

A NUMERICAL RATING OF A.I.D. PROJECTS COVERED IN
A.I.D. IMPACT EVALUATION REPORTS

by

RAYMOND COHEN
PPC/E-DIU

April 1984

BEST AVAILABLE COPY

TABLE OF CONTENTS

	<u>Page</u>
Foreword	ii
Executive Summary	iv
Impact Evaluation Reports Used in This Study	vii
I. Objectives	1
II. Background	2
III. How the Study Was Conducted	7
IV. Analysis	16
Introduction	16
World Wide (All Projects).....	20
Regions	25
Sectors	30
Project Start	37
Project Duration	42
Kinds and Levels of Funding	46
Beneficiary Participation	50
Elapsed Time after Completion	53
V. Conclusions	55
VI. Recommendations	66

FOREWARD

This study was undertaken in the summer of 1983, following discussions and agreement between the then management of PPC/E and this writer about the feasibility and utility of translating the narratives in the issued Impact Evaluation reports into numerical ratings as a means for providing Agency management (managers) with readily available and useable data and supportable judgements regarding AID's performance. Two such attempts had previously been undertaken. One, an academic work, was in process when the study by this writer was begun. The other, an effort by PPC/E's then Director was aborted because of the pressures of other work requirements and for other reasons.

This study has evolved into a "test" or "prototype" model for an approach to accomplishment performance assessment. It is hoped that this "model" will be examined and, if found potentially useful, improved, implemented and employed by AID management and staff for the limited, but important, purpose to which it seems suited. This purpose is intended, broadly, to facilitate and strengthen AID's communication in-house and with others regarding our past performance.

The study also indicates some kinds of analysis regarding the relationships between project characteristics (e.g., length and funding level) and project accomplishments which might be pursued using the "model" outlined in this report. The utility of this analysis has not yet been demonstrated.

A second objective, to informally assess the usefulness of the IE reports generally, was introduced into the study when it was already underway. The study effort in pursuit of that objective centered on informal discussions between this writer and a fairly large number of other AID staffers.

The cooperation of colleagues throughout the Agency, particularly in sharing their knowledge and understanding with me is much appreciated.

SUMMARY

The AID Impact Evaluation program, during the four year period 1979-1983, produced a substantial number of Impact Evaluation and other reports. Forty-five of the Impact Evaluation reports evaluating 73 projects seemed to have enough in common, e.g., in terms of kinds of impact looked at and kinds of projects evaluated (essentially Development Assistance funded), to warrant the reports being treated as a database for an attempt to assess AID accomplishment performance in numerical terms.

Toward the objective of conducting such an assessment, 11 "accomplishment factors" were established (covering such matters as delivering benefits to the poor and training needed manpower); a rating or "scoring" scale running from 1 (worst) to 10 (best) was also established. The factors were then scored for each project by this writer and, in a roughly estimated 50% of the factors, by a second person. The data (scores) were then computerized and analyzed.

On the basis of this analysis, it appears that AID's accomplishment performance over the years has been very good. In broadest terms, 13 projects (18% of the 73 projects included in this report) were judged on the basis of Impact Evaluation reports, to be Marginal or Unsatisfactory; the remaining 60 projects (82%) were evaluated as Satisfactory or better.

In addition, only six of the 73 projects (8.2%) were evaluated as being Marginal or Unsatisfactory in terms of their impact on poor people; for institution building and manpower training impact the Marginal and Unsatisfactory performance level was almost as low: 14.0% and 12.5% respectively. Put positively, 91.8%, 86.0% and 87.5% of the evaluated projects were rated Satisfactory or better in terms of benefiting the poor, establishing or up-grading institutions and training needed manpower.

On the other hand, about a third and a quarter of the projects were rated Marginal to Unsatisfactory in terms of their impact on women and on government policies.

The data extracted from the IE reports seem sufficient to support analysis needed to reach firm judgements regarding AID accomplishment performance. If this idea is accepted, means might be developed and employed to broaden the base of information used in such analysis and increase the relevance, reliability and credibility of its judgements. The Conclusions and Recommendations sections of this report outline ways by which the Agency might move in that direction.

When the accomplishment ratings are considered in terms of certain project defining characteristics, e.g., project duration, sector, and level of funding, other details of performance are suggested.

For example, the numerical ratings suggest that projects started in the 1950s and 1960s performed better than those started in the 1970s. Even more interesting (and, I believe, in need of further examination), projects started in 1973 and prior years were rated higher than projects started in 1974 and later years in terms of both overall project performance and in such specific accomplishment areas as benefitting the poor, institution building, and affecting policy.

Similarly, projects at the highest funding level (AID input) and the longest duration scored highest while, in both instances projects at the opposite extreme (lowest funding level-\$1.0 million or less; and shortest duration-five years or less) scored lowest.

Much more analysis of these kinds could be done. However, the useful applications and utility of such analysis and conclusions flowing from it are uncertain.

List of Impact Evaluation Reports Used in this Study

<u>REPORT NO.</u>	<u>REPORT DATE</u>	<u>COUNTRY</u>	<u>PROJECT(S)</u>
1	12/79	Columbia	Small Farmer Market Access
2	12/79	Kenya	Kitale Maize
3	05/80	Thailand	Potable Water
4	05/80	Philippines	Small Scale Irrigation
5	05/80	Kenya	Rural Water Supply
6	06/80	Liberia	Rural Roads I&II
	06/80	Liberia	Rural Roads II
7	06/80	Sierra Leone	Rural Penetration Roads
8	08/80	Morocco	Food Aid & Nutrition Education
9	10/80	Senegal	Rural Health Care
10	10/80	Tunisia	Care Water Projects
11	11/80	Jamaica	Feeder Roads
12	12/80	Korea	Irrigation
13	12/80	Thailand	Rural Roads
14	12/80	Central Amer.	Small Farmer Cropping Systems
15	12/80	Philippines	Rural Electri- fication
16	12/80	Bolivia	Rural Electri- fication
17	01/81	Honduras	Rural Roads
18	03/81	Philippines	Rural Roads I&II
	03/81	Philippines	Rural Roads II
19	05/81	Nepal	US Aid To Educa- tion
20	05/81	Korea	Potable Water System
21	06/81	Ecuador	Rural Electri- fication
22	10/81	Costa Rica	Rural Electri- fication
23	09/81	Nigeria	Teacher Education
24	10/81	Peru	Water Health Services
25	10/81	Thailand	Mobile Trade Training Units
26	01/82	Kenya	Rural Roads
27	01/82	Korea	Agricultural Research

List of Impact Evaluation Reports Used in this Study

<u>REPORT NO.</u>	<u>REPORT DATE</u>	<u>COUNTRY</u>	<u>PROJECT(S)</u>
28	01/82	Philippines	BICOL Integrated Area Develop
29	02/82	Indonesia	Small Scale Irrigation
30	02/82	Guatemala	Institute of Ag Sci and Tech
31	03/82	Sudan	RAHAD Irrigation
32	05/82	Panama	Rural Water
33	05/82	Nepal	Food Grain Tech- nology
34	05/82	Thailand	Agricultural Research
35	06/82	Pakistan	On-Farm Water Management
36	06/82	Korea	Health Demonstra- tion
37	08/82	Kenya	Radio Correspondence Education
38	09/82	Philippines	Low Cost Alt. To Primary Ed.
40	12/82	Francophone Africa	Entente Fund African Enterprise
41	03/83	Panama	Housing Guarantee Program
42	04/83	Bangladesh	Small Scale Irrigation
43	04/83	Egypt	Rural Improvement Services
44	05/83	West Africa	West Africa Rice R&D
46	06/83	Paraguay	Rural Education Development
48	10/83	Tunisia	Wheat Development

I. OBJECTIVES

The review of the entire body of completed Impact Evaluation reports on which this report is based was undertaken with the limited expectation that it might help to clarify some Agency views regarding its past performance in terms of development results or accomplishments achieved. (These views tend to be harsh, not only self-critical but broadly negative as well). Performance in "accomplishment" terms, i.e., as used in this report, is distinguished from performance in terms of implementation (operational or administrative) matters, e.g., project or program adherence to time schedules and budgetary guidelines. Similarly, performance in terms of "accomplishments" as used in this report is both more general than and distinguished from performance in traditional macro-oriented areas related, for example, to trade, balance of payments or economic growth rates.

In line with this limited expectation, the main objective of this report is to provide some numerical indicators of AID accomplishment performance over the years as that performance was assessed in the Agency's series of Impact Evaluation (IE) reports prepared during the past four years on selected AID-assisted projects. This report also attempts to analyze the numerical indicators with a view to identifying broad project characteristics (such as sector and funding

level) which might correspond with better or worse than average accomplishment levels.

During the course of the review of the IE reports (and attendant discussions) a second and ancillary objective was established. It is to provide a general assessment of the utility or use level to date of the IE reports.

II. BACKGROUND

In 1979, AID-initiated a new and fairly substantial effort (the Impact Evaluation program) to evaluate AID-assisted projects grouped into several major sectors or sub-sectors selected by AID management. In broad terms, the main objectives of the Impact Evaluation (IE) program were:

- to provide senior officials of the Agency "documented" information and judgements regarding the effect ("impact") of AID efforts on the economic development of cooperating countries, with "development" seen largely in terms related to improved conditions of life for the poor and/or improved government (primarily) and private capability to effect further improvements in those conditions); and

- to identify those relationships between the evaluated projects and their impact which might have wider relevance and applicability and which, therefore, could be used by Agency staff as guidance or at least reference points in the planning, implementation and management of other projects in other locations.

