Evaluations.
- o Ex-post evaluations rated projects to be fairly sustainable, but only reported three cases for the sustainability variable.
- o "Other evaluations" (2) were rated as not sustainable.

These results can generally be interpreted to infer that regional bureaus' evaluations tend to rate projects as not sustainable, especially in the Africa bureau. The only exception is the Asia bureau, which like the Central Bureaus generated ratings tending towards more sustainability.

When examined by their technical activities, most evaluations reviewed regarding sustainability were rated not sustainable:

this is probably due to the dominance of projects from the regional bureaus across a wider spectrum of those activities.

Finally, the later the time of the evaluation in the project cycle, the greater the probability of the project being judged as sustainable.

Table 19 depicts the distribution of sustainability by the major headings. The headings of Design, Implementation and Impact tend to score lower than Institution-Building and Data. The apparent anomaly of these results versus what might be expected for Data and Impact is mitigated by the paucity of reports with Data findings that addressed sustainability (only three cases in all) and is offset further by the fact that the majority of reports with Impact findings scored a two value (somewhat sustainable).

Table 19: Major Headings By Sustainability						
MAJOR HEADINGS	SUSTAINABILITY (N)					TOTAL
	(ROW PCT)					
	0	1	2	3	4	
	4	10	10	6	0	30
DESIGN	13.3	33.3	33.3	20.0	0.0	33.7
	44.4	35.7	34.5	28.6	0.0	
	2	8	7	6	1	24
IMPLEMENTATION	8.3	33.3	29.2	25.0	4.2	27.0
	22.2	28.6	24.1	28.6	50.0	
	2	6	6	8	0	22
INSTITUTION BUILDING	9.1	27.3	27.3	36.4	0.0	24.7
	22.2	21.4	20.7	38.1	0.0	
	0	1	0	1	1	3
DATA	0.0	33.3	0.0	33.3	33.3	3.4
	0.0	3.6	0.0	4.8	50.0	
	1	3	6	0	0	10
IMPACT	10.0	30.0	60.0	0.0	0.0	11.2
	11.1	10.7	20.7	0.0	0.0	
	9	28	29	21	2	89
TOTAL	10.1	31.5	32.6	23.6	2.3	100.0

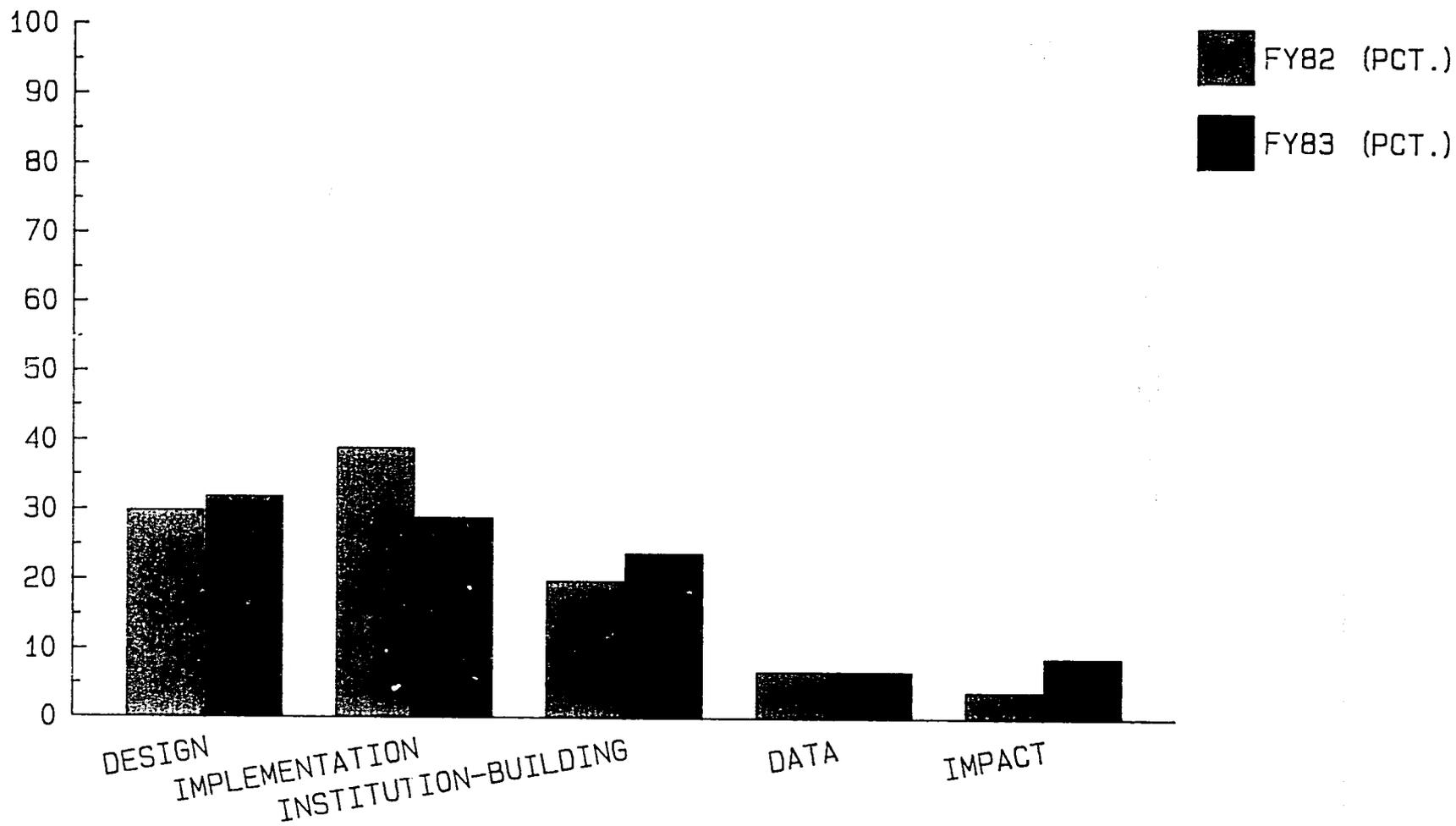
Table 20 shows the distribution of sustainability by the cost of the evaluation:

