

PN-AAP-698/62

ISN-34156

USAID MISSION CAPACITY FOR DATA RELATED ACTIVITIES:  
CURRENT STATUS AND OPTIONS FOR IMPROVEMENT

by

Dr. Chris Hermann

153-5

Prepared for AID/PPC/PDPR/RD  
under Purchase Order Number  
OTR-0091-0-00-2314-00

November 1983

USAID MISSION CAPACITY FOR DATA RELATED ACTIVITIES:  
CURRENT STATUS AND OPTIONS FOR IMPROVEMENT

Executive Summary

This report concerns the capability of USAID missions to:

1. use quantitative data for program and project purposes,
2. support the data collection and analysis components of AID projects, and
3. develop the institutional capacity of host countries for data related activities - data collection, analysis and use of findings for planning and policy purposes.

The report presents the findings of a study of the data related activities of six selected missions: USAID/Nepal, USAID/Egypt, USAID/Honduras, USAID/Panama, USAID/Mali and USAID/Zimbabwe. Various AID program and project documents combined with more than one hundred interviews with mission and AID/Washington staff, contractors and host country staff provided the basic information used for this report.

The principal conclusions and recommendations concerning improved support for data related activities are as follows:

1. The Agency needs to address three key institutional constraints to improve and expand data use in AID operations.
  - Clarify and strengthen the Agency's commitment to meeting its internal information requirements with the appropriate data. Such a commitment is essential for AID to strike a better balance between acting on a timely basis and acting on an informed basis. Greater emphasis on matching information needs to the appropriate data will encourage missions to better identify and obtain the various types of data needed (e.g., economic, agricultural, nutritional, etc.) at the necessary level of disaggregation (e.g., national, regional, provincial, village).

- Strengthen and better target the Agency's efforts to develop the institutional capacity of LDCs for data related activities. AID should capitalize on the comparative advantage the U.S. has in many areas of data collection and analysis. The Agency should take a more active role in transferring appropriate technologies to host countries to develop the capability of these countries to collect and analyze data for their own information requirements. In many LDCs, however, the need for accurate and timely data far outstrips the country's human and financial resources devoted to data collection and analysis. Therefore, the Agency should concentrate its efforts on transferring only affordable and sustainable systems and techniques to host countries.
  
- Eliminate existing disincentives toward better or greater data use by AID staff. At the same time, capitalize on staff interest and capability for improving data use in mission operations. At present, staff involvement with data related activities is frequently treated by senior management as though this were outside the interests or objectives of AID, regardless of the potential benefit to mission operations that better data bases and better data use offer. AID should encourage mission staff with quantitative skills to improve mission data use and support for data collection and analysis by providing worktime and appropriate incentives (i.e., career advancement). Analytically skilled staff could function as an in-house source of technical assistance for the mission's data related activities if such tasks were included in their job descriptions. Similarly, AID should capitalize on the existing interest and capability of mission staff for improving data use within mission activities.

In short, improvements entailing relatively small investments of time and funds could be made by many USAID missions if AID dealt effectively with the above institutional constraints.

2. USAID mission capacity for data related activities is a function of both in-house and host country capabilities. The principal in-house factors are
  - program size
  - senior management support
  - staff skills
  - availability of technical assistance
  - mission access to automated data processing equipment
  - availability of data.

Host country factors include

- an established tradition of using data for decision-making
- current interest in improving data collection and analysis
- human and financial resources
- availability of automated data processing equipment
- political context
- geography, climate and infrastructure.

The combination of these twelve factors establishes whether a mission has a low, medium or high capacity for data related activities. Therefore, mission capacity can be improved by changing any of these factors in a positive direction (e.g., greater senior management support for data related activities). Though some improvements could be made independent of AID/Washington, there are very definite limits on what can realistically be expected of even high capability missions without additional support. Therefore, AID/Washington should develop data support services to assist the missions with data related activities.

However, the only long-term solution to inadequate information in LDCs is developing host country capability for data collection and analysis. The potential for making the comparatively greatest improvements exists where the USAID mission has a medium to high capability for data related activities and the host country has a low to medium capability. The Agency should increase its efforts to expand host country capabilities where these conditions prevail.

3. The Agency should take a more pragmatic approach to data collection and analysis in LDCs. A pragmatic approach entails compromising statistical standards of data quality to accommodate data collection and analysis to the conditions under which AID operates. How far statistical standards must be compromised is dependent on the specific information needs, resources and other pertinent considerations of the project. A pragmatic approach, however, is not an excuse for half-hearted efforts on the part of AID for data collection and analysis. Rather, AID should be firmly committed to obtaining data of the highest quality that resources and conditions allow. A pragmatic approach should strongly influence decisions pertaining to data sources to be used, data quality, the representativeness of the data, the amount of data to be collected, and the frequency of data collection.

4. To obtain systematic, empirical evidence of project effectiveness, a number of projects reviewed for this study plan baseline - follow-up data collection. However, serious problems concerning the baseline - follow-up approach will substantially reduce or, in some cases, eliminate the utility of these data for project evaluation purposes. These problems can be attributed to inadequate planning, flawed designs, limited host country capability, USAID management practices, the extensive and complex information requirements of integrated rural development projects, and the time required for project effects to become measurable. To improve the utility of baseline - follow-up data collection for project evaluation purposes, AID should restrict its use of the design to projects where it is both necessary and feasible to collect data at more than one point in time and analyze the data collected on a timely basis.

As an alternative or adjunct to AID's current heavy reliance on data collection on a project-by-project basis, the Agency should give greater consideration to sector level evaluations of program success in achieving sector-wide objectives. Sector level evaluations are already planned by some missions using the non-projectized assistance mode, such as USAID/Zimbabwe. In general, a mission using this approach would have to select key criteria which accurately measure the types of changes that should result from the projects the mission supports in a given sector. Data for such evaluations would typically be broader and more comprehensive than that ordinarily collected to evaluate specific project outputs and effects.

5. The following actions are also recommended:
- establish a Data Support Division within PPC or S&T to assist the missions with data collection and analysis;
  - establish Regional Support Centers in selected missions that would then assist all missions in the region with data related activities;

PPC should take the following actions:

- develop model analyses of routine tasks for programs offices in a format amenable to microcomputer systems,

- provide guidance concerning subnational, socio-economic analyses missions are expected to perform,
- require an information strategy statement in the ABS,
- include in Handbook 3 guidance concerning the planning and implementation of data collection and analysis in AID projects, and
- develop an information policy statement clarifying the Agency's commitment to matching information needs to the appropriate data and the importance of developing host country capability for data related activities through technology transfer appropriate for the host country's needs and resources;

USAID missions should take the following actions:

- support separate institution-building projects to improve data collection and analysis in sectors where the mission's program is concentrated,
- make as much use as possible of local private sector firms to develop an in-country capacity for data collection and analysis,
- give greater emphasis to data collection for project design,
- use staff skills to create an in-house source of technical assistance for data related activities,
- try to coordinate data collection and analysis among projects and offices, and
- exploit any opportunity that arises for coordinating data related activities between the mission and other international development agencies.

Overall, there is much the Agency can do to improve mission support for data related activities. Highest priority should be given to building host country capability for data collection and analysis. More effective institution-building at that level will, however, require strengthening mission capacity for data related activities as well.

## Acronyms

- AID - Throughout this report, AID is used to refer to the entire Agency for International Development including both AID/Washington and USAID missions. AID is used interchangeably with the Agency.
- USAID - USAID is used to refer exclusively to the Agency's field missions and to distinguish the missions from AID/Washington.
- PPC - The Bureau of Program and Policy Coordination
- IRM/MPS - The Division of Mission and Program Support in the Office of Information Resource Management.
- BuGen/ SEU- The Survey and Evaluation Unit of the Bureau of the Census.
- BEST - USAID/Zimbabwe's Education Sector Program - Basic Education and Skills Training.
- PID - Project Identification Document.
- PP - Project Paper.
- PAAD - Program Assistance Approval Document.