While the Agency has conducted and continues to conduct other project evaluations, the IE program is unique in its world-wide focus and its methods and products.

Since inception of the IE program in late 1979, the Agency has conducted 61 IE studies in 34 countries (and for three regional entities) representing all four AID geographic bureaus: Africa, eight countries and two regional entities; Asia, 10 countries; Latin America and the Caribbean (LAC), 11 countries and one regional entity; and Near East/North Africa (NENA), five countries. Fifty-two of these studies have resulted in individually issued IE reports covering 26 countries (and two regional entities) and again, distributed among all geographic bureaus.

Forty-eight of these reports were completed in time to be reviewed as part of this study. However, only 45 of them, each of which (with one exception) deals with Development Assistance (DA) projects, were included in this report. The three IE

reports excluded deal with PL480 Title I programs. We concluded (after considerable deliberation) that the exclusion was called for because of differences in the kinds and range of primary objectives and methods of implementation between DA projects and Title I programs. On the other hand, we did include one IE report which deals with two Housing Investment Guaranty projects because their objectives, implementation and defined time-focus seemed to give them much in common with the DA projects for the purposes of this study.

All of the IE studies covered in this report were conducted and their reports prepared primarily by AID direct-hire staff, representing virtually all major units of the Agency, in cooperation with selected American academic and commercial contractors, direct-hire employees of other USG agencies and academics and government employees of cooperating host countries. The reports therefore, reflect a broad range of viewpoints and a tendency toward balanced, objective assessment. (Put differently, the reports impress this writer as being informed, in tone and substance, more by a critical skepticism and realism than by a charitable or congratulatory indulgence or defensiveness).

A range of topics to be covered in the IE reports and a general report format, including a preferred length have been prescribed by AID/W (PPC/E). However, no attempt has been made either to standardize the reports to facilitate comparative

(inter-report) analysis or to have the reports present direct quantified assessments or measures of impact in terms of planning document-stated objectives or other desired or perceived objectives. While the quality of the reports is, inevitably, varied, virtually every one of them is well-focused and thoughtful and presents information, insights and opinions that would seem to reward careful reading.

On the other hand, it should perhaps be noted that while the reports might yield a consensus on some matters, there are many development issues treated in one or more of the reports on which significant differences of views are argued and remain beyond consensus. From the relative importance and validity of "equity" or "wealth redistribution" as AID objectives to the role of market forces in education activities, to the "need" for or value of government subsidies, many positions are taken, sharply contrasted and strongly advocated. The reports make it clear (to me, at any rate) that uncertainty (of needs, objectives, required actions and results) will inevitably characterize AID efforts in the "real world," and frustrate our attempts to distill our highly varied experience into neat theories to guide our future actions. At the risk of being taken to task for moralizing, I would suggest that we (AID) must learn to be more comfortable with this circumstance and less taken with the notion that we can and must change it. We should accept the truism that "theory follows practice" and

appreciate our forward position and role as innovative (if self-doubting) practitioners of the fine^{art} ~~art~~ of aid in circumstances of incredible diversity and difficulty.

The 45 IE reports covered in this study evaluate 73 projects, most of which fall into six sectors (or subsectors), with project starting dates as early as 1954 and as recent as 1978. Funding levels, completion dates, levels of accomplishment and other project characteristics are also quite varied. Thus, while the reports deal with only a very small selection from AID's cumulative project portfolio and do not collectively comprise a scientifically determined (or statistically validated) "sample" of that portfolio, they do present data on what seems a representative cross section of AID development activities and accomplishments randomly distributed in time, place, size and substance.

In the summer of 1983, PPC/E decided to conduct a study designed to "grade" or "score" AID's past "accomplishment" performance using the available IE reports as the information source. The study as conceived and conducted is based on a rendering of the narrative assessment of project impact presented in 45 of those reports into quantified numerical ratings for a number of major indicators of accomplishment, e.g., delivering benefits to the poor, affecting policies, training people. These ratings were then reviewed and analyzed in terms of several major (and obvious) project "defining"

characteristics such as field of activity (sector or sub-sector) and funding level. Again, the chief aim of the study was simply to gain a "documented" sense of or feel for the quality of Agency performance in terms of "impact" or accomplishment; a lesser aim was to try to identify relationships between that performance and some common project characteristics.

The subordinate objective of assessing the usefulness of the IE reports was introduced after the study was underway and represents, essentially, an informal insertion to it.

III. HOW THE STUDY WAS CONDUCTED

As a first step in conducting the study, a set of generalized accomplishment factors (a "report card") was developed consisting of 11 development results-oriented conditions or indicators by which AID traditionally has defined its institutional and work objectives and measured its activity effectiveness/accomplishments. The 11 indicators (or factors as I have chosen to call them), do not, of course, comprise an exhaustive list of broad accomplishment indicators. They do comprise a list of what this writer and AID staffers with whom he has consulted consider both among the most "important" factors and the most useful within the practical constraints imposed by information available in the IE reports and the previously stated objectives of this study. The 11 factors with a brief description or definition of each follows:

- A. General impression, a broad summary perception of the project impact; not an average of the other 10 factors. (For some aspects of the analysis, General impression is treated separately from the other 10 factors).

- B. Benefits delivered to or impact on targetted beneficiaries or, if no beneficiaries were targetted, to the rural poor (defined in broad terms, i.e., not necessarily limited to the "poorest of the poor" or the "poorest 10% of the population").

- C. Benefits delivered to or impact on women, particularly but not exclusively, poor rural women.

- D. Effect on host country policies, including technical, social, political and economic ones of local or central authorities; in this context "effect on" could include 'support for" as well as "changes in" existing policies.

- E. Effect on government institutions and capabilities, in physical, procedural or other terms.

- F. Effect on the "stock" of trained manpower and skills, whether resulting from efforts in-country, in the US or in third countries.

- G. Appropriateness and "take" of the technology that project was intended to transfer.

- H. Effect on the host country private sector in terms of either increasing its capabilities or utilizing it in the project or on other development activities.

- I. Effect on the physical environment. (This factor was often evaluated in damage limitations terms, e.g., a road project that caused little or no erosion or drainage problems might warrant a relatively high rating.)

- J. Substantive activities/benefits of project are continuing to be implemented/delivered to the target population and/or other poor people. (Sustainability).

- K. Substantive objectives of project are being pursued or planned in other locations without AID assistance and using project in whole or in part as a model. (Replicability).

For any project not yet completed when the IE study of it was done, the factors were rated, as appropriate, on the basis of expected/prospective rather than achieved impact as reflected in the relevant IE report.

In addition to these 11 factors, a project average (or average of factors) score is also used to consider project performance. As the name implies, the project average (preferably) or factors average score simply represents the averaging of the 11 factors (or some lesser number in some circumstances where all factors could not be scored/rated) for a given project.

While the 11 factors used here are traditional in the sense that they represent long standing AID values and perceptions of its objectives, they also appear to this writer as being reflective of both 1973 "New Directions" and current AID conceptions of its role, capabilities and directions.

Separate reviews of the IE reports were then conducted by this writer and, where feasible, (for approximately half the reports) by one person representing the team which had prepared the report under consideration. On the basis of these two reviews, two separate sets of ratings of the factors were prepared using a rating scale of 1 to 10, with a "1" rating being warranted where the project concerned was terminated

because of the negative impact of the factor under consideration (which did not occur in this set of 73 projects) and "10" where the positive impact of the factor was extremely high (superior or outstanding). These two rating sets were then reviewed by this writer and a consolidated rating or "report card" prepared.

The closeness or similarity of patterns of the two separate rating sets for most projects and individual factors was striking and, among other things, made quite easy the task of resolving differences between the two score sets, largely by averaging. In only a very small number of cases was consultation between the two raters necessary; in an even smaller number of instances this writer exercised his right of last review by using a score for a factor based on his personal knowledge of the project. Beyond this, the closeness of the two separate score sets reflects creditably, I believe, on the straight-forwardness and communication effectiveness of the IE reports.

The stops on the 10 point scale used in this study are simply and lossely defined as follows:

- 1: project terminated for cause related to the considered factor;

- 2: significant net negative impact;

- 3: very unsatisfactory;
- 4: unsatisfactory;
- 5: marginal;
- 6: satisfactory;
- 7: fully satisfactory;
- 8: highly satisfactory;
- 9: superior; and
- 10: outstanding.

Where a particular factor was not relevant, i.e., the concerned project did not attempt to affect the factor and/or did not have a perceived impact on that factor, a "not applicable" was used. Where a factor was affected (or "impacted") by the project, even if unintentionally so, a numerical score was given to the factor.

After all 73 projects were "scored" for each accomplishment factor (or as many as information in the relevant IE report warranted in the view of this writer and, where one existed,

the second "scorer"), another scale was developed to condense the 10 point scale into one with a lesser number of "stops," each defined in verbal terms. This was done primarily to try to simplify and facilitate communication about the ratings analysis. The verbal scale has five stops or categories: Unsatisfactory (which equates with numbers 1-4 on the 10 point scale); Marginal (number 5); Satisfactory (6); Highly Satisfactory (7&8); and Outstanding (9&10).

As a device for achieving maximum simplicity, we then established a two point scale: Satisfactory (6-10 on the 10 point scale) and Unsatisfactory (1-5).