Table 20: Cost of Evaluation By Sustainability ^{1/}						
COST OF EVALUATION	SUSTAINABILITY (N)					TOTAL
	(ROW PCT)					
	0	1	2	3	4	
	1	3	3	2	0	9
\$1-4,999	11.1	33.3	33.3	22.2	---	19.6
	33.3	21.4	15.0	22.2	---	
\$5,000-9,999	0	1	5	1	0	7
	0.0	14.3	71.4	14.3	000	15.2
	0.0	7.11	25.0	11.1	---	
\$10,000-14,999	0	4	3	3	0	10
	0.0	40.0	30.0	130.0	---	21.7
	0.0	28.6	15.0	33.3	---	
\$15,000-19,999	9	0	4	6	2	12
	0.0	33.3	50.0	16.7	---	26.1
	0.0	28.6	30.0	22.2	---	
\$30,000-34,999	0	1	0	1	00	2
	0.0	50.0	0.0	50.0	---	4.4
	0.0	7.1	0.0	11.1	---	
\$45,000-49,000	2	0	2	0	0	4
	50.0	0.0	50.0	0.0	---	8.7
	66.7	0.0	10.0	0.0	---	
\$55,000-59,999	0	1	0	0	0	1
	0.0	100.0	0.0	0.0	---	2.1
	0.0	7.1	0.0	0.0	---	
TOTAL	3	14	20	9	---	46
	6.52	30.43	43.48	19.57	---	100.0

Table 21 presents the distribution of sustainability by the total cost of the project (including the cost of the evaluation) here. There nine such cost categories, which produced a total of 38 cases for total cost versus sustainability. These range in amounts up to \$13,000. The following observations were drawn from the above tables:

^{1/} Gaps in the numerical sequence of "evaluation cost" categories reflect on absence of such cases.

DISTRIBUTION OF FINDINGS FOR THE FY82 AND FY83 META-EVALUATION PROJECTS



10

Table 21: Total Cost By Sustainability

TOTAL COST	SUSTAINABILITY (N)					TOTAL
	(ROW PCT)					
	0	1	2	3	4	
	6	5	3	1	0	15
\$1-999	40.0 66.7	33.3 55.6	20.0 27.3	6.7 12.5	0.0 0.0	39.5
\$1,000- 1,999	1 16.7 11.1	1 16.7 11.1	3 50.0 27.3	1 16.7 12.5	0 0.0 0.0	6 15.8
\$2,000- 2,999	0 0.0 0.0	0 0.0 0.0	0 0.0 0.0	4 100.0 50.0	0 0.0 0.0	4 10.5
\$3,000- 3,999	1 50.0 11.1	0 0.0 0.0	1 50.0 9.1	0 0.0 0.0	0 0.0 0.0	2 5.3
\$4,000- 4,999	0 0.0 0.0	0 0.0 0.0	1 50.0 9.1	1 50.0 12.5	0 0.0 0.0	2 5.3
\$5,000- 5,999	0 0.0 0.0	1 50.0 11.1	0 0.0 0.0	0 0.0 0.0	1 50.0 100.0	2 5.3
\$6,000- 6,999	1 20.0 11.1	1 20.0 11.1	2 40.0 18.2	1 20.0 12.5	0 0.0 0.0	5 13.2
\$7,000- 7,999	0 0.0 0.0	0 0.0 0.0	1 100.0 9.1	0 0.0 0.0	0 0.0 0.0	1 2.6
\$12,000- 12,999	0 0.0 0.0	1 100.0 11.1	0 0.0 0.0	0 0.0 0.0	0 0.0 0.0	1 2.6
TOTAL	9 23.7	9 23.7	11 29.0	8 21.1	1 2.6	38 100.0

- o Evaluations reports that consider sustainability represent an extremely wide range of funding.
- o There are no apparent relationships between the cost of the project and how sustainable the project was rated.
- o There is a slight tendency for the evaluations which were examined for total cost and which addressed sustainability to rate on the low side of sustainability.

If Table 21 is recast into low, medium and high project cost (by intervals of \$5,000), it produces the following table:

Table 22: Total Cost By Sustainability						
	SUSTAINABILITY (N) (ROW PCT) (COL PCT)					
TOTAL COST	0	1	2	3	4	5
\$1-5,999	8 27.6 88.9	6 20.7 66.7	8 27.6 72.7	7 24.1 87.5	0 00.0 00.0	29 100.0
\$5,000-9,999	1 12.5	2 25.0	3 37.5	1 12.5	1 12.5	8 100.0
\$10,000-15,000	0 00.0 00.0	1 100.0 11.1	0 00.0 00.0	0 00.0 00.0	0 00.0 00.0	1 100.0
TOTAL	9 23.7 100.0	9 23.7 100.0	11 29.0 100.0	8 21.1 100.0	1 2.6 100.0	38 100.0

This categorization reinforces the generalized skewing towards less sustainability noted from Table 21.

Time Series Analysis

The following graph shows the distribution of findings from both Metaevaluation Projects (FY82 and FY83). The graph indicates

that evaluation reports produced relatively fewer implementation findings and more Institution-Building and Impact findings during the FY83 metævaluation period. This is presented in Table 23, which is a frequency distribution of the categories for both years. Appendix 1.c shows a percent distribution of findings types for each bureau for the FY 82 and FY 83 evaluations. The findings types within each bureau show little significant change from one year to the next.

Table 23: Frequency Distribution of Categories FY82-FY83

		FRE-		FY82	FRE-		FY83
		QUENCY	PERCENT	CAT	QUENCY	PERCENT	CAT
DESIGN	Overly ambitious objectives	17	1.5	11	40	3.3	
	Conflicting objectives	11	1.0	12	20	1.6	
	Failed assumptions	100	8.8	13	68	5.6	
	Missing inputs and outputs	48	4.2	14	28	2.3	
	Scheduling & budget	50	4.4	15	65	5.3	
	Recommendations & planned changes	113	9.9	16	165	13.6	
IMPLEMENTATION	Problems finding US contractors & personnel	7	0.6	21	14	1.2	
	Problems finding host country contractors & personnel	16	1.4	22	14	1.2	
	Commitment & performance of US contractors & personnel	113	9.9	23	66	5.4	
	Commitment & performance of host country gov't, contractors & personnel	108	9.5	24	87	7.2	
	Commitment & performance of both US and host country contractors & personnel	24	2.1	25	56	4.6	
	AID reporting requirements	14	1.2	31	5	0.4	
	Contracting & funding procedures	43	3.8	32	16	1.3	
	Coordination between AID & contractor	36	3.2	33	13	1.1	
	Procurement of commodities	31	2.7	34	21	1.7	
	Delay litanies	44	3.9	35	34	2.8	
	Coordination between AID & contractor	7	0.6	36	23	1.9	
INSTITUTION BUILDING	Progress at the central level	48	4.2	41	23	1.9	
	Progress with decentralization	11	1.0	42	12	1.0	
	Progress at the community level	53	4.7	43	71	5.8	
	Progress with training	33	2.9	44	53	4.4	
	Self-sufficiency & recurring costs	26	2.3	51	26	2.1	
	Strategies & structures	39	3.4	52	80	6.6	
DATA	Training problems	20	1.8	53	25	2.1	
	Collection & analysis of data	54	4.8	61	67	5.5	
	Plans developed via that analysis	5	0.4	62	15	1.2	
IMPACT	Disseminating information	17	1.5	63	6	0.5	
	Production impact	12	1.1	71	29	2.4	
	Economic impact	21	1.9	72	31	2.6	
	Social impact	11	1.0	73	37	3.0	
TOTAL	Spread/imitation effects	6	0.5	74	7	0.6	
		1138	100.0		1217	100.0	

SUMMARY OF FINDINGS

The primary findings presented in detail previously in this chapter can be summarized as follows:

Major Headings

- o "Design" relatively accounts for the most
 - problems with design
 - responsiveness to revision

Bureau

- o Little differentiation in findings patterns were described by regional bureau distributions
- o A strong correlation of central bureau (Impact and S&T) with "major groupings" of findings due to those bureaus clear focus/mission as regards to their evaluations. FVA resembles a regional bureau in its overall distribution, but has proportionately fewer implementation findings.