## Table of Contents

1. Introduction	1-7
1.1 Purpose of the Report	1
1.2 Information Used for the Report	4
2. Institutional Constraints Affecting USAID Mission Support for Data Collection and Analysis	7-24
2.1 The Agency's Commitment to Meeting Its Information Needs to the Appropriate Data is Ambiguous.	8
2.2 AID's Commitment to Building the Institutional Capacity of LDC's for Data Related Activities Needs to Be Strengthened.	12
2.3 The Agency Should Eliminate Disincentives Toward Better Data Use and Capitalize on Staff Interest in Improving Data Use in Mission Activities.	18
3. Factors Affecting USAID Mission Capacity for Data Related Activities: In-house and Host Country Capabilities	24-64
3.1 Mission Capability	26
3.2 Host Country Capability	37
3.3 Summary of USAID Mission Capacity for Data Related Activities	47
3.4 Mission Strategies for Improving Host Country Capability for Data Related Activities	53
3.5 Conclusions	63
4. A Pragmatic Approach to Data Sources, Quality, Representativeness, Amount and Frequency of Data Collection	65-85
4.1 Data Sources	68
4.1.1 Sample Surveys	68
4.1.2 Operations Data	74
4.2 Data Quality	77
4.3 Representativeness	79
4.4 Amount of Data	82
4.5 Frequency of Data Collection	84
5. Project Versus Sector Level Data Collection and Analysis	85-108
5.1 The Baseline - Follow-up Design: Good Intentions Gone Awry	85
5.1.1 Planning	87
5.1.2 Design	89
5.1.3 Host Country Capability	90
5.1.4 USAID Management	91
5.1.5 Integrated Rural Development Projects	93
5.1.6 Short-term Effects Versus Long-term Impacts	94

5.2	Improving AID's Use of Baseline - Follow-up Designs	97
5.3	Sector Level Analyses and Evaluations as an Alternative or Adjunct to AID's Project by Project Approach	99
6.	Recommendations for Improving AID's Current Data Related Activities	108-126
6.1	AID/Washington	110
6.1.1	A Data Support Division	110
6.1.2	Regional Support Centers	112
6.1.3	PPC	114
6.2	USAID Missions	115
6.2.1	Stand-alone Institution-Building Projects	115
6.2.2	Local Contracting Firms	119
6.2.3	Data for Project Design	121
6.2.4	In-house Support for Data Related Activities	122
6.2.5	Coordination of Data Related Activities in USAID missions	123
6.2.6	Coordination Between USAID Missions and Other International Development Agencies	124

VIII

## 1. Introduction

### 1.1 Purpose of the Report

This report concerns the current capability of USAID missions to support data collection and analysis activities designed to 1) meet program and project information needs and 2) to develop the institutional capacities of host countries to use quantitative data for administrative and planning purposes. AID/Washington's policies and procedures for collecting and managing quantitative, socio-economic data were reviewed in late 1981. A major recommendation of that review was that the Agency clarify its position on the importance it attributes to data related activities in a formal policy statement. To be effective, the Agency's information policy would have to recognize the range of USAID mission capabilities for data related activities. It was necessary, therefore, to review data collection and analysis activities currently supported by representative USAID missions to determine what the key factors are which affect the missions' capabilities for such work. This report presents the findings of that study. Its primary purposes are 1) to suggest how the Agency could improve the effectiveness and utility of data related activities USAID missions support and 2) to provide guidance for developing an information policy determination for the Agency.

The function of an information policy determination should be to provide the basis for the central and regional bureaux to

take the actions necessary to improve support to the missions for data related activities. An important objective of this report, therefore, is to suggest where improvements need to be made in light of the key factors and issues affecting current mission involvement with data collection and analysis. The need for improvement and, in particular, for better support systems for data related activities is clear from the information obtained from the missions selected for this study. The problem can be stated in terms of a simple ratio: the number of USAID activities where adequate use of data and analysis was made (e.g., where data collection and analysis genuinely contributed to project success) compared to the number of activities where data use or support for data collection and analysis was inadequate (e.g., where faulty or insufficient data contributed to poor project design, monitoring and/or evaluation). This study has found numerous examples of both adequate and inadequate data use for USAID program and project purposes, yet it is fair to say that there is considerable room for improvement to shift the balance of that ratio more to the positive side and less to the negative. Stated simply, the Agency should practice what it preaches - if improved data collection and analysis is beneficial to the development planning and government operations of the host country, then the same logic certainly applies to AID's own planning and operations.

A major conclusion of this study is that inadequate support for data related activities is the central problem the Agency needs to address. Better support for data collection and analysis to meet AID's own information requirements as well as those of the host country is certainly warranted in light of the increasing demand for better information about Agency operations. Pressure for demonstrating the effectiveness of AID's programs has, if anything, increased for those in government who oversee foreign assistance. The Agency's own internal demand for information within AID/Washington, within USAID missions and between the field and the central and regional bureaux is also substantial. Moreover, under the constraints of shrinking budgets and decreasing staff size, better data use in Agency activities will be essential to do more with less. Significant demands for information requiring data collection and analysis already exist in policy papers, CDSS guidance, the Project Assistance Handbook and project monitoring and evaluation requirements. Yet it is highly questionable whether the capacity to comply with these demands also exists in the missions. Perhaps most disconcerting, AID staff themselves link inadequate data use for program and project planning to the implementation problems arising from poor project design. In short, the main argument of this report is that basic data support systems and services should be established now if future improvements in meeting AID's informa-

tion needs are to be achieved.

### 1.2 Information Sources Used for this Report

USAID missions vary widely in terms of their capabilities for data related activities. The staff size of missions range from a few individuals to more than one hundred. Funding levels are as low as several million dollars and as high as \$750 million. Similarly, some LDC's have only marginal capacity for data collection and analysis while others are quite advanced. To cover the range of variation in mission capability, a set of missions was selected which differed substantially in regard to staff size, funding, host country capacity and geographic location. The missions selected were USAID/Nepal, USAID/Egypt, USAID/Honduras, USAID/Panama, USAID/Mali and USAID/Zimbabwe. The primary source of information for this study was loosely structured interviews with mission staff and contractors concerning the data related components of on-going projects and other mission activities. (See Annex A for a copy of the questions used to guide the interviews.) From these six missions, approximately one hundred USAID staffers and ten contractors were interviewed. During the course of the study, some sixty projects and sector programs were reviewed. CDSS's (including Annexes and Updates), PID's, PP's, evaluation reports, sector assessments and other pertinent materials provided background information. Fifteen AID/Washington staff knowledgeable about the programs of the missions

selected were also interviewed. USAID/Zimbabwe was an exception to this procedure. The mission has a staff of eight, only four of whom (including the mission director and the recently appointed deputy director) are directly involved with the substantive (as opposed to the support) requirements of the mission's program. The human resource officer for USAID/Zimbabwe was interviewed in Washington. Because of the nature of USAID/Zimbabwe's program - primarily sector funding and a large CIP - and its small staff size, it was decided that sufficient information could be obtained from available documents and AID/Washington staff and contractors who had recently worked in the mission. The current status of data collection and analysis underway in each mission was described in a series of separate reports noting the factors which appeared to affect the progress of those activities. This report summarizes and draws from the information obtained from each of the six missions.

The accuracy of this report depends on whether the experience of the selected missions concerning data related activities is representative of other USAID missions. The representativeness of this set of missions cannot be ascertained precisely. But two points are worth considering: 1) whether the factors affecting data use in these six missions also operate in other USAID missions and 2) whether each of the six missions accurately represents a larger group of similar missions.

There is good reason to believe that the missions used for this study do illustrate the common factors affecting the capability of USAID missions in general for data related activities. The six missions selected are a fairly diverse group, yet recurrent issues and problems were found which affected the data collection and analysis supported by each mission. Moreover, successful data collection and analysis efforts shared certain common characteristics (e.g., simplicity of design) as did unsuccessful efforts (e.g., techniques inappropriate for host country capabilities).