Next, seven basic or traditional project-defining characteristics, e.g., sector and level of funding, were selected against which to review the factor and project scores in an effort to identify significant relationships that might assist AID staffers in future project planning and in clarifying notions regarding the effect of these characteristics on Agency accomplishment performance. Again, this list of characteristics is both meager and obvious; though we believe each of the characteristics is important, others, e.g., implementor (PVO, university, private firm) and level of host country commitment, might have been considered as well. The composition of the list used in this report was determined, to a very large extent, by information available in the IE reports.

The seven characteristics used are:

- Geographic region of project;
- Project sector (or sub-sector);
- AID funding level and kind of funds (grant, loan, mixed);
- Project duration;
- Beneficiary participation level;
- Elapsed time since project was completed; and
- Year of project start (by decade and by pre and post- "New Directions Mandate" period).

Beyond this, some consideration was given to establishing a small number of project implementation characteristics against which to review the project and individual accomplishment factor scores. The chief implementation characteristics in which we were interested dealt with budgets, staffing, and scheduling. However, there was insufficient data in the IE reports to enable us to pursue this line of analysis. Nonetheless, it seems to this writer that an examination and evaluation of AID project and program performance in terms of

implementation with a view to gaining an improved understanding of the relationships between "good" or "bad" implementation and good or bad "accomplishment" (in both quantitative and qualitative terms) might be useful.

All of the data generated through the above outlined process was then computerized. Analysis of the computerized data presented in this report consists largely of comparing individual accomplishment factor scores (including those for project average) with other accomplishment factor scores within a project-defining characteristic and with World wide (i.e., all 73 projects) accomplishment factor scores.

In pursuing the second objective of this report - broadly assessing the usefulness of the IE reports - we relied on informal but relatively well structured discussion with 22 field-oriented AID staffers in three of the four regional bureaus, and S&T and PPC. All of these discussants have had both considerable experience with AID and various levels of exposure to or experience with project evaluation. All were familiar, again to varying degrees, with the IE reports; several had particular views regarding the purposes of the IE reports though none disagreed in substance with the purposes as presented to them by me much as outlined above under Section II Background. (Actually, the number of discussants was more than

30; however, the discussions with several of these was relatively brief and not particularly informative or instructive for purposes of this report). In a sense, the people with whom I discussed the matter represent something of a sample of AID's older (I hesitate to say that) and more experienced staff and therefore, perhaps well-considered and "real-world" - tempered views on this matter.

The structure of the discussions was based largely on responses by my interlocutors to two main questions posed by me on the IE reports and the subsequent dialogue flowing from those responses. In a majority of the discussions a "set" of other questions and a pattern of dialogue emerged. The two questions (not always phrased precisely the same) were:

- What is your general impression of the quality and utility of the IE reports?

- To what specific uses have the reports, individually or collectively, been put?

IV. ANALYSIS AND DISCUSSION

Introduction

Of the 73 projects included in this study, 50 (68%) were evaluated by teams which included one or more non-AID persons. The lowest scoring of the 73 projects had no non-AID members; the highest scoring did. Of the four lowest scoring projects

(rated under 5.0, i.e., Unsatisfactory) three had outside (non-AID) members and one did not; of the five highest scoring projects (above 9.0, i.e., outstanding) four had outsiders and one did not. Finally, the average score of the projects evaluated with outside team members was almost identical to that of the projects evaluated by teams comprised of AID staff only; both were, of course, very close to the World Wide project average score. In statistical terms (and in practical ones as well) the presence (or absence) of non-AID persons on the IE teams seems to have had no significance in terms of ratings or scores.

As indicated earlier, about half the 73 projects were numerically "scored" by two separate raters working independently and using to a very large extent, only the relevant IE report. These separate ratings were very similar for most accomplishment factors within any individual project and within the entire sample.

All of the above suggests to this writer that the IE reports individually and collectively, represent professionally objective, balanced and reliable assessments of the projects considered. In effect, the reports constitute a "database" which well serves the main purpose of this study and might readily be expanded to serve other purposes, particularly that dealing with the provision of information on AID performance to interested AID and non-AID groups and individuals .

On the other hand, it should be noted that the ratings for one or more factors in most projects are, in fact, based on extremely limited information or views in the IE reports; such ratings are, to a large extent based on impression rather than analysis of textual data and judgement. It should also be noted that the 73 projects comprising the "sample" for this study represent an extremely small proportion of the projects comprising the AID portfolio in any one year let alone the roughly 30 years covered by the IE program and thus this study.

Project Sample Profile

The 73 projects treated in this study are distributed among 29 countries in four regions: Africa, 13 projects in 8 countries; Asia, 35 projects in 7 countries; LAC, 18 projects in 11 countries; and NENA, 7 projects in 3 countries.

The 73 projects represent six discrete sectors and sub-sectors (63 projects) and one "Other" category. The 63 projects are distributed among the six sectors as follows: Agricultural Research, 8 projects; Education 15; Irrigation 8; Potable Water 10; Rural Electrification 10; and Rural Roads 12.

Thirty-six of the projects (virtually 50%) were started in FY 73 and prior years. The other half, those started since 1974, represent the "New Directions Mandate" period generation

of projects. By decade: five projects were started in the 1950s; 19 in the 60s; and 49 in the 70s, all in 1978 and prior 1970 years.

In terms of duration, 22 projects (30% of the sample) lasted (in funding terms) five years or less; 36 (about half the sample) lasted 6-10 years; and 15 (about 21%) lasted more than 10 years.

The US funding input pattern, comprised of four categories, breaks-down as follows: 18 projects (25%) funded at \$1.0 million or less; 37 projects (50%) funded at \$1.1 - 10.0 million; 11 projects (15%) funded at \$10.1 - 20.0 million; and 7 projects (10%) funded at \$20.1 million or more.

Thirty-four of the 73 projects in the study were financed (AID input) with grant funds only; 30 were financed with loan funds only and the remainder, nine projects, were financed with a mix of grant and loan funds.

Seventeen of the 73 projects (23%) were not yet completed at the time their respective IE study was conducted. Of the remaining 56 projects, 25 (35% of the 73 project total) had been completed from one month to three years; 22 (30%) had been completed four to eight years; and the remainder, nine projects (12%) had been completed eight years or more.

Finally, 24 of the 73 study projects (33%) were considered to have had little (or low) beneficiary participation in terms of project planning or implementation; 21 projects (29%) had a moderate (medium) degree of beneficiary participation; and 19 projects (26%) a good or better (high) degree of participation. No judgement on beneficiary participation was reached on the remaining 9 projects (12% of the sample).

World-Wide (All Projects) Analysis

The World-wide project average score (covering all 73 projects and all factors where rated) on the 10 point rating scale is 7.28, virtually identical to the World-wide Factor A, General impression, average score of 7.29. This closeness of scores would seem to validate the "realism" of the General impression rating which was intended as a broad "sense" or judgement of the project's impact, reached independently and not by an averaging of the scores of the other factors.

The highest scoring factor World-wide is Factor F, Effect on the "stock" of trained manpower and skills, with an average score of 7.86. This factor scored highest in three of the four geographic regions (see the "Regions" section below). In this connection, it might be noted that despite the factor's high score, very few of the IE reports devote much attention to the matter of training, either its accomplishments or problems.

Similarly, very few of the reports offer "lessons learned" or recommendations dealing with training. The entire area seems to be dominated by a widely accepted (if not fully articulated) set of attitudes centered on the ideas that training is close to an absolute good and necessity in development terms; AID is handling it well; and AID and the concerned LDCs are getting good pay-off from it. The matter generally seems shielded from the high "adrenalin flow" argumentation that seems to attach to many other aid (and AID) matters. The next highest World-wide factor score, 7.64, was recorded by Factor J, Substantive activities/benefits of project are continuing (Sustainability).

The lowest scoring factor World-wide is Factor I, Effect on the physical environment, with an average score of 5.84. The second lowest factor score, 6.48, is recorded by Factor C, Benefits delivered to or impact on women. Both of these factors scored at or near the bottom in all four regions, with one exception for Factor C.

The relatively low score achieved by Factor I to a large extent seems to reflect the newness of physical environment as a subject within our common everyday consciousness and the attendant lack of (or weak) understanding regarding the environmental implications of many projects. The effects of certain road projects or irrigation projects, for example, on land stability, distribution (of ownership) and use were apparently not adequately anticipated or appreciated (nor could

they have been) during the planning and implementation of those projects; nor were the social dynamics and their subsequent or further effect on the physical environment. The raising of these kinds of issues (or, often, "second generation" problems) in a number of the IE reports is encouraging; it seems to reflect a "raised consciousness" level and might offer some prospect of improved performance in this area in the late 1980s and beyond by the portfolio of projects started in the 1970s and later years.

The remaining six factors are ranked with their World-wide average scores as follows:

<u>Factor B, Benefits to poor</u>	7.60
<u>Factor K, "Replicability"</u>	7.59
<u>Factor E, Effect on government institutions</u>	7.49
<u>Factor G, Appropriateness of technology</u>	7.15
<u>Factor D, Effect on host country policies</u>	7.14
<u>Factor H, Effect on host country private sector</u>	6.76

The evidence available in the IE reports used in this study is not sufficient for this writer to account for the relatively low ranking achieved by ~~some~~ factors H, Effect on host country private sector; D, Effect on host country policies; and G, Appropriateness of Technology (ranked ninth, eighth and seventh respectively among the 11 factors). Nor do the reports provide sufficient indications of possible "trade-offs" between

different kinds of accomplishments or likely changes in their relative and absolute performance in the future.