Technical Focus

- o "Major headings" of findings vary more by this factor than by bureau: different types of projects produce different "profiles" (proportions) of findings throughout the project cycle, demonstrating "strengths" and "weaknesses" in the projects the reports are evaluating.
- o Overall, there is more attention to data and impact.

Evaltime

- o Interim and Final evaluations both look at Institution-Building in fairly equal proportions.

(Contrary to the hypothesis that this should be a relatively greater focus of Final evaluations).

Sustainability

- o Regional Bureaus score relatively lower
- o Final evaluations score relatively higher
- o Technical activities tends to score somewhat towards less sustainability.

III. RECOMMENDATIONS

RECOMMENDATIONS FOR FUTURE EVALUATION REPORTS

There are three major recommendations drawn from the analysis of findings presented in Chapter II regarding USAID evaluation reports themselves.

The first of these pertains to "Data." This topic was consistently the weakest in the evaluation reports and is one which can be addressed through a variety of techniques.

AID has developed a number of expeditious evaluation techniques, of which perhaps the most effective is "rapid rural appraisal." AID evaluation scopes-of-work should include suggestions for data-collecting methodologies drawn from these techniques. Contractors of all types, having to respond to such techniques as a stated part of the evaluation's scope of work, would thus be required to include some aspect of data collection in their evaluation report.

Since the methodologies developed for data correction would have been drawn from AID's repertoire, and given the proviso that alterations to these methodologies would have to be justified by more than "environmental" factors (e.g., the rainy season)--factors that should have initially been taken into account--the evaluation reports produced with this requirement would begin to provide AID with standardized, community-level information relevant to socio-cultural and economic conditions. This type of generalized data correction would not only facilitate the future design of projects, since particular areas of need would have been identified, but would also provide specific indicators against

which future evaluations of the projects could meaningfully measure impact and change.

To summarize, AID needs to standardize acceptable data collection methodologies and to compel, as much as possible, evaluators to follow methodologies from that universe.

The second of these recommendations pertains to "Design" and "Implementation." These are the two largest major categories of findings headings found in the evaluation reports, but are also the most negative ones in their comments. Most of the categories pertain to specific problems in either "Design" or "Implementation," and most fault AID, the contractors, the host country government, or some combination of the three. It seems clear that AID still has major difficulties designing, implementing and managing development projects based on the consensus of these findings.

Recommendations to improve these categories are divided by "design" versus "implementation and management."

Design

- o AID should modify the logical framework to include a more comprehensive cultural review and to incorporate advice from the host country itself and experts in those technical fields. What this requires is not so much a longer logframe format as one which is holistic in its orientation as opposed to causal and linear in its depiction of the project.
- o AID should prepare RFPs which are very clear as to the tasks required and the types of information necessary for the successful review of submissions: tasks and information that emphasize the necessity to incorporate culture-specific values rather than economic or technical ones.

- o AID should require a more stringent review process of the designs submitted to reduce the need to rework a project because of overly ambitious or conflicting objectives and, in particular, failed assumptions. This review process should be performed by the mission personnel, ideally those staff members who will be responsible for the project's oversight.

Implementation & Management

- o As part of the RFP submission, or in the initial stages of project implementation, a task flow chart should be drawn up by AID and the contractor, with specific responsibilities and deadlines established. Liaisons should be established at that time to facilitate communication. All parties involved would then have a clear idea of their respective responsibilities.
- o Project monitoring should be done by individuals in the mission with expertise in the particular technical activity at issue. In order to facilitate project monitoring, required reporting procedures should not only have the parameters and deadlines recommended above, but should be sufficiently streamlined so as to reduce the paperwork burden on the officer-in-charge.
- o Site visits by the project officer should be both scheduled and unscheduled, and of sufficient length and variety to provide a comprehensive overview of the actual (as opposed to reported) management and status of the project.
- o Projects done previously by the mission regarding the same technical activity should be used by both contractor and project officers to take advantage of lessons already learned. Other non-written resources, such as AID or host country personnel, should also be contacted to incorporate their experiences and ideas into the ongoing project.
- o Communication at every level of project implementation cannot be emphasized too strongly and must be a vital part of any project/evaluation management structure.

Specific issues to be addressed on a bureau- or technical activity-wide basis should be discussed with the evaluation staff

by the project officer at the beginning of the project. Sustainability, for example, is a variable which could have been addressed more widely by the evaluation reports reviewed for this study if that particular issue had been incorporated into the evaluation process, rather than imposed once the evaluation was over.

If specific issues are to be addressed, they must be well-defined, as there are a variety of meanings attached to most development terminology, and the use varies not only with context, but with author. The definition must be extremely clear and the analysis must be equally directed.

Once a specific issue has been defined, its use by decision-makers must also be specific, as this will, in turn, determine the types of analyses which can be performed to address that issue in a given evaluation.

RECOMMENDATIONS FOR FUTURE FINDINGS STUDIES

The metaevaluation project's findings' report began initially in FY 82 as a by-product of the scoring report and, thus, originally lacked both the extensive development of a conceptual instrument, and a specific focus/purpose.

Future findings reports similar to this one should be as specific as possible. They need to be performed to address a particular issue or question, and the evaluations used to support that analysis should be a sample of all evaluations performed, not the entire population of evaluation reports for a specific year. Sampling is more efficient, while defining the topics/issues to be studied will serve to narrow the population from which the samples should be drawn.

If findings are to be studied as they have been drawn from a metaevaluation project such as this one, then their categorization scheme should not necessarily reflect the entire project cycle, but perhaps only the specific issue(s) to be addressed. This would serve to focus the information collection and analysis, and to reduce the complexity of the categorization scheme.

Additional areas which should be examined in future findings reports are the incorporation of mission comments into the review of evaluation report findings, follow-up comments based on contractor/mission reactions to the evaluation report, and the impact of evaluation report recommendations and utilization of recommendations.