Whether each mission is representative of some larger group of USAID missions is more difficult to determine. For example, USAID/Mali is probably similar to other Sahel missions in regard to the availability of data and host country capability. Whether the mission is representative of other low capability missions in Africa or elsewhere remains to be seen. Similarly, USAID/Egypt is unlike most other missions in that the size and diversity of its program will generate more data and analysis than perhaps all other missions. It is probably the case that every mission has some unique or special characteristics which distinguish it from others. On the other hand, even though missions are not identical, they can have comparable capabilities for data related activities. Part Three of this report will point out that mission capability is a function of internal, in-house factors and external, host

country factors. From this perspective, different missions can have similar capabilities for data use and support to the host country but for different reasons. In this regard, the selected missions are probably indicative of the general categories - i.e., high, medium and low - of USAID mission capability for data related activities.

## 2. Institutional Constraints Affecting USAID Mission Support for Data Collection and Analysis

A preceding review of AID's policies and procedures for managing quantitative, socio-economic data was conducted in late 1981. The problems identified in that report as impeding better use of quantitative data were categorized as institutional, organizational or operational in nature. It was argued that the institutional constraints were most important because they were central to improving the Agency's data-related activities. The institutional constraints were described as follows:

- 1) The Agency's commitment to improved and expanded use of quantitative data is ambiguous. While AID expects analytically sounder work from the missions, it fails to provide the resources necessary to do so. Action is needed to better assure that AID's diverse information requirements are met by the appropriate data.
- 2) The Agency's commitment to promoting self-sufficiency for data related activities in LDC's through institution-building should be

strengthened. AID needs to clarify and reinforce its position that the capacity of LDC's to use and manage their own data is a legitimate target for institution-building of importance equal to that of other objectives.

3) The allocation of funds, staff, worktime and the reward system of career advancement constitute disincentives toward improved use of quantitative data. To meet the demand for analytically sounder work, AID needs to overcome these impediments.

Because of the institutional nature of these constraints - i.e., they are systemic and influence data related activities at all levels within the Agency - it seemed very likely that they would also affect USAID mission operations. It is no surprise, therefore, to find clear evidence of a mission-level equivalent for each of the institutional constraints. This section briefly discusses the effects of these constraints on the data related activities of the USAID missions.

### 2.1 The Agency's Commitment to Meeting Its Information Needs With the Appropriate Data is Ambiguous

The most fundamental and pervasive problem USAID missions presently confront regarding data use is meeting their need for various types of information with the appropriate data. The severity of the problem varies from mission to mission. USAID/Honduras, for example, represents missions which have been fairly successful in

collecting and using data for program and project purposes. USAID/Mali, on the other hand, represents the opposite extreme. But even in the more successful missions, it appears that matching information needs to the appropriate data is problematic and, in some instances, does not happen.

A principal cause of this problem is AID's ambiguous commitment to improved and expanded data use. On the one hand, numerous demands are placed on the missions to meet acceptable standards of information use. These demands are explicitly stated in CDSS guidance, the Project Assistance Handbook and project evaluation requirements. There are also substantial demands for expanded data collection and analysis implicit in the Agency's recent Policy Papers. On the other hand, the Agency has taken practically no action to guarantee that the missions will be able to comply with those demands. Some of the most basic requirements for data related activities - e.g., mission staff who have quantitative skills and whose workload includes oversight and assistance for data collection and analysis - has not been provided. Nor is there any one office or division in AID/Washington charged with the responsibility of providing technical assistance for data collection and analysis or for backstopping in any number of ways the data related activities missions undertake. Recently a very significant re-organization and re-direction of SER/DM has been accomplished through its replacement by SER/IRM (Information Resource Management). The Agency will

now be far more responsive to meeting the computer hardware and software needs of the missions. In particular, the Mission and Program Support Division (MPS) of SER/IRM has developed an excellent system to assist missions to acquire and maintain microcomputers. Similarly, the DP offices have supported more extensive use of cost-benefit analysis for project design. Yet little if any action has been taken by other parts of the Agency to support the implementation of data collection and analysis components of projects or other data related activities of the missions.

Mission staff are most directly affected by the present ambiguous situation. They are well aware of the problem and their position in it. In each of the missions selected for this study, staff described how for one project or another, they had no alternative other than to work with whatever data were at hand regardless of its quality, timeliness or whether it was disaggregated to the appropriate level. Worse yet are their reports of simply muddling through as best as their expertise allowed essentially working without adequate information. The specific reasons for this have been cited in preceding reports: the unavailability of data from the host country, budget and time constraints precluding minimum data collection and analysis, mission staff without the necessary skills for data collection and analysis management, no technical support from AID/Washington for data related activities, no worktime allocated to data related activities, lack of support from senior management, and so on. The effects of the problem are most apparent at the project level, beginning with project identification and continuing

through to the final evaluation. Some of the most telling observations pertaining to data use in USAID missions were made by some of the Agency's most experienced staff:

- the deputy mission director who stated mission staff were so completely consumed by implementation problems that there is simply no time for more careful use of data to guide planning;
- the assistant office director who remarked how dismaying it is to witness how projects come to be identified and designed without basic data. As evidence, this person pointed to the number of projects which encounter implementation problems at the very outset and continue to experience such problems throughout the course of the project, consuming an inordinate amount of staff time; and
- the program officer who frankly stated in a cable to AID/Washington that a recent review of the implementation status of one project was "...a litany of problems common to complex integrated rural development projects built on incomplete and faulty information."

Unfortunately, the conclusion that must be drawn from this and other supporting evidence is that the situation will not improve until the Agency takes action necessary to assure that the missions will have better access to the appropriate data to meet their information needs.

2.2 AID's Commitment to Building the Institutional Capacity of LDC's for Data Related Activities Needs to be Strengthened.

Of the three major constraints to improving the effectiveness to AID's data related activities, strengthening the Agency's commitment to developing a sustainable capacity within the host government for data collection and analysis which is commensurate with their needs and resources is most consistent with AID's development goals. The Agency needs to make a more concerted effort in this area because accomplishing this objective in many countries proves far more difficult than it might first appear.

A very difficult and, regrettably, common problem USAID missions confront in their institution-building efforts for better data use is a lack of genuine interest on the part of the host country. Disinterest in improved data collection and analysis by the host country typically reflects the absence of a tradition of planning and decision-making based on empirical information. Instead, overriding political concerns, particularly maintaining political stability, dominate government policy and planning. The host country will agree to projects which ostensibly will develop ministry capabilities to use data more effectively. However, their real interest is apparently in the physical outputs of the projects (e.g., new buildings, roads, vehicles, etc.). The statistical institution-building component of the project is perceived as something foisted

upon them as part of the process and, in some instances, to be treated as inconsequential.

In this situation, if the data related component is to succeed, the USAID mission is essentially forced to try to modify the behavior of the host government in regard to planning procedures. Institutionalizing data collection and analysis becomes a matter of changing the decision-making process so that data use is an integral component. The common expression bandied about in these circumstances is "creating a felt need" on the part of the host country for adequate information (and adequate by the standards of Western planners and public administrators). At the same time, the mission's own project information needs are tied to the actions of the host country. Consequently, the predictable inaction on the part of the host country means that USAID information needs go unmet.

Two health projects in Egypt which have substantial data collection and analysis components and are to institutionalize improved data use in the Ministry of Health are perfect examples of the problem. In one project, after years of promises by the Egyptian contractor responsible for the statistical institution-building component, no genuine progress had been made toward improving data use in the ministry. There was an almost game-like quality in this project. The MOH was not really interested in the data component. One mission staffer involved with the design of the project said that disinterest

was understood at the outset and no one (?) really expected institutionalization of data use to occur. Yet plans to improve data use in the ministry figured prominently in the project paper and an evaluation team working on the project acted as though they firmly believed these improvements were to have been made. These various elements are simply inconsistent and contradictory. In the other health project, first round analyses which were to have guided later phases of the project were literally years behind schedule. An evaluation of the project emphasized the importance AID attributes to data analysis, yet funding was extended for several more years. The messages here are again inconsistent - data use is important; however, even though it has been neglected, funding will be continued. The question here is whether project funding should be suspended because the host country has failed to make a serious attempt at the data component of the project. Cutting off funds seems a rather extreme action (especially in USAID/Egypt where such tremendous pressure to "keep the money moving" exists). But how should the Agency demonstrate the importance it ostensibly attributes to data related activities when those activities are ignored by the host country? That question returns us to the first institutional constraint of what precisely is AID's commitment to improved data use.