The newness of the first two mentioned factors (H & D) as areas of deliberate and articulated (as opposed to almost "covert") emphasis might be part of the explanation. Another contributing circumstance might be the widely-held perception of short-term inconsistencies between the pursuit of these kinds of accomplishments, on the one hand, and those centered on the direct delivery of benefits to the poor and women, on the other. In recent years, the Agency has tended to try to resolve these "inconsistencies" (as a matter of general policy) in favor of the latter kinds of accomplishments or objectives.

In this regard, the majority of the IE reports are (or to this writer appear to be) supportive of the concepts of aid as an instrument to help improve the conditions of life of the poor immediately and directly and to achieve some improved degree of economic equity through redistribution of wealth. However, a number of reports (with varying degrees of apparent defensiveness) suggest broad alternative (and equally traditional and semi-ideological) concepts, i.e., aid as a means of helping to create lasting and increasing opportunities for the poor or, a variation, aid as a means for facilitating development processes. Some of the reports dealing with Agricultural Research, Feeder Roads, Rural Electrification and (even) Potable Water make points to this effect.

The view is perhaps best stated in a report on a roads project which argues that while "A road does not change a social system (redistribute wealth) ... it (does) widen the range of opportunities. The opening of the roads has made nobody worse off; but the rich have obviously gotten richer. They were in a better starting position. However, in terms of percentage of disposable income - of ready cash - that of the poorest people has quite likely increased most and they, rather than the more affluent, are newly within reach of health, credit and technical service for their farms, and education for their children." And, finally on this point (and quoted also because of its implied caution regarding our sometimes facile reading of others' expectations), an IE report that states: ". . . . there are (in Thailand) clearly recognized patron-client relationships, mutual obligations in a wide range of social and economic situations, which place no stress on equal access to the benefits of economic development. The rich and -- with or without great means -- the ambitious and the venturesome have more quickly and more successfully taken advantage of opportunities created by new all weather roads dramatically increasing the range and volume of effective transfer of goods, and of ideas. But the poor have also gained. Indeed in terms of relative gains theirs may be the greater"

A study - I would suggest it be done by a university with good sociology/anthropology/economics capability under a three to five year AID contract - of the real and apparent

differences in these concepts or approaches and of AID's application of them, would, I believe, be instructive and helpful.

On the five point verbal scale, "translated" or compressed from the 10 point scale, five projects (7.0% of the 73 project base) were rated Unsatisfactory; eight (11%) were rated Marginal; 11 (15%) were Satisfactory; 44 (60%), Highly Satisfactory; and five (7.0%), Superior.

Condensed into a simple Satisfactory or Unsatisfactory assessment, 60 of the total 73 projects (82%) were Satisfactory (or better) and 13 (18%) were Unsatisfactory (Marginal or worse on the five and 10 point scales).

Regions

The project average score and the General impression factor score on the 10 point rating scale by region are: Africa, 6.45 and 6.46 respectively; Asia, 7.50 and 7.40; LAC, 7.88 and 8.17; and NENA, 6.21 and 6.00. These region scores compare with the World wide project average and General impression factor scores of 7.28 and 7.29.

As can be seen, the Asia and LAC bureaus score higher in these two regards than the World wide average; Africa and NENA score lower.

It will also be noted that the General impression score is higher than the project average score in LAC and lower in NENA. Among other things, this might reflect a relatively high sense of development movement and promise in LAC (the "general impression" is that the project, the "actionable" and controllable part of the more abstract development process leading to results, is working or has worked, perhaps better than any individual accomplishment or aggregation of accomplishments, might suggest). The opposite (a kind of limited expectations attitude) might help to explain the reverse relationship of the project average and General impression factor scores in NENA.

In all the regions, except LAC, the highest single factor score was achieved by Factor F, Trained manpower and skills. In two of these regions, Africa and Asia, the score for this factor was substantially above the respective region's project average score; in NENA it was moderately higher. In LAC, Factor F ranked only eighth among the eleven factors, scoring 7.50 compared to the LAC project average of 7.88. Measured against the World Wide average for this factor of 7.86, only Asia with an 8.39 did better.

Factor B, Benefits to poor and Factor C, Benefits to women were ranked by the LAC Bureau first and fifth respectively among the eleven factors. Factor B scored 8.67, well above the LAC project average and the World Wide average for this

factor. Asia also scored above both its project average (modestly) and the World Wide average (barely) in this factor. The Africa bureau's Factor B score was modestly higher than its project average; the NENA score was slightly lower. Both scored below the World-wide average for this factor.

LAC's score of 7.84 for Factor C, Benefits to women against an LAC project average of 7.88 and a World Wide average for Factor C of 6.48 was very good relative to the scores for this factor achieved by the other regions. Asia's Factor C score of 6.43, like the Africa and NENA scores in this factor, was below both the region's project average and the World Wide average for the factor.

The scores (and the consequent rank among the 10 factors - excluding Factor A) achieved by those two factors (B & C) in all regions except LAC were disappointing and surprising to this writer, in the case of Factor B, and disappointing (but not surprising) in the case of Factor C. The ranking (or place within the 10 factors) is as follows:

TABLE I

	<u>World Wide</u>	<u>Africa</u>	<u>Asia</u>	<u>LAC</u>	<u>NENA</u>
<u>Factor B</u>	3	4	6	1	5
<u>Factor C</u>	9	8	8	5	9

To this writer's pleasant surprise, all regions scored well to very well, (with two partial exceptions) in Factor J, Sustainability and Factor K, Replicability which ranked second and fourth among the eleven factors on a World Wide basis. The rankings for these two factors are as follows:

Table II

	<u>World Wide</u>	<u>Africa</u>	<u>Asia</u>	<u>LAC</u>	<u>NENA</u>
<u>Factor J</u>	2	7	2	4	3
<u>Factor K</u>	4	3	3	6	4

While inter-region comparisons do not seem warranted within the context of the limited data available, some patterns of accomplishment factor scores might be noted.

- The LAC pattern seems quite different from that of Africa and NENA particularly and of Asia to a lesser degree. For example, LAC did poorest (excluding the Environment factor) in terms of impact on policies, institutions and trained manpower. The other three bureaus did either best or well in these three accomplishment factors.

- On the other hand, LAC scored very well in terms of impact on the poor, women, the private sector, and technology while the other three scored only poor to fair in these areas.

- Despite variations in their scoring patterns, the LAC and Asia projects did well in virtually all factors relative to the World-wide scores. And, of course, Africa and NENA did poorly in comparison with those scores (and even more poorly relative to the LAC and Asia scores themselves).

The IE reports do not provide much information, direct or indirect, that would help to explain this pattern variation. One obvious line of speculation (supported to some extent by a perhaps too careful reading of the reports) suggests that because of the LAC region's relatively advanced institutional and manpower infrastructure, and its long history of and familiarity with the substance, purposes, processes and rhetoric of US assistance, the LAC Bureau/Region tended to under-emphasize the institutional factors in favor of more direct delivery of benefits to the poor and women. To a lesser extent the same situation and reaction seem to have prevailed in Asia.

Viewed simplistically, perhaps, this suggests that institutional and trained manpower capability (and, some would argue, physical infrastructural support) function, among other ways, as something of a precondition for or "enabler" of

successful efforts to provide benefits directly to targeted poor populations. In a sense, one might note a correspondence (suggested to me orally by an IE team member) of this view with the cliché that "it takes money (in this instance institutional, skills and physical infrastructure capital) to make money (in this instance benefits delivered to the rural poor)."

The "phenomenon" of substantially higher levels of all accomplishment factor scores achieved by LAC and Asia compared with those achieved by Africa and NENA appears significant to this writer and perhaps warrants further study. In my own view, if sufficient data were available by country (as was not the case in this study) intra and inter-country performance comparisons/analysis would probably be helpful in trying to assess the value of universal as opposed to region - specific or to country^{-specific} aid strategies.

Sectors

Sixty-three projects (86% of the 73 project universe) are distributed among six major sectors: Agricultural Research, 8 projects; Education, 15 projects; Irrigation, 8; Potable Water, 10; Rural Electrification, 10; and Rural Roads, 10. The remaining 10 projects (primarily in Health, Nutrition and Area Development) comprise a miscellaneous "Other" category or sector. (Ten of the 15 Education projects were covered in one IE report on one country).

In terms of project average score by sector, the Education sector, with 7.99, ranked highest among the seven sectors (including the "Other" sector). It was followed by Rural Electrification with 7.82; Agricultural Research with 7.37; "Other," with 7.21; Rural Roads, 7.14; Potable Water 6.60; and Irrigation 6.55. These sector project average scores compare with the World-wide average project score of 7.28. Excluding the "Other" sector, three sectors scored above the World-wide project average and three below.

The Education sector scored particularly well in Factor D, Effect on host country policies and Factor F, Training skilled manpower, each of which scored 8.53. This sector also did very well in Factor B, Helping the poor (8.40), Factor E, Effect on institutions (8.27), and Factor K, Replicability (8.21). These scores in the "8" range might be compared with the World Wide averages in the cited factors: D, 7.14; F, 7.86; B, 7.60; E, 7.49; and K, 7.59.

The Education sector scored worst (and in the first instance, unaccountably, from this writer's point of view) in Factor I, Effect on physical environment, (5.00 against a World Wide score for this factor of 5.84) and Factor H, Effect on host country private sector, (6.33 against a World Wide score of 6.76).