AID has already produced other findings reports focused on particular issues or technical activities. Of particular note are the Program Evaluation Discussion Papers Nos. 11 and 13. No. 11 is a review of Effective Institution-Building and No. 13 is a review of Agricultural Research. Because of the specificity of the topics, they can be studied with greater depth and more relevant, specific recommendations can be made with greater utilization potential.

IV. METHODOLOGY

OVERVIEW OF PROJECT

The "metaevaluation project" began in 1982 and has comprised three separate contracts. The first contract entailed the development of the scoring instrument^{1/}, the second contract applied the instrument to FY82 evaluation reports^{2/} and, finally, this contract applied a revised version of the instrument to the FY83 evaluation reports^{3/}.

The scoring instrument was developed as a means to assess internal characteristics which form a "good" evaluation. These characteristics were derived through interviews and questionnaires and were defined as factors over which the evaluation team had control. That is, the characteristics were internal to the evaluation itself. These included criteria for data collection, data analysis and overall presentation and verification of information.

Also included as an integral part of the scoring instrument was an application of the logical framework, which is a management tool used in the design of projects. The "logframe" concept attempts to establish a casual relationship between the various components of project design and also serves as a reference to assess the evaluation's analysis of each of those components.

-
- 1/ Final Report: Development of a Quality/Completeness Scoring Instrument for USAID Evaluation Reports. Contract No. AID/SOD/PDC-0391, Work Order No. 1.
 - 2/ Final Report: Analysis of the Quality of FY80-82 AID Evaluation Reports. Contract No. AID/SOD/PDC-0391, Work Order No. 2.
 - 3/ Final Report: Analysis of FY83 AID Evaluation Reports. Contract No. OTR-0000-C-00-3482-00.

The first metaevaluation effort used this scoring instrument developed by TRITON to assess the quality and completeness of all FY82 evaluation reports, as well as a sample of FY80 and FY81 evaluation reports for a time-series analysis. The internal characteristics used to analyze quality and completeness were given weighted values according to their relative importance, applied to the reports to give characteristic scores and then summed to produce a total score.

External characteristics were also analyzed against these scores to determine any statistical relationships. The principal external characteristics examined were: bureau, technical focus, time of the evaluation, unit originating the evaluation, cost of the evaluation, host country participation, contractors, and number of individuals/firms participating in the evaluation effort.

Two additional components were performed with this (FY82) metaevaluation task. The first consisted of a compilation^{1/} of project design abstracts (provided by AID's computerized data base) and the findings generated by the coders during the metaevaluation process. "Findings" had been defined as short, concise statements referring to major conclusions or recommendations made in the evaluation report. "Coders" refers to TRITON's staff responsible for the actual application of the scoring instrument during the metaevaluation project.

The second additional component was an analysis of the distribution of findings by external characteristics; in particular, bureau, time of evaluation, technical focus and coder.^{2/}

^{1/} Final Report: Findings Compendium of the FY82 Metaevaluation. Volumes I and II, Contract No. AID/SOD/PDC-0391, Work Order No. 2.

^{2/} Final Report: Findings Compendium and Analysis of FY82 AID Evaluation Reports. Contract No. AID/SOD/PDC-0391, Work Order No. 2.

The subsequent metaevaluation project of FY83 evaluation reports utilized modified scoring instrument. Major modifications were based on a shift from quantitative to qualitative variables: the experimental hypotheses connected with the logical framework analysis were replaced by variables assessing the impact and extent of external factors, as well as a particular focus on the issue of sustainability. Additional external variables were added to include the presence/absence of a scope of work, reallocation of inputs and outputs, as well as the redefinition of "host country participation" to mean host country participation on the evaluation team. As part of the shift from quantitative to qualitative analyses, the analysis of the distribution of findings was given precedence over the distribution of scores, and two smaller adjunct studies were performed, one to focus on contractor perceptions of evaluation reports and the second to determine the types and distribution of innovative techniques used in the FY83 evaluation reports.

FINDINGS' CATEGORIZATION SCHEME

The findings from the FY83 metaevaluation project were categorized using a coding scheme developed for the FY82 metaevaluation project. As further information can be found in Final Report: Findings Compendium and Analysis of FY82 AID Evaluation Reports (Contract No. AID/SOD/PDC-0391, Work Order No. 2), that coding scheme will be only briefly summarized here.

The categorization process was performed inductively. Findings were read and then grouped together by topics. From those groupings, general category names were derived for a total of thirty-one (31) categories. These categories encompassed the entire project cycle, from design through impact.

This categorization scheme was applied to the findings of the FY83 metaevaluation project. Definitions and examples of each

category are presented for each of the thirty-one categories (see Figure IV-1).

These categories are grouped into five major headings. The headings are: Design, Implementation, Institution-Building, Data and Impact. These headings refer closely to the various stages of the project cycle and appear throughout the analysis to provide a more generic description of the findings' content. The major headings and their respective categories appear in Figure IV-1.

Since these categories were inductively derived, they represent the distributions already present in the data (the FY82 USAID evaluation reports), rather than responding to a pre-formulated categorization scheme. This results in a categorization scheme which is, in essence, an artifact of the data. In a recent comparison with other findings-type reports^{1/}, sufficient similarities were ascertained to determine that while the categories are artifacts of the data, they are sufficiently representative of the overall project cycle to be treated as valid indicators of the presence/absence of different aspects of that cycle.

DEFINITIONS AND EXAMPLES OF FINDINGS CATEGORIES

DESIGN

Overly ambitious objectives. This category occurs in the design phase of the project, although it naturally has an impact on project implementation if not remedied while still in that stage. The category can be described as an indication of attributing an unrealistic amount of importance to one particular

^{1/} Comparison of TRITON, DIU and Asia Bureau Findings. Contract No. OTR-0000-C-00-3482-00.

facet of the project, usually overestimating management's capability of dealing with that aspect.

- o Project design was overly ambitious for the time available and decisions will have to be made as to which aspect of the project is given priority. (Project No. 6080166)
- o Project design was too ambitious at goal/level purpose -- AID can't change the complex education system of Nepal simply with this project. (Project No. 3670123)
- o Project design was overly optimistic in terms of time it takes to build an institution and the level of cooperation between donors necessary. (Project No. 6490112)

Conflicting Objectives. This occurs when the project does not clearly define exactly what it wants to accomplish or, for example, when its objectives are so broad that there is no way to manage one aspect of the project without negatively impacting on another, usually with regards to areas of responsibility or authority and local systems.