Developing the institutional capacity of LDC's for data related activities is obviously directed towards helping the country meet its own information needs. Determining the amount and quality of data needed in relation to the country's ability to afford the costs of collection, analysis and management of data can be a complex question. One common view is that <sup>as</sup> a country develops, it increasingly needs data of higher quality, greater comprehensiveness and more detail. There is much to support this view. With development, the major and most obvious constraints are overcome; impediments to further growth might be more intricate and more difficult to fully understand. The national economic system also becomes more complex and the need to monitor economic conditions on a routine basis increases. Finally, as the rate of development quickens, data must be collected and analyzed more frequently and on a more timely basis. It is assumed in this perspective that the capacity to afford data related activities grows at a commensurate rate and that the greater importance of timely accurate data is appreciated by the developing country.

One implication of this perspective is that the least developed countries have the least need for data. In some instances, this is probably true. For example, some problems least developed countries face are so apparent that large scale data collection efforts are unnecessary and possibly counterproductive. But identifying major

problems and constraints is not the same as devising appropriate strategies and designing effective projects. For those purposes, systematically collected data of reasonably good quality is needed regardless of the level of host country development. What the country lacks is not the need for data, but the ability to afford it. Mali exemplifies precisely this situation. Very poor countries like Mali are in an economically precarious position and cannot afford to make unsound decisions which waste extremely limited resources and impede development. In short the need for data to guide decision-making in countries like Mali is not commensurate with their level of development. Rather, their needs far exceed the human and financial resources they have available for data collection and analysis.

The fundamental inability of many countries to bear the re-  
of  
current costs/improved or expanded data collection and analysis must figure prominently in the Agency's policy concerning institution building for data related activities. AID's strategy must differentiate between opportunities for sustainable institution-building within the budget of the host country and other instances where alternative actions are needed to meet mission and host country information needs.

Where legitimate opportunities for institution-building are found, the Agency must further consider whether the skills, techniques and systems - i.e., the technology being transferred to the host

country - are appropriate just as it does for other types of technology transfer. The six missions selected for this study indicate that USAID missions are very much involved with efforts to develop the host country's capacity for improved administration and planning through better data collection and analysis. Yet some mission staff seriously questioned whether the techniques and systems being developed were appropriate given the limited staff skill level and interest of ministry officials. Some believed that the institution-building component was providing technology which was too sophisticated to be sustained after the project by the host country.

A "level of development" perspective seems to be the best guide to resolving this problem. That is, greater statistical sophistication becomes increasingly appropriate as the country develops. But here again the matter is not that simple. The Agricultural Resource Inventory Project in Nepal illustrates the point. The project will establish a monitoring capability in the Government of Nepal based on remote sensing data. This is a rather "high tech" solution for such a "low tech" country. But in this instance, the solution is most appropriate. Nepal must be able to locate and track its most serious environmental problems. It also desperately needs to monitor cropping patterns and predict production levels. Given the physical geography of the country, remote sensing would appear to be the only viable solution.

The conclusion that ought to be drawn from this is that AID's support for institution-building which involves improving data use for administrative and planning purposes should be guided by sophisticated (in the sense of enlightened) strategies. Uniform, undifferentiated approaches which in effect ignore the contingencies of the situation, particularly host country needs, interests, resources and capabilities, will doubtlessly lead to attempted improvements which are not sustainable because they are, in one way or another, inappropriate for the host country. Rather, AID needs to clarify and make consistent its commitment to institution building for improved data use in LDC's in such a way that USAID missions can choose among alternative solutions to meet development information requirements.

2.3 The Agency Should Eliminate Disincentives Toward Better Data Use and Capitalize on Staff Interest in Improving Data Use in Mission Activities.

A major impediment to improving data related activities undertaken or supported by USAID missions is the overriding importance placed on obligating annual budgets within the fiscal year. Disincentives toward increased involvement with data related activities by mission staff appear to be the logical consequence of the pressure to move money on a rigid schedule. These disincentives include:

- 1) lack of support or interest on the part of senior management

in improving data use; 2) insufficient allocation of worktime for data related activities by mission staff; and 3) a reward system which focuses on managerial performance at the expense of improving data bases and analytic work beneficial to mission operations. Mission management responds directly to the policies and directives issued by the Agency including the message to obligate funds on time. Greater involvement with data collection and analysis is easily perceived as jeopardizing mission compliance with moving money. Lack of senior management support is reflected by inadequate work-time for skilled staff to engage in data related activities. Moreover, mission staff certainly recognize that career advancement is heavily dependent on managerial criteria which do not reward significant involvement with improving data use within the mission. In short, the system is oriented toward expediency of action, not careful analysis for planning action.

Even given the valid reasons that can be mustered in defense of moving money according to the fiscal calendar, the Agency still must find room to incorporate better information use in the process. Simply getting the money out the door on time is no guarantee that it has the desired development impact. AID will have to strike a better balance between acting on a timely basis and acting on an informed basis.

The reverse side of the coin is for AID to capitalize on the positive factors which could facilitate better information systems in mission operations. A number of mission staff have quantitative, analytic training and experience which are directly applicable to meeting AID information requirements. Equally important, many with such skills firmly believe that data collection and analysis are not only relevant to but necessary for improving the Agency's operations. In addition to the information gathered for this study, the following responses of several USAID missions to a cable concerning the performance of BuCen/SEU and the perceived utility of data related activities are insightful:

USAID/Zimbabwe:

"In general mission believes AID should give greater attention to data collection and analysis particularly in its priority areas such as agriculture, population, health, and human resources. Attention should be given to assisting host countries develop these tools of development."

USAID/Bolivia:

"...the current managers and staff concur in placing data collection and analysis high on the list of priorities for the mission and for development projects of the country...it is of an utmost importance that AID

(and other donors) actively promote programs of this nature. In their absence neither the mission nor the host country have available the needed basic information on which can be based sound policy decisions and development program(s)."

USAID/Indonesia:

"... AID has, in the past, given too little emphasis on surveys and evaluation but awareness of the value of effective evaluation of projects is growing."

USAID/Barbados:

"Too little emphasis given to surveys and evaluation activities to provide data for effective policies and programs."

The project officer of the Regional Education Analysis in USAID/Kenya cabled the following response:

"I believe such projects constitute one of the most effective kinds of policy dialogue called for by the Administrator. They transfer technology and help build key institutions. If successful, they help lay the basis for more coherent and effective sector policies. This approach constitutes an alternative to the piecemeal project by project approach which places excessive financial and administrative burdens on the LDC."

How representative such views are of all USAID mission staff is uncertain, but they are perfectly consistent with the opinions expressed by many mission staff interviewed for this study.

Interest in improving data use is also evident in the increasing number of microcomputers being purchased for projects and in-house purposes by USAID missions. IRM/MPS has clearly explained the procedure missions should follow to purchase microcomputers. MPS must simply review and concur on all such acquisitions. They have evaluated several systems that they can support; however, they remain flexible about which microcomputer the mission may purchase when circumstances indicate an alternative brand (e.g., Sierra Leone has a local NCR dealer but no Apple or IBM repair centers; therefore a NCR system might be a more logical choice though it is not one of the AID supported machines.) At a recent workshop IRM sponsored in Abidjan, it was very apparent that the number of microcomputers in the missions will proliferate to the point where missions will have several for staff use.

The potential impact microcomputers could have on mission data use could be very significant if the Agency takes action to overcome other impediments. The availability of microcomputers will introduce an alternative means for improving the quality and use of quantitative data. The standard approach to improving data bases has essentially followed a problem oriented, research strategy.