The Rural Electrification sector (or sub-sector) scored best and very well (in the "8" range) in factors: J, Sustainability; K, Replicability; F, Trained manpower; E, Institutions; and C, Impact on women (in this last, scoring 8.00 - the highest score among the sectors for this factor and, quite high when measured against a World Wide score in this factor of 6.48). The Rural Electrification sub-sector scored worst in Factors I, Environment and G, Appropriateness and "take" of the applied technology. Suprisingly and disappointingly to this writer the sub-sector also did relatively poorly in Factor H, Effect on the host country private sector which ranked eighth among the sub-sector's 10 factors, (though with a score of 7.60 it was higher than the World Wide score of 6.76 for this factor).

The Agricultural Research sector did remarkably well in Factor F, Trained manpower, scoring 9.29 (against a World Wide average of 7.86) which was the highest score for any single factor in the sector cut or arrangement. This sector also did well in Factor E, Institutions and Factor G, Technology. It did least well in Factor D, Policies, Factor C, Women and Factor K, Replicability, (these last represent a worrisome situation which might indicate a need to reduce the pace of centralized agricultural research institution building and high level manpower training relative to the pace of effort in extending research activitiy/results and benefits to the local

level and targetted groups. Further analysis in this "popular" sector seems called for).

Eight of the 10 factor scores in the Rural Roads sector set were very close to their corresponding World Wide factor average scores (three slightly above and five slightly below). In the remaining two factors J, Sustainability and D, Policies, the sector scored appreciably lower.

The Irrigation sector scored below the World Wide average in all factors. Within the sector's own factor rankings it did best (and respectably, i.e., close to the World Wide scores) in Factor J, Sustainability; Factor K, Replicability; and Factor B, Benefitting the poor.

The Potable Water sector also fell short of the World Wide averages in all 10 factors. Within its own set it did best with factors K, Replicability and F, Trained manpower. It did least well with factors I, Environment and H, Private sector. Another low scoring factor (a surprise and disappointment and perhaps a methodological anomaly) was factor C, Impact on women which ranked eighth among the sector's 10 factors (the General impression factor (Factor A) was excluded from this and all other intra-sector rankings)

The low score of Factor C, Impact on Women in the Potable Water sector might reflect the general circumstance expressly noted in ^{one of the IE reports:} "Women were not a specific source of concern in the planning of the

project." It might also reflect a problem confronting virtually all efforts to assess substantive accomplishments. This has to do with determining the point (in terms of time and consequences) at which a condition is viewed as an achieved or "completed" accomplishment (as opposed, say at its most extreme, to a cause of a subsequent disbenefit). This point is treated instructively in two separate IE reports on potable water projects. The first of these reports rates Factor C very high and supports this rating by such statements on the matter as "By bringing water closer to rural households, the new water systems made life easier for women." Also, women "appreciated the convenience of having sufficient water for personal and household needs and they expressed satisfaction to be participating in the amenities of modern life."

The second report rated Factor C as fair. This rating was justified by an extension of the ^{line} ~~reasoning~~ of reasoning given above. As presented in the report, the extension holds that the potable water project resulted in "... three hours of released time for each housewife in the village. Most of the women reported using this time for household chores and child care ..., for tending animals ..., for their families ... (and others might add for farming or second-income generating chores)."

In another report, on feeder roads, the evaluators go so far as to say they are not sure we can ever accurately gauge the success of a project in a final sense. Whether verbally or

numerically measured, our accomplishments can (and should) be judged tentatively at any point in time. Not only our accomplishments but even our attempted "solutions" to development problems should be seen as highly tentative and, even more potentially discouraging, over time and from varied perspectives, perhaps "counter productive." From the perspective of an early worker in the field of evaluation (an academic cum politician) AID should accept this "ordering" of things; it reflects the real but uncertain context in which "practitioners" of development assistance are constantly challenged for "true solutions," yet development "like politics requires accommodation and not solutions which are the frequent, if misguided object of research.))

On the five stop verbal scale the Potable Water and the "Other" sectors were the only ones to record "Unsatisfactory" projects: four of the 10 Potable Water projects fell in this category and one of the "Other." Eight additional projects were rated "Marginal": one each in Agricultural Research, Rural Roads and "Other" (a Nutrition project); two in Potable Water and three in Irrigation.

Of the remaining 60 projects, 11 were rated "Satisfactory," 44, "Highly Satisfactory" and five, "Superior." The following Table III shows the incidence and percentage in each performance category by Sector.

TABLE III
Sector Performance

<u>Sector Number Percentage</u>	<u>Unsatisfactory</u>	<u>Marginal</u>	<u>Satisfactory</u>	<u>Highly Satisfactory</u>	<u>Superior</u>
Ag Research					
No.	-	1	2	4	1
%	-	12.5	25.0	50.0	12.5
Education					
No.	-	-	1	14	-
%	-	-	7.0	93.0	-
Irrigation					
No.	-	3	1	4	-
%	-	37.5	12.5	50.0	-
Potable Water					
No.	4	2	-	1	3
%	40.0	20.0	-	10.0	30.0
Rural Electrification					
No.	-	-	-	9	1
%	-	-	-	90.0	10.0
Rural Roads					
No.	-	1	5	6	-
%	-	8.0	42.0	50.0	-
Other					
No.	1	1	2	6	-
%	10.0	10.0	20.0	60.6	-
Totals					
No.	5	8	11	44	5
%	7.0	11.0	15.0	60.6	7.0

Table IV

<u>Totals</u>	<u>Marginal or Lower</u>	<u>Satisfactory or Better</u>
No.	13	60
%	18.0	82.0

Project Start

Projects started in the 1950s had an average project score of 7.57 and a Factor B, Benefit to the poor score of 8.20. Comparable scores for the 1960's and 1970s were 8.06 and 6.95 for project average and 8.37 and 7.24 for Factor B. Thus the 1960 projects scored highest in overall project average and in Benefitting the poor; the 1950 projects ranked second in these two areas and the 1970 projects last. The same ranking applies to factors: C, Women; D, Policies; E, Institutions; F, Trained manpower; and J, Sustainability. In factors: G, Technology and K, Replicability the 1970 projects ranked second, ahead of the 1950 projects, but by very small margins.

Comparing projects started in FY 73 and prior years with projects started in FY 74 and subsequent years (a 36-37 split) much the same unexpected (to this writer) pattern emerges. The '73 and prior projects score higher in 10 of the 11 factors (including Factor A, General impression) and in project average. In Factor B, Benefits to poor, for example, the '73 and prior projects score 8.08 to the post-73 projects 7.14. In factors: C, Women; D, Policies; E, Institutions; and F, Trained manpower, the '73 and prior scores are substantially/ significantly (almost a full integer - or more) higher than their post-73 counterparts.

The reason (or reasons) for this "unexpected" situation are not evident to this writer. Two circumstances or hypothesis do suggest themselves and would seem to warrant further consideration. The first of these has to do with the idea of "benefits maturation." There is considerable anecdotal evidence (this writer is not conversant with the relevant literature) suggesting that development project achievements tend to increase as the elapsed time after project completion increases or extends. A number of the IE reports, for example, expressly make this point. If the point is valid, it would seem to argue that a realistic reading of the impact of AID-assisted projects begun in the 1970s might not be available before, say, the mid- to late- 1980s. The data in this study appear to support that view. For example, the project average score for projects which had been completed for more than eight years when their IE studies were done was 7.99 as compared with 6.81 for projects completed three years or less at the time their IE reports were issued.

In the same vein, the idea that benefits increase after the project is completed might suggest the need for a more positive assessment of the "sustainability" record of (and prospects for) AID-assisted projects. While there are, I understand, numerous indications (or "horror stories") in apparent support of the opposite view (consistent with AID's perennial concern with and pessimism regarding this "question of sustainability"), the evidence in the IE reports seems to support the positive

view. Of the 11 factors (including Factor A, General impression) rated on a world wide basis, Factor J, Sustainability ranked second highest, following only Factor F, Trained manpower in factor scores. It scored or ranked first among factors in three of the seven sector ratings and second, third and fourth in Asia, LAC and NENA respectively in their regional ratings; (it ranked only seventh in Africa).

The second "hypothesis" (closely related to the first) centers on the still more complex and controversial idea that 1973 and earlier projects, precisely because they were geared (arguably) to more attitudinal and contextual developmental changes, to longer term development objectives, and to more indirect delivery of benefits to poor people than their post-73 counterparts, were better able to generate more lasting benefits in high priority achievement areas. Put another way, a way suggested in several IE reports: "...the trickle-down theory does work" and the pursuit of "deeper" (contextual, attitudinal, institutional, "development-gearing-up") objectives might be as "effective" as more direct, "frontal" attempts in providing benefits to the poor, for example.

Looked-at in yet another way,^a variation of the above, the 1973 and earlier accomplishment performance might reflect the constancy and soundness of AID's purpose or mission and the dependable realization of accomplishment results over time based on that mission. In effect, this perspective suggests that the general development assistance approach followed by AID in FY 74 and later years was in many (perhaps most) essentials much the same as that followed in earlier years. To the extent there were differences they tend to reflect matters of emphases and conceptual formulations (rhetoric, perhaps) rather than fundamental substance.

In any case, it might be appropriate and timely to begin soon to develop a long term plan (running say for five to 10 years starting, say, in 1986 or 1987) designed to assess the differences and relative merits of pre-and post 1973 conceptual and practical approaches to LDC economic development and AID assistance thereto. The same study or, more likely, series of studies, might also review and assess another matter that is touched-on in a number of IE reports: the success of AID efforts to achieve "wealth redistribution" (as opposed - if opposition is inherent in development situations/strategies/objectives - to "increasing national wealth") and the appropriateness of "redistribution" as an AID objective.