- o The conflicting objectives in the project design led to compromises and poor scheduling of events. (Project No. 2630065)
- o It is still unclear yet as to whether the scope of project activities are proportionate to the scale of environmental problems. (Project No. 3670132)
- o During project selection at field level, considerable influence may be exerted by special interest groups or individuals. (Project No. 4930315)

Failed Assumptions. This category refers to something taken for granted not materializing, usually with negative impacts on project implementation. These assumptions may be either internal and refer to the project design, or external and refer to such factors as climate or the world economy.

- o Many of the initial loans became insufficient to cover the cost of the housing construction due to the February 1982 devaluation of the peso and the establishment of a parallel dollar market, resulting in unfinished additions. (Project No. 5110510)
- o Project did not take cultural context and language requirements of the beneficiaries into account in design and implementation: this greatly hampered chances for project success. (Project No. 6150157)
- o Participants were selected because of availability, not English language ability, with the result that a great deal of time was spent teaching English before content training could get underway. (Project No. 6600052)

Missing Inputs and Outputs. This occurred when the necessary logical steps in the daily management of the project had not been developed and, thus, critical things were omitted. It generally resulted in delaying the project while a makeshift replacement could be developed or, in some cases, in seriously affecting the motivation of the implementing agency and the beneficiaries, as well as the situation which had prompted that project.

- o There is a question as to whether there will actually be job openings for participants: more emphasis on placement is crucial. (Project No. 6080157)
- o Training in management and leadership is necessary to build any institution. (Project No. 3670102)
- o Although most of the data needs for health planning have been identified, these have never been assembled into a cohesive, clearly stated health information base for planning at the district, provincial, and other levels of the MOH. (Project No. 6150187)

Scheduling and Budget. This is the most formalistic of the design categories, since it pertains directly to a responsibility of the managing unit that does not rely on host country variables: If the contractor was doing something "stateside," it would still have to provide this type of data. However, such data is missing

or faulty because of a variety of factors, ranging from the funding pipeline to unrealistic expectations concerning subproject implementation.

- o Recommendations for revised systems of land registration, based on project tests and evaluations, will be difficult to implement during remaining duration of project. (Project No. 4970312)
- o The interval between the signing of the USAID/NCIH agreement, and the signing of the NCIH/CARICOM subagreement has created a dilemma for those three, as the two agreements, which are interdependent, are scheduled to terminate at different times, some eight months apart. (Project No. 5380054)
- o Due to inflation and rising construction costs, target number of warehouses was delayed in construction while additional funds were negotiated. (Project No. 6320210)

Recommendations and Planned Changes. This category differs from the others in that it is not simply descriptive, but rather prescriptive: That is, it offers suggestions on remedying the problems or else relates what has been decided to correct the shortcomings in the project design. Despite project difficulties, once the managers are aware of problems, steps are taken to rectify them or when it doesn't appear as though any reasonable strategy will help, recommendations are made that the project be terminated.

- o Outreach messages involving local farmers and market women should be initiated. (Project No. 3670118)
- o The Solomon Islands projects need to involve the men, to make the workshops more practical and to leave something behind on the ground. (Project No. 8790251)
- o More emphasis on stimulating private sector involvement in forestry is necessary. (Project No. 9365519)

IMPLEMENTATION

Contractors and Personnel:

Problems Finding U.S. Contractors and Personnel. Before a project can be implemented, the requisite staff for the tasks involved have to be found. It is often a long process, due to lack of language skills, or the level and type of skills needed. This was true in some of the projects for which evaluations were reviewed regarding U.S. or third country contractors, while other problems presented themselves to host country contractors.

- o Severe delays in technical assistance provision were partially responsible for necessary reorganization of nature and type of technical assistance. (Project No. 6490101)
- o Rangeland research is limited due to insufficient staffing. (Project No. 6490108)
- o Project was delayed for eighteen months due to difficulties in hiring long-term technical assistance. (Project No. 6570006)

Problems Finding Host Country Contractors and Personnel.

These types of problems range from initially finding qualified personnel since the pool of skilled labor is usually very limited to the personnel who are qualified not being released (or only reluctantly) from their government positions. While a language problem is not specifically mentioned, in many of the larger countries a government official would probably not know the local languages of many of the beneficiaries.

- o Considerable delay in implementation was caused by inability to recruit competent host country staff. Educational levels had to be reduced and will require additional technical assistance towards staff development. (Project No. 2780245)

- o The recruitment of personnel is difficult in the Project Area. A combination of factors, the isolation of the dispensaries -- making it difficult to attract staff who with that level of training prefer to live in larger towns or join the army or police force, the variety of languages making it highly desirable to recruit from within the local village community, etc., prevented dispensaries from meeting target numbers. (Project No. 6310001)
- o Counterparts were not provided in a timely manner due to the chronic over-allocation of personnel. (Project No. 6570006)

Commitment and Performance of U.S. Contractors and Personnel. This category refers to primary responsibility for the implementation and was used if fault or praise could be leveled at one contractor in particular with regards to project design or implementation. This category does not refer to the project evaluation contractor.

- o The two studies generated are of dubious utility, took an excessive amount of time to complete, used an untested survey instrument and would have been more useful if the authors had worked cooperatively in the research activities. (Project No. 6080147)
- o PMO management has been successful in promoting good working relationships between its staff, consultants and contractors and, through closer field supervision, has coordinated implementation well. (Project No. 4920310)
- o Problems have arisen with graduate students as "researchers"; they are preoccupied with their own research and tend not to gain the cooperation of their counterparts. (Project No. 6310007)

Commitment and Performance of Host Country Contractors, Government and Personnel. This category focuses on the willingness or national capacities of the host country governments to expend their own manpower: while many of the findings reflect negative impacts -- lack of coordination or personnel, or near

total disinterest, there is a growing realization of cooperation as the most effective means to achieve objectives.

- o An apparent lack of commitment, both in timeliness and decision-making ability of the Directorate of Land Registration to meet project purpose has slowed project implementation. (Project No. 4970312)
- o Staffing still poses a problem due to a combination of pay and organizational problems: over 83% of the professional staff have announced their intention to resign if this is not corrected. (Project No. 5320065)
- o High quality and dedication of counterpart staff/trainers contributed to output achievement. (Project No. 6310044)

Commitment and Performance/Coordination of Both Host Country and U.S. Contractor, Personnel and Government. This category was derived after the other single-fault categories were deemed too narrow to focus on the problems that beset any project in which more than one contractor is involved and, thus, there is no single entity to commend or criticize.