The basic idea is to adapt standard research methods to difficult field conditions as best as available resources permit. The main problems are identified and data needs are specified as clearly as possible. An appropriate methodology is selected or developed. Careful attention is paid to all aspects of the data collection process to obtain data of acceptable quality. As sound as this type of approach appears, it unfortunately seems to have been something less than a resounding success. Data in many LDC's where this type of approach has been attempted still have remarkably unreliable and inaccurate data. In part the effectiveness of the traditional approach has been handicapped by computer systems which were generally inappropriate for developing countries. As a result, much of the important analysis is frequently done in the U.S. and not in the LDC. This has undercut the training - technology transfer goals of such projects.

The microcomputer, however, has the potential to correct the hardware problem affecting the traditional research approach to improving data bases. By decentralizing data use and facilitating data manipulation, those more directly involved with policy and planning see first hand the importance of better data quality. In effect, the machine plays a central role in making improvements in data possible instead of merely playing a secondary, supporting role as computers have in the past.

Decentralized computing combined with staff interest in better data use for mission operations have clear implications for AID's data related activities. Significant improvements could be achieved if AID capitalizes on the full potential of staff skills, interest and the capabilities of microcomputer technology. However, what is missing at this time is technical support to the missions to pull together these and other elements so that real improvements do occur. Therefore, the Agency should carefully consider the options available to it for providing data support services to the missions.

### 3. Factors Affecting USAID Mission Capacity for Data Related Activities:

#### In-house and Host Country Capabilities

This part of the report deals with the current capacity of USAID missions to 1) meet internal, in-house information needs and 2) support efforts to develop the institutional capacity of the host country for data related activities. The in-house factors which affect mission capability include: 1) program size; 2) the availability and accessibility of quantitative data; 3) senior management support; 4) the quantitative, analytic skills of mission staff; 5) technical assistance from U.S. and local sources; and 6) the availability of automated data processing equipment, particularly mini- and microcomputers. These factors also define the level of support the mission can provide to the host country for data collection and analysis activities. Conversely, the data related

activities of USAID missions are to a significant degree dependent on host country capability for data collection and analysis. This dependency is apparent in the availability of data for mission information needs, designing assistance to the host country for data related activities and the outcome of projects wherein the host country has primary responsibility for data collection and analysis. A second set of host country factors, therefore, also influences mission capacity for data related activities; these include: 1) a tradition of collecting, analyzing and using data in government operations; 2) a genuine interest in improving or expanding data use; 3) human and financial resources available for data related activities; the availability and adequacy of automated data processing equipment; 5) current political factors; particularly internal stability and the host government's position on access and dissemination of data; and 6) geographical and climatic conditions.

The importance of each of these twelve factors to AID's data related activities is discussed briefly in sections 3.1 and 3.2. For heuristic purposes, the missions are ranked as high, medium or low on each of these factors. Overall mission capacity is estimated by calculating an average rank for each mission. If the necessary information were available, rough estimates of this type could be made for all USAID missions. This would further clarify the Agency's

present capabilities for data related activities, the distribution of mission capacities (i.e., the percentage of missions with a high, medium or low capability) and geographic patterns (e.g., Sahel missions have a low capability). Such a system could guide data support services designed to assist the missions with data related activities.<sup>1</sup>

Ranking the capacities of all USAID missions, however, goes beyond the scope of this report. Instead only the six missions selected for this study are ranked as a means of summarizing the results of this study. In section 3.4, a typology of mission capacity identifies three general categories of mission support for data related activities by crosstabulating mission capability with host country capability. General strategies for improving host country capability for data related activities are then discussed. A major conclusion of Part Three is that AID's opportunities for making significant improvements in the data related activities of LDC's could be expanded by strengthening in-house mission capability.

### 3.1 Mission Capability

Information obtained from the six missions selected for this study indicates that the following factors are central to USAID mission capability for data related activities. In this section, the bearing each of these factors has on data collection and analysis is briefly discussed. The six missions are then ranked on each factor to estimate their relative capability for data related activities.

---

<sup>1</sup> See Annex B for a brief discussion of mission ranking for data related activities.

1) Program Size Of the six factors pertaining to mission capability, program size is the least manipulatable. It is a "given" which establishes the basic parameters within which data related activities are undertaken. Program size consists of three components. The first is mission funding. The second is the number of direct hire staff and contract personnel in the mission who are responsible for the design and implementation of the mission's program (i.e., program and project staff). Though substantive staff (as opposed to support staff - e.g., the comptroller's office) are most directly involved with data related activities at the project level, support staff are also important because they frequently are most knowledgeable about computer systems and software for data management and analysis. Third, the diversity of the program, that is, the number of projects and the number of different sectors in which the mission is active, determines which data bases the mission must use and try to improve. These three components usually increase together - as funding increases or decreases, staff size and program diversity change commensurately. There are exceptions. USAID/Zimbabwe's program, for example, is substantial - approximately \$75 million annually. But the principal funding mechanisms for the mission are sector grants and a Commodity Import Program. This allows the mission to operate with a staff of eight. The foregoing suggests that program size ought to be viewed as a composite of funding level, staff size and sector diversity.

2) Senior Management Support Senior management support for data related activities is fundamental to improving in-house data use

and providing assistance to the host country for better data collection and analysis. In particular, the importance the mission director attributes, or allows mission staff to attribute, to better data use in mission and host country operations sets the overall orientation toward data related activities. The clearest evidence of senior management support is the allocation of staff worktime for such activities. A change in management's perspective or a change in senior level staff can have a significant impact on data related activities. For example, the current director of one of the missions selected for this study gives higher priority to economic analysis than did his predecessor. As a result, mission staff<sup>are</sup> expected to be able to make substantial improvements in this area. This suggests that improvements in AID's data related activities could result from greater receptivity on the part of mission directors and other senior level staff to the utility and importance of such work for mission operations. A clear policy statement specifying the importance the Agency attributes to data collection and analysis to meet mission and LDC information requirements should facilitate senior management support. Similarly, enforcement of existing demands for sounder analytic work in mission operations would provide additional impetus.

### 3) Mission Staff with Quantitative, Analytic Skills

In all of the missions selected for this study, there was at least one and, in most cases, several mission staffers who had

training and work experience which involved data collection and quantitative analysis. Given the range of mission staff size among the six missions selected for this study ( USAID/Egypt - 125+ versus USAID/Zimbabwe - 8), it is likely that most missions have at least one person with the skills necessary to assist with the mission's data related activities. Furthermore, the re-newed emphasis the Agency is currently placing on economic analysis and the recent hiring and placement of economists in USAID missions should increase the availability of quantitative skills. As noted earlier, many of those with quantitative skills expressed interest in improving mission data use and support for data collection and analysis. USAID/Honduras' experience indicates that even one person working in this area can have a significant impact. The key to capitalizing on existing in-house expertise is providing worktime for these staffers to use their skills to the benefit of the mission. The responsibility for providing support for the mission's data related activities should also be stated in staff job descriptions. To further strengthen the analytic and data management capability of missions, staff who have rudimentary quantitative skills could be given in-service training to refine their competency in this area.

#### 4) Availability of Data

The principal sources of data for USAID missions are the host country, AID funded activities (e.g., project generated data, sector assessments, special studies) and other international development agencies. Most important is the host country. Its capacity to collect and analyze data of reasonably good quality on a timely

basis combined with a willingness to make data available is central to the mission adequately meeting its own information needs. Ideally, the host country maintains data bases for each major sector. Needless to say, the number of such data rich LDC's where AID is active can be counted on the fingers of one hand. It is far more common for missions to find that usable data exists only for a limited number of sectors at best and that most data available from the host country in other areas is out of date, seriously flawed, geographically limited, etc. if it exists at all. Even determining precisely what data are held by the host country is difficult for some missions. Complicating the situation further, some countries are reluctant to make data it holds accessible to the mission or other users. In some instances, the mission obtains access to data only as a result of personal contacts staffers have established with their host country counterparts.