Project Duration

Under this characteristic, projects were organized into three categories; the first of these covers projects of five year duration or less (22 projects); the second, those of six to 10 years duration (36); and the third, those of 10 years or more duration (15). In terms of project average and Factor A, General impression, the last duration category (10 years or more) scored best (7.51 and 7.47 respectively), the second group or category (six to 10 years duration) scored second, and the first category (projects of five or less years duration) scored last with scores of 6.93 and 7.05.

Similarly, the 10 years and more projects scored best and the six to 10 years projects second best in Factor B, Benefits for the poor; Factor E, Institution building; Factor F, Manpower training; Factor H, Effect on host country private sector; and Factor J, Sustainability. In these five factors (and in the project average and Factor A, General impression factor) these two duration categories scored above and the five years or less duration category below the corresponding World Wide averages.

In four of the remaining five factors: C, Women; G, Technology; I, Environment; and K, Replicability, the six to 10 years category scored highest, the 10 years and over, second and five years or less, third. And again, in these four

factors the two longer duration categories scored above and the shortest duration category below the corresponding World Wide averages. Only in Factor D, Policies, did the five years or less category score above either of the other two categories, scoring marginally higher than the third place 10 years and over duration category.

The strong correlation between highest overall project and several factor scores (as mentioned above), on the one hand, and the longest duration projects category on the other seems to support the view offered in several IE reports to the effect that a major reason for "success" is the steady AID commitment and application of resources to the project (or program) over an extended period of time. This view calls into question the currently prevailing attitude that merit attaches to project "shortness" of duration and the consequent reduction in aid levels or greater quantitative availability of "opportunities for intervention." It would also seem to have implications regarding the soundness of AID's traditional practice of terminating its involvement in AID-assisted projects as soon as possible after AID funds have been expended. (In this regard, at least one IE report recommends that AID maintain some involvement in completed projects for an indefinite time even if AID has no plans to provide further assistance to any of those projects themselves or to potential follow-up projects. It might be useful to examine the manpower requirements for maintaining this kind of continuing involvement and other costs (benefits likely to result from it).

On the other hand, the "good" performance of the five years or less category in Factor D, Policies (relative to the other two duration categories) might reflect an AID ability to "reach" policy makers in shorter periods of time than is required to effect and demonstrate other kinds of impact. The point here may well be simply that merit or value does not attach to any particular project length; it is incumbent on us therefore to try to make judgements regarding project duration on the basis of specific, i.e., primarily country level, circumstances and independently of pre-conceived notions regarding desirable length.

It might also be noted that the shortest duration projects category had the highest portion of its portfolio implemented by PVOs. The implications of this association (of accomplishment and implementator) might be worth exploring. In this particular instance (not examined closely), the relatively poor performance of the short projects might reflect the inherent difficulties of those kinds of projects, the limited resources applied in them and the (often) mixed objectives which they expressly or implicitly pursue, e.g., PVO capability development, relief, "religious," "experimental," in addition to developmental.

Beyond this, it should be noted that while the two longer project duration categories scored above the World Wide

in all factors, the scores of factors: C, Women; D,

Policies; H, Private sector; and I, Environment, were not "good" relative to the other factor scores in these duration categories.

In terms of the five stop verbal scale, five projects (23%) of the 22 projects in the first category (five years or less duration) were Unsatisfactory (3 projects) or Marginal (two). This compares with no Unsatisfactory projects in the other two duration categories and one Marginal in the second category (six to 10 years) and none in the third (more than ten years).

Beyond this, the scores and rankings are mixed.

- At the Satisfactory level the percentages were fairly close in all three categories: 14.0, 11.0 and 20.0 respectively.
- At the next highest level (Highly Satisfactory) the six to 10 year group registered 61.0% (22 of 36 projects) while the five years or less group, registered 27.0% (six of 22 projects) and the over 10 years group 20% (three of 15 projects, the same as at the Satisfactory level).
- At the Superior level (the highest) the over 10 years projects registered 60.0% (nine of 15 projects), the six to 10 years group, 25.0% (nine of 36) and the five years or less group, 36.0% (eight of 22).

- Combining the two highest categories, (as Highly Satisfactory or better), the ranking of the three categories by percentage of projects is: five years or less, 64.0%; over 10 years, 80.0%; and six to 10 years, 86.0%.

Kinds and Levels of Funding

The 73 projects included in this study were grouped into four funding level categories. The categories and the project distribution among them are as follows:

Table V

<u>Category No.</u>	<u>LOP Funding Level</u>	<u>No. of Projects</u>
1	Less than \$1.0m	18
2	\$1.0 - 10.0m	37
3	\$10.1 - 20.0m	11
4	More than \$20.0	7

In terms of project average, category 4 (more than \$20.0m) ranked first with a score of 8.22; category 2 (\$1.0-10.0m) ranked second with 7.58; category 1 ranked third with 6.74; and

category 3 (\$10.1-20.0 m) ranked last with 6.58.

These scores compare with the World Wide project average of 7.28. It is interesting to note that projects funded at the \$10-20 million level scored lower than projects funded at both lower and higher levels. No explanation for this "phenomenon" readily suggests itself.

Category 4 (more than \$20.0m) scored highest in seven of the 10 rating factors (excluding Factor A, General impression), and second highest in the remaining three factors and in Factor A. Category 2 scored highest in the three factors and Factor A in which category 4 scored second (factors B, Benefits to poor; H, Private sector; and I, Environment), and second in the other seven factors (in which category 4 scored highest). In every instance, both categories 2 and 4 scored above the corresponding factor World Wide average.

Conversely, in all factors, categories 1 and 3 scored lower than the corresponding World Wide factor average. Each of these categories itself scored lowest in five factors and next to lowest in the other five factors.

Category 4 scored remarkably well (in the 8 and 9 range) in five of the 10 factors: D, Policy impact; E, Institution building; F, Manpower training; J, Sustainability; and K, Replicability.

Category 2 scored remarkably well (in or near the 8 range) in Factor F, Manpower; B, The poor; E, Institutions; J, Sustainability; and K, Replicability.

In terms of the kinds of AID funds used (grant, loan or mixed), loan funded projects (30 of them) scored highest in project average and Factor A with 7.43 and 7.57 respectively. Grant only projects (34) scored second in these two areas. This writer was extremely surprised by (and continues to regard with particular skepticism) the lowest rank achieved by mixed grant/loan projects which scored 6.99 in project average and 6.78 in Factor A, both scores significantly below the World Wide scores in these areas of 7.28 and 7.29. (Three of the nine projects in the "mixed" category were in the Irrigation Sector and all three scored poorly, as did the sector as a whole relative to the other sectors. Also, seven of the nine projects were started in the "low-scoring" 1970s).

Generally, the rankings of the three kinds of funding categories in the 10 individual factors (excluding Factor A) seem haphazard and do not reveal a discernable pattern (unless the haphazardness is itself seen as the pattern).

For example, the mixed grant/loan category scores highest relative to the other two categories and very well (in the 8 range) in Factor J, Sustainability and Factor K, Replicability. Loan projects score second in Factor J, and

third in Factor K while grant projects reverse that rank, scoring third in Factor J and second in Factor K.

(The third rank in each factor is below its corresponding World Wide score). Some further exploration of likely reasons for the good performance of the mixed grant/loan projects in Factor J and Factor K might be revealing.

In other factors, Grant only projects scored highest relative to the loan and mixed categories in factors: D, Policies; E, Institutions; and F, Training. Loan projects scored highest in B, Benefits to poor; C, Women; G, Technology; and H, Private sector.

Within the individual funding categories, Grants scored best in Factor F, Training; Factor K, Replicability and Factor E, Institution building. Loans scored best in factors: B, Benefits to poor; J, Sustainability; and G, Technology. (The relationship of Factor B to kinds and levels of AID resource inputs should perhaps be examined more thoroughly). The mixed funding category did very well, as mentioned above, in factors J and K. It also did well (7.75) in Factor F, Training.

Beneficiary Participation Level (BPL)

The data for this project characteristic strike this writer as being truly shocking or in error, or in need of more

Beneficiary Participation Level (BPL)

The data for this project characteristic strike this writer as being truly shocking or in error, or in need of more sophisticated analysis than can be performed at this time. The 73 study projects were grouped into four categories under this project characteristic: those projects with high beneficiary participation (19 projects), those with a medium level of beneficiary participation (21), those with a low level (24) and those for which no judgement could be made regarding the level of beneficiary participation in the planning and/or the implementation of the concerned project (9).

Within this framework the high BPL category scored 6.07 and 6.05 in project average and Factor A, General impression respectively; the medium BPL category scored 7.59 and 7.14 in these two areas (project average and Factor A); and the low BPL category scored 7.95 and 8.13. The high BPL category also scored below the other two BPL categories and below the World Wide average in Factor B, Benefits to the poor and every other factor. The low BPL category scored above the World Wide averages in every factor and very well (above 8.00) in Factor B; Factor E, Institution building; Factor F, Manpower training; and Factor K, Replicability.

In the five point verbal scale high BPL projects had nine of its 19 projects (47.0%) rated Marginal or Unsatisfactory; six projects (32%) rated Satisfactory; and four projects (21%) Highly Satisfactory: No high BPL projects were rated Superior.

The medium level BPL category registered two projects (9.5%) Marginal or Unsatisfactory; two projects (9.5%), Satisfactory; 17 (81.0%), Highly Satisfactory; and none Superior.