- o Dissatisfaction has been expressed by every party involved in the implementation of this project. (Project No. 2790045)
- o There has been a lack of coordination between the participating sectors. (Project No. 5200272)
- o CEPAS' progress has been notable among IRT projects. The reason for its success in comparison with others appears two-fold. First, the recipient organization committed itself to the activity as a continuing program rather than as an exceptional activity. Second, it has the organizational capacity to carry out its stated objectives. (Project No. 698040715)

Procedures and Bureaucracy:

AID Reporting Requirements. The majority of findings in this

category stressed the need to simplify the reporting requirements but, at the same time, to clarify them so that it would be clear to the contractors what their responsibilities were and when those documents were to be delivered. This would result in an improvement in AID's ability to monitor the project in terms of sheer paperwork requirements, and the reduction of those at the same time might result in more on-site monitoring to prevent abuses of the projects.

- o Some inattention by AID to its responsibilities for monitoring and for providing answers to requests for interpretation of grant terminology and AID policy regarding procurement and contracting requirements delayed project implementation. (Project No. 6980412)
- o Implementors are nearly always over-optimistic about the good they do, while outside evaluators working under severe time and data limitations usually omit certain class benefits that are indirect and hard to document. (Project No. 9380250)

Contracting and Funding Procedures. This category again reflects the paperwork burden imposed by -- and on -- AID. There are some innovations in the types of funding, not all successful, and some projects apparently devise their own procedures for contracting, also with varying degrees of success. Many of the less successful ones delayed or hindered the overall implementation of the projects, but they are nonetheless interesting for signaling a willingness to innovate in solving a problem for the benefit of both AID and the contractors.

- o Outputs are behind schedule, due in large measure to FUNDACED's inexperience with management and with AID's regulations and requirements. (Project No. 5200269)
- o Lack of pre-proposal orientation on USAID regulations, procurement waivers, contract modifications and the significance of PACD caused implementation problems for

several of the IRT proposers. (Project No. 698040705 et al.)

Coordination Between AID and Host Country. This category refers both to the reporting procedures and high staff turnover, which makes coordination -- and certainly continuity -- very difficult, and the lack of language ability among that staff. Many of the problems derive from a lack of communication between AID staff and the host country nationals.

- o USAID was praised by the Department of Agricultural Extension staff for the advice, support and encouragement provided to the training program. (Project No. 4930280)
- o The project was not well-guided by either the Indonesian agency responsible nor by USAID during its inaugural period. (Project No. 4970308)
- o Much of the goodwill generated by the provision of the ferry has dissipated due to the lack of understanding between AID and GROD as to whether it was a temporary or permanent measure. (Project No. 698041024)

Procurement of Commodities. A project cannot start without the basic inputs: This category is basically a recitation of complaints, whether about AID's policies, or problems with host country procedures, or the actual mechanics of getting the commodity from the warehouses to the targeted sites.

- o Seed production is the cornerstone of this project: if the government cannot furnish a sufficient quantity, then it should contract out. (Project No. 3670118)
- o Procurement delivery delays have plagued the project from its inception. (Project No. 6850247)
- o A major problem identified by the evaluation team is that even though the ministries and other government agencies have the capable manpower to operate microcomputers, they do not have the funds for purchasing computer hardware and

software. (Or, more precisely, they have difficulty buying imported computer -- hardware and software -- technology.) (Project No. 9365728)

Delay Litanies. This category was derived when it was apparent that some findings -- unless they were individually dissected clause by clause -- were essentially a long array of all the things that could "go wrong" with a project. These ranged from commodity problems through host government constraints, to financial problems within the mission, or to sheer logistical constraints in transporting the materials to sites.

- o Project was delayed for one year because of contract problems; even though it was extended, it has still run into more problems with consultants, personnel, construction, seed production, government cooperation and coordination between implementing organizations. (Project No. 3670118)
- o Causes of delays include:
 - Insufficiently specific guidance and criteria for approval of sub-projects
 - Inadequate staff attention both by the Government of India and USAID
 - Delays in the Government of India approval process. (Project No. 3860465)
- o Government agency's bureaucratic slowness in combination with petrol and transportation shortages diminished quantities/quality of outputs. (Project No. 6410064)

Coordination Between AID and Contractor. This category was added when the gap between the contractor problems and some of the areas of coordination was recognized. This category was designed to focus on the need for more and better coordination between AID and the implementing organizations, which often seemed to be functioning at crossed-purposes.

- o Failure to assist in project logistics by USAID/Sanaa caused unnecessary delay in technical completion of the soil survey program. (Project No. 2790042)
- o AID responded in a timely manner to all of the changes required by the redesign. (Project No. 5150158)
- o Although relationships between the voluntary agencies and USAID missions are almost universally cordial and mutually respectful, but, except in which a Food for Peace or Voluntary Agency Officer is adequately informed and has the time, there is a minimum of communication on substantive issues. Relationships between AID/W and the voluntary agencies are smooth but there are unavoidable delays in expediting project approvals and amendments. (Project No. 9040006)

INSTITUTION BUILDING

Progress:

Progress at the Central Level. This category assesses the changes made -- usually attributed to the project -- in the host government's activities/capabilities at the national level. All of the findings within this topic and through the next three reflect positive improvements made at some level of the project.

- o Institute is a strong and self-sustaining institution with excellence in facilities, faculty, curricula and the students it trains. It will continue long beyond life of project. (Project No. 2760019)
- o Research council is an action agency that stimulates and encourages research, creating an enthusiasm among researchers to work toward national goals. (Project No. 3880051)
- o Project achieved most of its outputs and accomplished its stated purpose of strengthening government's capacity to plan and delivery health services. (Project No. 6600057)

Progress with Decentralization. This category was interpreted as incorporating some of the extension elements into a sub-national level, although not yet at the community level. It stresses the outreach aspects of programs.

- o Decentralization of responsibility to a regional level also facilitated the disaster recovery. (Project No. 1500025)
- o New marketing system for fertilizer distribution has been introduced. Switch from public to private distribution system is occurring and price de-control being introduced. Private distribution system seems to be working but will need a long time to become established -- institutional transfer can't be done overnight. (Project No. 3880024)
- o This AID-financed portion of the project has stimulated enthusiasm and expectation in a credit union movement that was largely dormant. (Project No. 6930220)

Progress at the Community Level. This category continues the description of progress to smaller units, this time focusing on efforts at the community level.

- o The women are very eager to continue these money-making endeavors, and to formalize their working together with a cooperative. (Project No. 6080166)
- o The number of schools implementing sub-projects was lower than expected, but the establishment of school projects' target was still exceeded. (Project No. 5220170)
- o Project has developed, and is developing a community development infrastructure in a region where no such enterprise had previously been tried. (Project No. 6310010)

Progress with Training. One of the vital components of institution-building is training host country persons.