The obvious answer to a lack of necessary data is for the mission to fund a survey or other type of data collection effort. But to do this, the data must be considered important enough to warrant the expense. Perhaps more important is whether there is enough time for the mission to undertake data collection. Pressure to meet deadlines sometimes forces missions to push ahead without the necessary data with less than desirable results.

Greater flexibility on meeting deadlines - especially for missions with large programs operating in a variety of sectors - to allow for obtaining basic data would help. Better forward planning on the part of the missions concerning the types of data their programs require is also warranted. In part, improved planning for data acquisition should be expected as a result of the CDSS process. Ultimately, the only practical solution to improving data quality and coverage consistent with AID's development objectives is developing the institutional capacity of the host country for such work. Unfortunately, there are no magic bullets to bring this about in the near-term for most LDC's. Rather, a sustained, long-term commitment will be required to achieve institution-building of this sort. Moreover, the Agency should be prepared in making such a commitment to recognize and live with the fact that progress will consist of a series of small steps forward followed by periodic setbacks. Nonetheless, such a commitment is unquestionably called for in light of the pronounced comparative advantage the United States has in the area of data related activities. No other country, for example, can match American expertise in census operations, agricultural statistics ( in particular, area frame sampling), and population/ demographic research. Furthermore, these and other types of data collection methodologies are constantly supported by U.S. government agencies and U.S. universities whose expertise is readily available to AID.

### 5) U.S. and Local Technical Assistance for Data Related Activities

USAID missions use U.S. and local (i.e., in-country) consulting firms, freelance consultants, university teams and other U.S. government agencies for technical assistance to data related activities. Missions require outside support for various purposes. Some data related activities are too time-consuming for mission staff to perform; contractors then serve as an appropriate substitute. A common use of these suppliers is for activities requiring very specialized technical skills which mission staff do not have. For example, even though mission staff might have basic quantitative skills, and outside experts on the more esoteric aspects of sampling design from BuCen/SEU is often times necessary. The AID staffer still plays an key role as a facilitator - recognizing when specialized expertise is needed and later acting as the liaison between the mission or host country which needs the data and the specialist with the technical skills required to obtain it. Again, in-house/capability is very important in this situation to bridge the gap between the substantive information needed / the technical skills involved.

Effective use of technical assistance for data related activities is contingent upon 1) sufficient funding to cover costs for design, field operations, data processing, analysis and interpretation of findings; 2) adequate planning including a clear scope of work which anticipates the type, amount and timing of assistance required; 3) identification of appropriate suppliers; and 4) the availability of services when they are needed. These are essentially the same steps involved with procuring other types of services for AID activities. The difference is merely in the type of technical skills - e.g., quantitative, statistical methods - required. However, it seems that in a number of instances, data related activities have suffered because of inadequate attention to these details. This suggests that better planning for technical assistance is warranted on the part of the missions. Limited staff time and/or lack of pertinent skills in some missions also indicates a need for assistance from AID/Washington or regional centers to assist with procuring data support services.

#### 6) Automated Data Processing Equipment

The availability of a mini- or microcomputer to mission staff obviously is important for improving data use. As noted earlier, IRM has developed procedures for the missions to easily obtain microcomputers. This should lead to a proliferation of micros in the missions. But the mere presence of a computer system in the mission

is no assurance that data processing and analysis will be supported. Several years ago, USAID/Honduras obtained a Wang minicomputer justified on the basis of its cost-effectiveness to the mission for data entry, processing and analysis. For a variety of reasons, the use of the machine for analytic purposes has virtually been eliminated. Restricted hours for using the computer has limited user access. AID's new Mission Accounting System (MACS) consumes a considerable proportion of the machine's capacity when it is running. During the work day, the system is occupied by word processing.

Finally, no easy to use statistical package has been installed (this is also true for USAID/Egypt's minicomputer). Apparently, USAID missions need to include better support for analytic uses of data in their planning for computer systems.

A comparative ranking of the six missions selected for this study is now made using the criteria discussed above. Table 3-1 provides the rough estimate of/in-house capability of each mission.

Table 3-1 Mission Capability

<u>USAID Mission</u>	<u>Program Size</u>	<u>Management Support</u>	<u>Staff Skills</u>	<u>Tech. Assist.</u>	<u>ADP</u>	<u>Available Data</u>	<u>Overall Capability</u>
Mali	1	1	1	1	2	1	1.2
Nepal	2	1	3	1	2	1	1.7
Panama	1	2	1	2	2	2	1.7
Egypt	3	1	3	3	3	2	2.5
Zimbabwe	3	3	3	3	2	2	2.7
Honduras	3	3	2	3	3	3	2.8

Code: small/low - 1; medium - 2; large/high - 3

It should be kept in mind that the missions are ranked on a comparative basis; for example, USAID/Nepal has a medium sized program in comparison to the other five missions, and so on. Second, the rankings should be interpreted only as ordinal scores; that is, USAID/Zimbabwe has a larger program in comparison to USAID/Mali, but a "3" score does not mean it is actually three times larger. Similarly, a low score on senior management support does not imply there is no support for data related activities or even opposition to better data use. Rather, in comparison to the other five missions, there appears to be less support than found elsewhere. In other words, a mission's rank is only a relative and not an absolute measure. The estimate of overall mission capability (an average or mean of the six rankings) should also be interpreted conservatively as suggesting where each mission stands vis a vis the others.

The rankings for program size are fairly straightforward. They are a composite of funding, staff size and program diversity. Senior management support is much more difficult to estimate. Support can vary because of the individuals involved (e.g., mission director X places little value on data related activities); the demands of the program (e.g., major implementation problems have arisen and taken precedence over other activities); or some combination of these and other factors. The rankings on management support also suggests a degree of uniformity of opinion throughout the mission which might

be misleading. For example, the mission director might not be interested in devoting more staff time to data related activities while office directors believe such work would be very beneficial. Though high, medium and low support is indicated, actually a dichotomous yes - no split is suggested by the scores. The division between the high ranking missions - Zimbabwe and Honduras - and the low ranking missions - Nepal, Egypt and Mali - was fairly apparent by the end of the study. Only USAID/Panama ranks an intermediate position. The current mission director has expressed interest in increasing economic analysis in mission operations, but actual improvements remain to be accomplished.

Staff skills were estimated on the basis of the number of mission staffers who reported having quantitative skills or data management experience. USAID/Zimbabwe, for example, has at least two staffers who have such skills, one of whom is highly proficient at data related activities. Technical assistance to augment mission capability is weighted or biased toward the availability of local contracting firms. Though USAID/Nepal and USAID/Mali have been able to get high quality technical assistance from U.S. firms, local contractors have at best minimum competency for data collection and analysis (but they are improving, which is important).

Automated data processing equipment is based on current and soon to be acquired computer systems in the mission. USAID/Egypt and USAID/Honduras both have Wang minicomputers. USAID/Panama has a Wang word

processor and as of November, 1982, planned to obtain a microcomputer. Neither USAID/Mali nor USAID/Zimbabwe presently have ADP equipment, but both intend to purchase micro's in the near future. USAID/Nepal is something of an oddity. It is too remote to make Wang equipment a practical solution (i.e., there is no Wang service in Kathmandu). As of September, 1982, the mission had not purchased microcomputers nor were there any plans to do so. However, with the support of the population office director, a number of mission staff and contractors have purchased their own microcomputers. The office chief assists with maintenance problems. In other words, officially, there are no microcomputers in the mission; in fact, micros are used for mission activities.

The availability of data to the mission is largely a function of host country capabilities. But it is included with mission factors because available data is also a product of AID support and a determinant of the amount and type of data use missions are capable of performing. USAID/Nepal and USAID/Mali rank at the lowest end of the scale. The situation of USAID/Egypt and USAID/Panama is somewhat better. USAID/Zimbabwe has access to very uneven data bases in terms of coverage. Excellent data are available from the GOZ on the modern sector, but virtually no reliable data exist on the economic and social conditions of the traditional sector.