The low BPL category had two projects (8.3%) rated Marginal or Unsatisfactory; three (12.5%) Satisfactory; 14 (58.0%) Highly Satisfactory; and Five (21%) Superior.

The scores and rankings outlined above are contrary to this writer's expectations (and current convictions). I am not able at this time to attempt an explanation for the phenomenon of projects having better performance when beneficiary involvement is low rather than high. One thought which suggests itself ^{as} a contributing factor has to do with the kinds of projects included in this sample. For example, one might speculate that relatively high capital intensive (say road) projects, for example, or projects focussed on developing central institutions (e.g., agricultural research projects), might indeed do best with little beneficiary participation.

Another potential explanatory factor might have to do with the imprecise way the characteristic was formulated and imposed on the accomplishment factors. It is perhaps worth noting also that at least two of the IE reports suggest that too much is often made of the merit of or even need for a high level of beneficiary participation (one report states that projects with a high level of beneficiary participation "are often poorly designed and executed"). These reports suggest that attention to beneficiary participation should be a function of the kind of project to be implemented and many other host country circumstances, including local sociological and even political perceptions on the matter rather than a function of AID policy. Opposed to this view is that strongly presented in another IE report and which states that "The development of physical infrastructure should be ancillary to social and political institution building (among other things to serve as) a corrective for those who would add 'a dash of participation to project designs'"

Elapsed Time between Project Completion and IE Study

Under this "characteristic" projects were grouped into four categories shown below with the distribution of projects among them.

Table VI

<u>Category</u>	<u>Elapsed Time</u>	<u>No. of Projects</u>
1	None (Projects not yet completed)	17
2	Three years or less	25
3	Four-eight years	22
4	More than eight years	9

In ~~five~~ conformance with this writer's expectations, categories 4 and 3 shared all rating honors. Category 4 scored highest in the project average, Factor A, General impression and five of the other 10 factors: B, D, E, G and H; category 3 scored second in each of these factors. In addition, category 3 scored highest in the five remaining factors: C, F, I, J and K while category 4 scored second in them. In all instances except one, both categories scored higher than the World Wide average. The exception was Factor I, Environment in which category 4 scored below the World Wide average. This last mentioned scoring phenomenon might constitute a clue to the

thought expressed in several IE reports (particularly those dealing with roads and irrigation) to the effect that, over time, [REDACTED] circumstances surrounding a completed ^{project} sometimes [REDACTED] change and contribute to the deterioration of the physical and social environment, e.g., through altered land use and soil erosion and changes in traditional social attitudes and arrangements. The relationships of this "hypothesis" to the one mentioned previously regarding "sustainability" of benefits and the idea that AID should continue its involvement with projects after they've been completed might warrant further consideration.

As something of an "anomaly," from this writer's perspective, category 1 (projects not yet completed) scored higher than category 2 (elapsed time of less than three years) in all factors but two and in the overall project average. An explanation for this situation might reside in the "optimism of expectations" with which AID staffers (and apparently others) are often afflicted. Because of this we tend to view not-yet completed projects as always rich in potential. Similarly, we seem often to suffer a "deflation of expectations" regarding completed project results in the short-run (often justified, perhaps, due to the complexities and difficulties of project management transfer and of integrating responsibilities/benefits into host country bureaucratic and other "systems").

IV. CONCLUSIONS AND QUESTIONS

AID Performance

- While universal or even widely accepted standards of "good" or "bad," satisfactory or unsatisfactory donor accomplishment performance are not available beyond the limited technique of benefit/cost analysis, the evidence available in, or "interpreted" from, the series of AID Impact Evaluation reports strongly suggests that AID performance has been "good." The main indicators supporting that judgement are the high ratio of satisfactory or better to marginal or worse performance ratings, 60 to 13 (82% to 18%) in terms of projects and seven to one (87% to 13%) measured in terms of AID funding inputs. It seems significant in this regard also that 49 of the 73 projects studied (67%) were considered to be Highly Satisfactory or better, with 11 projects (15%) simply Satisfactory. While these data do not (can not) quantify the development benefits delivered, they do indicate that a substantial level of such benefits, or "beneficial impact", was achieved. (No attempt was made to distinguish between "intended" or "unintended" benefits or impact; both were "toted-up" for purposes of this report).

- AID performance has been fairly consistent over the past 30 years; however, there has been some apparent decline during the past decade and this should perhaps be examined and, if possible, accounted for.

- The constancy of AID performance has occurred against a relatively constant set of Agency objectives as reflected for example in intended beneficiaries, sectors (and sub-sectors) of major attention and key elements of articulated rationales for aid generally and for individual projects and programs. While there have been changes over the years in the relative emphasis placed on particular objectives and our rhetoric of justification, the essential Agency mission has remained remarkably constant. More pronounced have been changes in our implementation strategies and procedures, e.g., partial movement toward decentralization, virtually total movement toward use of non-AID staff for project implementation, and the decline in available manpower and funding for field activities.

One key element in the constancy of our mission may be seen in the relationship of project selection and success, on the one hand, to "political" considerations on the other. Despite several changes in Agency (and US Government) administration during the past 25 years, only two of the 73 projects studied in this report were expressly related by the evaluation teams to political considerations, in one instance benignly and in the other in a negative way unflattering to AID, i.e., a

project was selected and implemented solely because perceived (real or imagined) political considerations argued for it and despite expressed technical arguments against it. This is, of course, an extremely complex issue in itself and is not, in fact, adequately defined simply in terms of "political considerations"; diplomatic, strategic, trade and other interests are also important. Yet, on the basis of a review of the IE reports (which were admittedly and, I would guess, by design, not intended to address this complex issue) it appears that AID actions in the areas of project development (selection) and implementation, are largely free of direct (or even pronounced indirect) influence by these kinds of interests/considerations. In effect, despite some rhetoric and effort to the contrary, AID appears to have (and is generally perceived by host governments and other donors in the field as having) something close to a "pure play" economic development rationale and program. The "real" advantages and disadvantages to the US and to economic development of both this perception and the "reality," which supports it might perhaps warrant further study at this time.

The broad constancy of Agency values and objectives over time is striking and seems to conflict with what seems to be a widely held notion that AID's values, objectives and operations have changed frequently in the past and remain subject to substantial, even dramatic, variation on short notice.

An examination of the reasons for and meanings of this constancy/inconstancy, reality/perception dichotomy might yield some interesting and helpful ideas, for example, in the area of how AID communicates with others in the Executive Branch and with Congress about AID.

- The "credibility" of the IE reports and the evaluation teams who prepared them is a central matter in determining the validity of the judgements regarding AID performance as outlined in the conclusions above. It is clear to this writer that the diversity of persons and organizations represented on the evaluation teams which prepared the IE reports, by itself, argues convincingly for their objectiveness and reliability. The generally critical tone of the reports and the sometimes demanding standards they set for the evaluated projects also suggest their "credibility" and balance - albeit slightly tilted against the project and AID. Some idea of what I mean by demanding in this context can be conveyed by reference to IE reports themselves, quotes from several of which follow (underlining is added by me). "(The project) ...resulted in a smaller number of residential and productive consumers than was possible." "Perhaps the major shortcoming of the project is its failure to have an impact throughout(the country)". "Programs such as ... (this irrigation one) can provide substantial benefits for the rural poor, but can not achieve redistribution of wealth." (The "concept" of wealth *redistribution*

is touched-on in a number of the IE reports).

"Farmers generally feel their lives have been improved but to understand why farmer income fell far short of what it might have been ..." (This severity of standards or demands on evaluated projects probably contributed to something of an under-rating of some projects by this writer).

- A very cursory look at the matter of implementation - accomplishment relationships during the conduct of this study suggests there is not a necessary positive correlation between, say, "good" implementation, i.e., adherence to staffing, budget, scheduling, and coordination plans and requirements on the one hand and a good level of accomplishment on the other. The correlation seems particularly uncertain with projects of which technical assistance is a significant part and in which changes in social and cultural attitudes and practices are sought (expressly or by implication). In effect, it appears that the number and intensity of problems frequently encountered during implementation of AID-assisted projects are not, in themselves, reliable indicators of the quantity, quality or sustainability of likely or eventual project accomplishments. A study of the relationships between implementation and accomplishment, including an examination of the relationships between AID project planning processes and documents, on the one hand, and implementation problems encountered and accomplishments achieved, on the other, might provide useful insights/guidance on AID planning and implementation needs, objectives and processes. (This last (AID processes) is seen by some - both in and out of AID - as

being both overly demanding in terms of functional or operational needs and objectives and too optimistic in terms of our expectations regarding the accuracy of plans, i.e., our predictive capabilities).

Regarding this last - what I for one see as an Agency addiction to and inflated confidence in detailed plans and planning - several reports express a view to the effect that "flexibility" regarding all aspects of project plans (but objectives and timing for example, and resource uses, more so than, say, resource levels) in the face of implementation realities is central to achievement of good accomplishment performance. Put negatively one report concludes that "The evaluation calls into question the utility of rigidly structured development schemes." Another report states that "The basic assumption of institution building projects should be that they will evolve in ways unforeseen during the design phase." And, finally, from a report whose rated project received one of the highest scores, "The project as designed on paper bore little resemblance to the implemented project ..."

- Independently, of the content of this report's conclusions regarding AID performance there is, in my view, a need to consider the following questions:

-- Are conclusions (and the information on which they are based) regarding AID aggregate performance likely to be useful (however defined) to AID management and/or staff?