- o The 85 lecture notes (notat) in Arabic may be the most comprehensive and practical collection of hands-on written

skills development texts for these technical areas in any language. (Project No. 2760079)

- o Project has been successful in meeting its quantitative output goals regarding teacher enrollment, radio/educational material disbursement and curriculum development. (Project No. 3670123)

Problems

This section reflects all the things that can "go wrong" with institution-building, from self-sufficiency through training. This category contrasts with the successes described on the last few categories discussed. It initially was set up in this dichotomy to make it easier to visualize the broad categories of "progress" versus "problems," with the basic types of each derived from examining the overall findings as to institution-building.

Problems with Self-Sufficiency and Recurrent Costs. This category marks one of the major problems with any project: How to provide the project with sufficient funds in the beginning and then channel the funds so as to generate enough revenue to sustain the project without renewals.

- o High loan delinquency rates lessen amount available for re-lending, hurt chances of getting additional funding, and generally harm project's potential for sustainability. Re-financing of delinquent loans is a big problem. (Project No. 5220177)
- o The project has a "give-away" nature which may have caused, according to the evaluator, more individualistic farming, false expectations for future programs, altered perceptions of credit/acquisition, and a burden on the Government of Gambia and AID to keep the program going. (Project No. 6350215)
- o Maintenance was only sporadically accomplished on the vessel due to 1) the French technician devoting little time to it and 2) the Djiboutian technicians being dis-

couraged from any initiatives as the keys to the engine room and cabin were kept locked by the French. (Project No. 698041024)

Problems with Strategies and Structures. This category deals with the difficulties encountered from using a particular technique that proved ineffective in the cultural setting under review, or due to diversion of trained personnel from their proposed job to another area of the country's needs.

- o The Solomon fisheries project needs work: turtle sanctuary component was destroyed by raiding war party and fisheries need much better to market service, as well as ice makers. (Project No. 8790251)
- o While the training is, generally, relevant and considered of value to the Government of Zaire, it is hard for the Service to keep its trained personnel in light of higher salaries elsewhere. While a sense of moral commitment may delay departure, it cannot prevent it. (Project No. 6600052)
- o Anti-development effects of food programs include dependency, passivity, a "give-me" attitude and, in sierra communities, the undermining of community work traditions. (Project No. 9380801)

Training Problems. This category is diametrically opposed to the one mentioned under "progress," and reflects all of the things that can "go wrong" with any training program, whether it involves simply not achieving a targeted output, or using an inappropriate curriculum.

- o Educational component is sound conceptually, but does not adequately address the problem of differential student preparation, background and interest and has not been institutionalized into the overall program. (Project No. 4930314)
- o Area farmers felt that the project was a wasted effort, as they were not taught techniques which would be useful to

them, and spent time performing the measurements for someone else's research. (Project No. 6310007)

- o The program might have taken into consideration more of the economic and social concerns of the communities and of Lesotho in designing and implementing the training program. (Project No. 9380184)

DATA COLLECTION AND ANALYSIS

It is interesting to note that while this subject comprised several of the internal variables for the metaevaluation's quality/completeness assessment instrument as to criteria for a "good" evaluation, proportionately there were very few findings under these categories, no matter how well or poorly the evaluation scored. The Office of Science and Technology is the only exception to this generalization: A concentration of findings under the category of data collection and analysis found for this bureau proved not to be an indicator of a "better" or "worse" quality evaluation report, but merely served as an indicator of the general finding subjects found within that bureau's reports.

Data collection. This category refers to the actual dynamics of finding and assembling a body of knowledge, and then conducting an analysis of it. The category also divides fairly neatly into things that went wrong and those that did not.

- o The Rural Dynamics Survey has provided an important and unique source of information about the process of economic, social, and institutional change in rural Indonesia, and there is a continuing need for further work of this nature. (Project No. 4970225)
- o There is a lack of survey data to measure progress -- in part because of transportation difficulties and the University of Ghana computer breakdown. (Project No. 6410064)

- o Hypothesis that agrotechnology transfer may be made by use of soil taxonomy was sufficiently tested by rigorous scientific and statistical methods for AID and others to use in technology transfer. (Project No. 9310582)

Plans. This category refers to the types of output the information has/has not been generating.

- o A reasonable and workable system for monitoring has been developed. (Project No. 3880051)
- o The Agricultural Sector Planning Unit, although no longer threatened by the Integrated Area Development Studies (IADS) as a rival, appears disinterested in using the IADS planning methodology in setting agricultural sector investment priorities. (Project No. 5200249)
- o CATIE has successfully developed a farm research methodology which is being used nationally. (Project No. 5960083)

Disseminating Information. This is another category which reveals what is done/not done with the information collected. By-and-large, the most prevalent problem, however, is getting the relevant data out of the "laboratories" and into the fields where it could be of most use.

- o Problem exists in disseminating results to other processing plants. (Project No. 49802501)
- o Key implementation/operations weakness is the lack of an adequate management/dissemination system. (Project No. 5960089)
- o Reports and efforts to disseminate data are taking place. (Project No. 6980420)

IMPACT

Production Impact. This category delineates any increases in the productive capacity of the target group/crop/project that are attributable to the mechanisms in the project.

- o Evaluation team feels that the project will have a substantial positive impact on the country's agricultural mechanization as a whole. (Project No. 2630031)
- o Both total production and marketing have risen in the project area as a direct consequence of the project. (Project No. 6790001)
- o Improved Rural Technology-type activities, while small, do address a development need in countries like Tanzania. Impact in relation to AID funding inputs is significant. (Project No. 698040705)

Economic Impact. This refers to any observable changes in the income/purchasing power of the targeted group, and is usually a result of some increase in production -- using the methods set forth in the project or related to expressed needs by the target population -- in the general area.

- o Project shows that credit on soft terms is feasible: demonstrated by high repayment rates and money put to good use in purchasing agricultural inputs. (Project No. 4920339)
- o Rural savings increased by 223%, more than double the targetted goal. This is one of the major strengths of La Merced, i.e., farmers' trust in its operations despite national economic and political instability. (Project No. 5110533)
- o Transition from a type of State Farm to private enterprise with tenant farmers and free market has resulted in increased production, sales, private shops and industry. (Project No. 6490103)

Social Impact. This category groups all of the "non-tangible" changes, such as health or increased goal achievement, together. These changes are generally reported as positive in the evaluation reports' findings.

- o Project has contributed substantially to the expansion of the basic health service network in Central Tunisia:
 - Health centers constructed/renovated, staffed and are providing in/out-patient care
 - Physicians trained in management
 - Paramedical workers trained in basic health care
 - Improvement of management/administrative systems. (Project No. 6640296)
- o Decreasing malarial incidence is due to early detection, prompt treatment by clinic personnel and new regimes of anti-malarial drugs developed and used throughout the country. (Project No. 4930305)
- o The project has been effective in reaching the target group: low-income families in the rural areas of Bolivia. (Project No. 5110510)

Spread/Imitation Effects. This category contains an excellent measure for determining project success: a finding that a design is replicated, formally or informally. The informal mechanisms usually involve the spread of a type of organization/innovation from one geographic area to another, while a more formal dynamic would be an invitation from a third country to initiate a similar project there.