An estimate of overall mission capability is made by calculating an average rank based on the set of six scores. The missions can be categorized as high, medium or low using the following cut-off points:

low : less than 1.5 - USAID/Mali

medium : 1.5 to 2.4 - USAID/Nepal, USAID/Panama

high : greater than 2.4 - USAID/Egypt, USAID/Zimbabwe, USAID/Honduras

As would be expected, the largest missions have the highest capability for data related activities. However, the mission with the largest program - USAID/Egypt - ranks lower than two other missions and is a borderline case between medium and high capability. It is also apparent that even though missions in the same category have comparable capabilities, this results from a somewhat different combinations of in-house factors.

### 3.2 Host Country Capability

The following factors seem to be the principal factors determining host country capability for data related activities:

#### 1) An Established Tradition of Using Data for Decision-making

The assumption is frequently made, and wrongly so in many countries that the importance of collecting and using pertinent data to guide decision-making is self-evident. Western societies in general operate under what is perceived as the "rational" approach to planning and decision-making. That is, having adequate information about conditions

is considered essential before action can be taken. This is true in both the public and private sectors. A manufacturer planning a new product typically conducts a market survey to determine potential demand. Stock market analysts use an array of data sources and analytic systems to guide investment. In the public sector, the amount of information marshalled for proposed action can be staggering - from OMB and CBO analyses of the effects of budget cuts to environmental impact and traffic flow studies for the construction of shopping malls.

The "rational" model of information use is far from universal and certainly not an established tradition in a number of developing countries. In some LDC's, data collection and analysis is viewed as one more Western aberration. Instead, decisions are made for any number of reasons other than what the "rational" model would define as empirical and objective. This would include purely political motivations which ignore economic justifications in favor of cronyism and personal self-aggrandizement. This is not to say such factors do not also influence and even determine decision-making in developed countries. The difference is in the blatancy with which this is done and the absence of self-righteous rhetoric about how decisions are made in an honest and above board fashion in countries which do not operate according to the "rational" model.

lack of a  
The /traditon of using data to guide actions and its corollary - no felt-need for better data collection and analysis - can have dire consequences for development projects which are to improve and expand

the host country's data related activities. In such situations, plans for data collection and analysis might be circumvented or ignored. Efforts to improve data use within host country ministries can meet with little if any success as a result.

## 2) Current Interest and Support for Data Related Activities

The success of data collection and analysis components of AID projects is highly dependent on the current interest and support of the host country for data related activities. Moreover, host country support is not uniform and can vary among ministries. Nor is host country support constant over time. Changes in government administrations can dramatically alter support, or the lack of it, for data collection and analysis. The Government of Honduras provides a good example of this. When the Suazo administration took office, many government officials were replaced. In some ministries, officials who had supported plans for data collection and analysis in USAID projects were replaced by individuals who clearly did not hold the same interests. The data related component of the Rural Technologies Project, for example, ran aground precisely because of such a change in GOH support. In other ministries, such as the Ministry of Education, the exact opposite occurred. Those appointed to head the MOE strongly support better data use and mission staff are now optimistic that improvements will be made.

Variation and instability in host country support for data related activities have clear implications for AID's efforts to improve key data bases. First, variation in support among ministries determines

where the most likely opportunities for institution-building efforts exist. That is, the mission should focus on those ministries which demonstrate genuine support for better data use and bide its time until similar favorable conditions prevail in other ministries. Second, changes in government can lead to establishing a stronger reliance and, in time, a tradition of data use for decision-making. In such instances, AID should encourage this very positive development through additional training and technical assistance to the host country. On the negative side, when the host country has neither a tradition nor an interest in developing its capability for better data use, missions should consider alternative means to meeting short-term program and project information needs other than through the host country. For example, such short-term solutions might entail data collection using contract, non-government/<sup>personnel</sup> involving little if any host country participation. Missions might also consider developing in-house data banks in lieu of host country support for maintaining key data bases. Similarly, loss of host country support for data related activities can play havoc with AID's institution-building efforts. On the other hand, re-newed interest by the host country in improving its information systems creates a prime opportunity for the mission to build host country capability in this area. In general, AID needs to make a long-term commitment to developing host country capability for data related activities while recognizing that progress will occur unevenly in a series of small advances followed by periodic setbacks.

### 3) Human and Financial Resources

Even given a genuine interest on the part of the host country, the majority of LDC's are constrained, some severely, by limited human and financial resources available for data related activities. At the extreme end of the scale are countries like Mali which are hard pressed to maintain current government services and meet present operating expenses. The idea of expanding data collection and analysis, as useful as that might be, is simply beyond the pale for these countries. The recurrent costs of such activities simply exceed the government's budget. Even in countries which have the interest and financial resources to improve their information systems, such as Zimbabwe, manpower and particularly skill staff are scarce. Many government workers who acquire technical skills leave government service for better paying jobs in the private sector. This problem is especially acute in countries like Jamaica where government salaries are abysmally low and opportunities in the private sector, either domestic or abroad, are readily available. Clearly staff turnover of this sort defeats the purpose of training programs intended to produce skilled manpower needed by the host government. Insufficient human and financial resources largely account for the lack of basic data for key sectors in many

LDC's. The lack of available data constrains the mission's in-house data use while the lack of skilled staff limits the types of data related activities the mission can support.

#### 4) Automated Data Processing Equipment

Automated data processing equipment, the software necessary for data analysis and a service facility to maintain computer systems are as essential to the operation of efficient information systems in LDC's as they are in developed countries. The computing capacity of LDC's vary widely, but in general, most need better computer systems than they currently have. Perhaps with the exception of Honduras, in the other five countries selected for this study, demand for computer support exceeds the capacity of existing systems. In the extreme cases like Mali, only the most essential, day to day tasks have been computerized (e.g., the government's payroll). Even in better equipped countries, such as Zimbabwe, there is a genuine need for more and better equipment. Limited finances to purchase computers requires that many activities which could be computerized are done by hand. The result is high inefficiency and staggering delays in completing tasks. Inadequate maintenance of existing equipment which causes excessive down-time further slows data processing. Data analysis, of course, receives short-shrift under such circumstances. In short, data use could be improved in many LDC's by computerizing existing operations and providing adequate maintenance without any additional data collection.

### 5) Political Context

To develop and maintain sound data bases, a minimum level of continuous support by the host country is necessary. Funds and staff obviously have to be provided for the daily operation of the systems. But equally important, the central government has to reinforce the message to its staff that the information aspects of their jobs are important.

Basic support for data related activities is highly subject to political factors. A change in government administrations and staff can significantly alter this support. Again, Honduras offers a good example. Mission staff reported that ministries had been making relatively good use of data to guide their planning, but recently some ministries had become more "action oriented" in an attempt to achieve some immediate improvements in high priority areas, such as land reform. In countries where governments change every twelve to twenty-four months, data related activities can become a very on again, off again proposition. As a result, plans for data collection and analysis as well as increasing the host country's capability for data use can be disrupted and, in some cases, eliminated completely.

Political stability or the threat of it can lead to heightened concern about access and use of available data. The possibility exists that data can reveal poor performance on the part of the government which in turn can fuel internal dissent. Sometimes such criticism is

justified, but not always. In the extreme cases, internal security concerns leads to strict controls over data the government views as potentially sensitive in nature. As a result, the USAID mission might have only limited access to data or encounter substantial delays in the release of data held by the host country. In areas where warfare has actually broken out, data collection, of course, has to be curtailed. For example, the level of violence in areas of Honduras along the border with Nicaragua or El Salvador makes data collection entirely too risky. Warfare also causes population displacements and interrupts routine economic and social activities. Even after fighting has stopped, it will take some time for normal economic and social patterns to be re-established. Unless there is some very special reason for doing otherwise, data collection should be worked around the political realities of the country.

#### 6) Geography, Climate and Infrastructure

The periodicity of an event or activity generally dictates when data should be collected. For example, agricultural data on farm inputs, labor use and crop yield have to be coordinated with the local cycle of farm activities. Being able to get to the project site and travel about the area is necessary for collecting data at the right time. But this is not always an easy matter due to the geography and climate of certain countries. This is particularly true where basic infrastructure, especially bridges and all weather roads, is inadequate

or do not exist in parts of the country. Nepal is a good case in point. The geography, climate and lack of basic infrastructure in Nepal make even the simplest data collection efforts a major undertaking. Though not as extreme as in Nepal, such conditions are found elsewhere and do impede data collection efforts. The result is, again, to limit the availability and quality of data.