-- If such conclusions and information are to be useful, to what extent must they be based on a continuing (larger) sample of AID activities to assure a wide acceptance of their validity?

-- How might AID best (cost effectively and reliably) obtain information and conclusions regarding AID aggregate performance?

Responses to these questions are attempted in the Recommendations section of this report.

Relationship between Accomplishment Factors, Project Characteristics and Performance

- The number of rated projects covered in this report comprises a very small portion of the project portfolio of any of the four geographic regions, particularly the Near East and North Africa region, and of any sector "portfolio." In most respects and for most purposes that portion strikes this writer as too small to support a satisfactory degree of confidence in conclusions based on inter- (or even intra-) region or sector analysis. The same situation pertains to the other project characteristics (funding level, project duration, etc.) with which this report deals. The analysis contained in this report regarding relationships between the factors, characteristics and performance ^{is} therefore, presented with considerable reservation and caution. Indeed, the analysis itself, and

conclusions drawn from it should be viewed primarily as a suggested analytical framework or approach and an indication of the kinds of information and relationships which might usefully be examined if a "sufficiently" large base of evaluated projects were available. To a lesser extent conclusions might be seen as "clues" which should perhaps be pursued through more extensive ratings and more extensive and thorough analysis. I would estimate a minimally "sufficient" number of projects for a good ratings analysis at, say, 100 a year, with few or no "characteristics" having less than, say, 15-20 samples. Below I suggest a method by which such a relatively large number of projects might be evaluated numerically at relatively low-cost and with good effect in terms of accuracy and justified potential-user confidence.

- I can not at this time comfortably or confidently reach any detailed conclusions regarding possible project or program planning implications based on the scores and rankings outlined in the Analysis section above. More analysis, first of all, would have to be done from both a project characteristics and an accomplishment factors perspective in order to understand these data better. But beyond this, I am not now convinced that these kinds of analysis and conclusions to be derived from them have utility in terms of future project and program planning. They might (no matter how large the project sample and database) be too general to have specific relevance or applicability. That is to say, that even if these data indicate, for example, that in the past, education projects did

very well, particularly "moderately-priced" (say \$1.0 to 10.0 million) education projects lasting six to 10 years and having low beneficiary involvement, we would not know why those projects did very well; nor would our ability to forecast how future education projects would do in any particular (unique, really) LDC environment ^{be improved.} Only country specific and current analysis are likely to be (can be) helpful in this regard. In brief, it seems likely that the kind of analysis and conclusions aimed at discovering or suggesting causal relations between project characteristics and accomplishments might be inherently either so conventional and broad, i.e., already widely understood and appreciated by AID as to be of little specific operational utility or so contentious and broad as to have proponents on both (or all) sides of the issue under consideration and thereby to be, again, of little operational use.

All of this is not to suggest that this kind of analysis and conclusions drawn from it can not be intellectually stimulating and useful to academics, say, and to AID staff, in expanding their understanding and sensitivities. Because of this, I would not, at this time, foreclose the possibility of further analysis of these kinds of data and submission of that analysis to AID colleagues for review. Helpful guidance for the future might, with that kind of input, be realizable.

- As opposed to my concern or reservation, as indicated above, I believe that the kind of data on which this report is

based can be useful in serving the main purpose of this report, i.e., providing AID management and others a "documented indication" of how well we've done in the past with regards to some important accomplishment objectives. If my positive view on this matter is accepted, I would recommend an alternative means for providing, and maintaining on a current basis, this kind of "documented indication." In brief, I would recommend employment of scientific sampling and polling techniques on an extensive basis (say 100 projects a year), using the kinds of project characteristics, accomplishment factors, and rating scales used in this report (all to be carefully reviewed and modified as appropriate).

Utility of IE Reports

- I think it important that a number of field-oriented AID-staff are dubious about the usefulness of the IE reports to AID in terms of the main objectives of the IE program: to provide substantive and easily "digested" and communicated performance information/judgements to AID management; and to provide information/guidance to AID staff in project planning and management.

It is also interesting to me that there is some (though largely muted) uncertainty or concern in AID regarding the appropriateness and effect (benefits and costs) of the rather wide distribution which the IE reports receive. As I understand it, copies of the reports are distributed

("unsantized") to universities, other aid donors, and the Congress and to other agencies of the Executive Branch. A study of the "impact" of the reports on the knowledge, understanding and actions of these entities regarding AID and its "mission" and efforts and economic development matters generally might be instructive.

- In this regard, the view seems not uncommon among AID staffers that the IE reports have been very useful as instruction or training tools ("case studies"), in university programs in such areas as development economics, geographic area studies and other social science disciplines. There is also a view or perception (though less widely held) that the IE reports have been useful to other donors in helping them "think-through" and further develop their own evaluation programs.

- Beyond this, I sense that the view is not uncommon within AID that the reports and the process by which they are planned and implemented have been successful in terms of enhancing intra-AID consideration of important development issues and thereby "up-grading," subtly and over time, the knowledge, understanding, and development planning and management capabilities of AID staff. On the other hand, there seems to be, almost literally, no support for the idea that the IE reports have contributed noticeably or traceably to the generally acknowledged improvement in the quality of AID project planning documents which has taken place during the

past several years. There is a virtually unanimous view among the persons with whom I've discussed these matters to the effect that the reports have not been useful to (or even seriously considered by) Agency management.

- The above suggests to me that we do not now have a sufficiently clear understanding of what needs the IE reports (or program) addresses or, conversely what objectives the program pursues; who its end-users are (or should be); and what degree of success it has had to date. Further examination of these matters, which would include obtaining the views of AID management, seems warranted to me.

VI RECOMMENDATIONS

- Shortly after distribution of this report, discussions should be held with in AID to obtain views as to the report's potential utility, if any. Assuming these discussions reveal a sense of potential usefulness, follow-up discussions should be held, say six to 12 months later to obtain views regarding realized utility, if any.

- Assuming the major objective of this report is deemed sufficiently important to pursue and that the numerical analysis approach used in this report is considered a feasible one in pursuing that objective, I recommend that AID consider discussing with key Congressional Committees (presumably, Appropriations and Foreign Affairs/Relations) and possibly the

NSC and OMB their participation in an AID performance evaluation program. Their participation would consist of their provision of staff needed to work with AID staff in implementing the field portion (the major part) of that program. (If necessary AID could/would conduct the program with or without that participation). The main element of this program calls for the fielding of teams of evaluators comprised of three to four persons drawn from eight major sources: US Congressional staff, non-technical staff of other Executive Branch agencies, US academic staff, US private business personnel, cooperating LDC academics, cooperating LDC government staff, cooperating LDC private business community, staff and AID. Perhaps four or five teams drawn from these sources would be fielded each year, each team for about three months - one month of study in the US and one month of fieldwork followed by one month of debriefings and report preparation in the US.

The month of study in the US would be used primarily to become conversant with the country in which they will work (selected by PPC/E); to select, say, between 10 and 20 completed AID-assisted projects for numerical rating from a list containing a larger number of projects prepared by AID (and with a view to contributing to the annual aggregate rating or evaluation of say 100 projects that constitute a representative sample of completed AID projects in terms of basic project defining characteristics, e.g., sector and funding level); to prepare a list of perhaps two to 10 persons

per project to be interviewed in the cooperating country, each person regarding one or more of the selected projects (and with PPC/E and the cooperating USAID primarily responsible for developing the "master" country list of potential interviewees from among host government, private sector, academic and USAID candidates as well as actual or potential beneficiaries); to become thoroughly familiar with the sets of project characteristics and accomplishment factors and the rating scales by which the factors will be "measured"; and to become familiar with preferred interviewing and information recording procedures.

The month in the field would be used to conduct the interviews and prepare the project ratings. (The above suggested level of effort might be excessive to need. Substantial adjustments down might be feasible in the numbers of interviewees and, possibly, projects without compromising the soundness of the approach and the benefits expected from it).

The final month in the US would be used to prepare the rating material for computerization and for debriefings/discussions with a wide range of Congressional, other Executive Branch and AID persons and to discuss the contents of a summary report.

Responsibility for preparing the analysis of the team's ratings and a brief report centered on it would rest with the AID member of the team and the PPC/E Coordinator for the sector. Responsibility for gaining some substantive use of the report would also be a prime responsibility of the PPC/E Coordinator.

Considerable refinement of the above outlined approach will be needed if it is to be tried or tested as a systematic and potentially long term continuing effort.

- Consideration should be given to undertaking studies of some of the matters indicated in earlier parts of this report. These might include:

-- development of a plan for a series of evaluations of "New Directions" (post - 1973) projects with a view ~~to~~ *to* comparing their performance with that of earlier projects.

-- a study, closely related to the one immediately above, of the rationale, objectives and understanding (within AID and selected host countries) of the basic pre- and post '73 aid concepts or strategies and AID's "compliance" with or adherence to them.

-- a study of the relationship of "beneficiary participation" in various kinds of projects and the performance of those projects.

-- a study of the relationship between project implementator in selected projects and the performance of those projects. (This kind of study would clearly have to be done by an outside (non-AID) contractor and might even be done for AID by the GAO, or possibly (if remotely so) by a group involving DAC members).

- A study to analyze the relationships between implementation problems encountered in already-completed AID-assisted projects and the accomplishment results achieved by those projects.

- Assuming that following a critical assessment of its usefulness, the IE program is continued in much its present form, consideration should be given to having the IE teams use the numerical rating approach to the project(s) being evaluated as a supplement to their present approach and product, i.e., the IE report.