- o Positive unplanned effects:
 - More than called for improvements by government in institute's physical plant
 - Curricula/material extended to other health institutions
 - Qualified faculty attached to institute from all over the region -- more than expected. (Project No. 2760019)

- o This community development model has proven replicable: similar projects have been started elsewhere based on lessons learned here. (Project No. 6310010)
- o A nascent shipbuilding industry has begun to meet the demand for boats which can use the new outboard motors and heavier nets. (Project No. 6570006)

DATA PROCEDURES

There are two primary statistical techniques used in the analysis of the findings' distributions. These are: 1) frequency distributions; and 2) "two-by-two" tables. These statistics describe the way the findings are grouped by the external variables of bureau, technical activity, evaluation time, and contractors. A frequency distribution can be defined as a table representing the overall numbers (frequencies) of types of items, grouped by those types of items. In the above analyses, these groupings are by bureaus, technical activities, evaluation time, or contractors. A frequency distribution is used to ascertain: 1) how many of each type analyzed are present under a specific condition; and 2) whether or not that number (frequency) is disproportionate with respect to all types. An example is the percentage of findings found by type of finding for each of the technical activities (see Table 6), and the subsequent analysis presented to identify disproportionate values.

The analysis procedures used to interpret these statistics are threefold: 1) description of data differences; 2) analysis of the statistical significance of those differences; and 3) hypotheses for the distributions and relative differences.

The operative questions addressed are:

- o How are the findings distributed?

- o What are the major groups within that distribution?
- o What can one infer from those groups?
- o How are the findings distributed with respect to three external variables: bureau, technical activity, and time of evaluation?
- o What are the major groups within those distributions?
- o What can one infer from those groups?
- o How do the distribution of findings and that of the issue of "sustainability" correlate?
- o Is the relationship between the distribution of findings and the issue of "sustainability" (if this issue is noted in a given evaluation report's findings) correlated with any of the three external factors (bureau, technical activity and time of evaluation) examined above?
- o How does the distribution of FY83 AID evaluation report findings differ from that of the FY82 findings?

These questions were addressed in Chapter I of this report. The basis for the answers were generated largely through the use of the Statistical Analysis System (SAS) software package run on AID's own mainframe (IBM 370) computer. Coded responses for the data were inputted and then analyses of frequency and variance were performed by TRITON.

This report is presented as a more qualitatively oriented analysis of the FY83 evaluation reports and their findings, while the Scoring Report (forthcoming) presents similar analyses for the distribution of the quality and completeness scores of the evaluation reports studied. The scores of the reports and a partial analysis of these scores are, however, presented in the appendices of this report as a reference tool. Additional references for

individual projects may be found in Final Report: Findings
Compendium of the FY83 Metaevaluation (TRITON: 1984).

APPENDIX 1.a: Techcode Distribution

TECHCODE	CODER				(N) (ROW PCT) (COL PCT)
	1	2	3	4	TOTAL
AGRICULTURE	30 40.0 33.0	23 30.7 33.3	8 10.7 23.5	14 18.7 27.5	75 100.00 30.6
RURAL NON AG	4 30.8 4.4	8 61.5 11.6	1 7.7 2.9	0 - -	13 100.0 5.3
RURAL MULTI- FUNCTION	5 29.4 5.5	5 29.4 7.2	4 23.5 11.8	3 17.6 5.9	17 100.00 6.9
NUTRITION	1 20.0 1.1	3 60.0 4.3	0 - -	1 20.0 2.0	5 100.0 2.0
POPULATION	2 25.0 2.2	3 37.5 4.3	1 12.5 2.9	2 25.0 3.9	8 100.0 3.3
HEALTH	9 34.6 9.9	9 34.6 13.0	6 23.1 17.6	2 7.7 3.9	26 100.0 10.6
EDUCATION	12 60.0 13.2	1 5.0 1.4	2 10.0 5.9	5 25.0 9.8	20 100.0 8.2
HRD	7 19.4 7.7	2 5.6 2.9	2 5.6 5.9	5 13.9 9.8	16 100.0 6.5
INFRA- STRUCTURE	3 17.6 3.3	6 35.3 8.7	4 23.5 11.8	4 23.5 7.8	17 100.0 6.9
OTHER	18 37.5 19.8	9 18.8 13.0	6 12.5 17.6	15 31.3 29.4	48 100.0 19.6
TOTAL	91 37.1 100.0	69 28.2 100.0	34 13.9 100.0	51 20.8 100.0	245 100.0 100.0

APPENDIX 1.b: Time of Evaluation Distribution

EVALTIME	CODER (N) (ROW PCT) (COL PCT)				TOTAL
	1	2	3	4	
INTERIM	64 37.9 70.3	50 29.6 72.5	21 12.4 61.8	34 20.1 66.7	169 100.0 69.0
FINAL	17 30.9 18.7	16 29.1 23.2	9 16.4 26.5	13 23.6 25.5	55 100.0 22.4
EX-POST	5 55.6 5.5	2 22.2 2.9	1 11.1 2.9	1 11.1 2.0	9 100.00 3.7
OTHER	5 41.7 5.5	1 8.3 1.4	3 25.0 8.8	3 25.0 5.9	12 100.0 4.9
TOTAL	91 37.1 100.0	69 28.2 100.0	34 13.9 100.0	51 20.8 100.0	245 100.0 100.0

65

APPENDIX 1.c: PERCENT DISTRIBUTION OF FINDINGS TYPES FOR EACH BUREAU

<u>BUREAU</u>	<u>TYPE OF FINDING</u> (FY82/FY83 PERCENT WITHIN BUREAU)				
	DESIGN	IMPLEMENTATION	INSTITUTION BUILDING	DATA	IMPACT
NEAR EAST	30.1/40.3	48.4/24.5	16.9/22.8	3.7/3.6	0.9/9.6
ASIA	32.3/36.5	41.9/25.3	18.6/25.0	4.8/8.6	2.4/4.7
LAC	35.1/29.2	35.6/37.5	21.5/19.6	3.7/8.2	4.2/5.7
AFRICA	31.9/26.5	38.6/32.3	20.6/25.7	4.2/6.0	4.7/9.5
IMPACT	14.3/30.0	15.7/13.2	40.0/13.3	11.4/1.9	18.6/24.5
SCITECH	20.4/39.7	41.6/24.3	7.1/14.9	26.5/22.3	4.4/9.3
FVA	26.3/36.0	28.9/24.0	36.8/22.7	5.3/9.3	2.6/8.0