Using the preceding factors, the six countries selected for this study can be ranked in terms of their capability for data related activities. Table 3-2 presents these rankings.

Table 3-2 Host Country Capability

<u>Country</u>	Tradition of Data Use	Current Interest	Hum. & Fin. Resources	ADP	Polit. Context*	Geogra. Climate*	Overall Capabilit
Mali	1	1	1	1	2	2	1.3
Nepal	1	2	1	1	2	1	1.3
Egypt	1	1	3	2	1	3	1.8
Panama	2	2	2	2	2	3	2.2
Honduras	3	2	3	3	2	2	2.5
Zimbabwe	3	3	3	3	3	2	2.8

Code: small/low - 1; medium - 2; large/high - 3

\* Political Context rated in terms of the degree to which political concerns of the host country impede data related activities. code: 1-substantial; 2-some; 3-little

\* Geographraphy & Climate rated in terms of the degree to which enivronmental factors impede data related activities. code:1-substantial; 2-some; 3-little.

A tradition of data use is contingent on having both an established perception of data and analysis as useful if not essential tools for decision-making and the resources to collect and analyze data. Egypt, for example, is ranked low because several ministries lack a genuine appreciation for the utility of data as guide to decision-making; while Nepal scores poorly in part due to a lack of resources and for data related activities. It appears that the tradition / interest in data use of a country are not necessarily correlated with its level of economic development. Other factors might account for its propensity to use data, such as the colonial pattern established before independence for some countries (e.g., former British colonies might make greater use of data than former French or Dutch colonies). Current interest in data use is difficult to estimate for some countries because interest can vary among ministries of the government. Honduras, for example, might deserve a higher rank, but mission staff reported some decline in support for data related activities recently. At the upper end of the scale, Zimbabwe most strongly encouraged better data use in government operations at the time this study was undertaken.

The capacity to collect and analyze data reflects both the host country's own resources as well as those provided by international development agencies. Heavily funded USAID programs, therefore, can make a significant impact in terms of substantially increasing the resources available to a country for data related activities. This

is the case for Egypt and Zimbabwe and will probably soon hold for Honduras as foreign assistance to that country increases. At the other end of the spectrum, small or moderate sized programs in very poor countries like Mali and Nepal are limited in terms of offsetting the lack of funds and skilled manpower for data collection and analysis. Automated data processing equipment is also tied to the host countries resource capability. It is fairly clear that Honduras and Zimbabwe are comparatively better equipped than the other countries. Egypt has acquired a number of new computer systems recently with USAID assistance and should rank higher as these systems become fully operational and adequately maintained.

One of the more difficult factors to estimate is the influence of political concerns on data related activities. In comparison to the other five countries, political issues seem to most interfere with data use in Egypt. This point was made by several USAID/Egypt staffers. The concern for maintaining political stability and the highly centralized system of information dissemination and control appear to impede the availability and use of data. At the other extreme, the Government of Zimbabwe is very open and even favorably inclined toward more data collection and analysis despite its own internal security concerns. Honduras is somewhat problematic. Though the government has made data readily available to USAID/Honduras, increased violence in border areas could lead to greater restrictions and certainly questions the viability of data collection in those areas.

The other three countries fall between these two extremes - political issues intrude to some limited extent on certain topics, but, for the most part, they remain in the background (which in itself is a positive situation).

Geographic and climatic factors clearly affect data collection and analysis most in Nepal. The lack of all-weather roads and flooding during the monsoon season isolates sections of the country making data collection impractical in those areas. Travel in certain parts of Mali, particularly in the northeast section of the country is problematic (it takes a long time), but it is not as difficult as is travel throughout Nepal. Zimbabwe's rainy season complicates data collection somewhat, but the country's relatively good transportation system minimizes the problem. Difficult access to certain parts of Honduras and the mountainous geography of the country (which interferes with aerial photography, for example) affects data collection to some degree. There are no significant impediments to data collection caused by geography or climate in Egypt and Panama.

### 3.3 Summary of USAID Mission Capacity for Data Related Activities

A rough estimate of the overall capacity of the USAID missions for data related activities can be made by adding the average (overall) rankings from tables 3-1 and 3-2 as follows.

Table 3-3 Overall Mission Capacity

<u>USAID Mission</u>	<u>Average Rank</u>
Mali	2.5
Nepal	3.0
Panama	3.9
Egypt	4.3
Honduras	5.3
Zimbabwe	5.5

The average rank should be interpreted only as an ordinal score indicating the relative position of each mission vis a vis the other five. Minor numeric differences are, therefore irrelevant.

Table 3-3 suggests that USAID/Zimbabwe and USAID/Honduras represent higher capacity missions; USAID/Egypt and USAID/Panama, medium capacity; and USAID/Nepal and USAID/Mali, lower capacity. The higher capacity missions will, in general, be ones where 1) a fairly sound in-house capability exists and 2) the host country is committed to better data use and has human and financial resources for doing so. Table 3-3 reiterates the importance of program size and host country capability. The three highest ranking missions - USAID/Zimbabwe, USAID/Honduras and USAID/Egypt - are the largest of the six in terms of program size. However, even though mission size is related to capacity for data related activities, the relationship is not perfect. USAID/Egypt has the largest program of any USAID mission, yet it ranks considerably lower than USAID/Zimbabwe and USAID/Honduras. This is

due in large part to the host country's low capability. In particular, several ministries are not especially interested in improving data use and the GOE places comparatively stricter controls on access to and release of government held data. The mission also contributes to the lower ranking. It is remarkable that a mission with the resources USAID/Egypt has does not provide better in-house support for data related activities, particularly in-house data use. A fair conclusion to be drawn from this is that the capability of the host country for data related activities places an upper limit on what is appropriate in terms of technology transfer and what is possible in terms of expected performance. This implies that developing the institutional capacity of the host country for data related activities should have beneficial effects on mission data use as well. That is, as the host country's capacity increases, it is reasonable to expect better information use by USAID missions because more data of higher quality should be available.

A comparison of the missions' capacities can be further summarized in terms of a crosstabulation between mission capability and host country capability as follows.

Table 3-4

		<u>Host Country Capability</u>		
		Low	Medium	High
<u>Mission Capability</u>	Low	Mali		
	Medium	Nepal	Panama	
	High		Egypt	Zimbabwe Honduras

Table 3-4 presents in more graphic form the composite rankings of the six missions discussed earlier. The table also illustrates the basic similarities missions share in dealing with data related activities. That is, missions in the same column confront similar problems in improving host country capability. Missions in the same row have comparable capabilities for in-house data use and support for data collection and analysis. Such similarities suggest that the table could be collapsed further based on the potential of missions to improve host country capabilities for data collection and analysis. Table 3-5 provides a typology differentiating three general categories of mission capacity/potential.

Table 3-5

		<u>Host Country Capability</u>		
		Low	Medium	High
<u>Mission Capability</u>	Low	<u>1</u>		<u>2</u>
	Medium	<u>3</u>		
	High			

Category One - low mission capability combined with low to medium host country capability - represents the worst case situation. Both the USAID mission and the host country have low capability for data related activities. For the most part, there is little the mission can do beyond trying to integrate data collection and analysis into suitable projects if sufficient financial and human resources can be made available. Often times this will not be possible even though

better data are needed. Institution-building will at best be gradual and the possibilities for improving available data will be very limited in the short-term. Missions operating under these conditions are the ones most in need of additional support from AID/Washington. Category Two - high capability host countries - should be the easiest for USAID missions to assist. Their basic strategy should be to build upon existing institutional capacity. Higher capability countries typically have a core of adequately trained staff. The major constraints to better data use are usually recognized and, for the most part, the country only needs some additional training for staff; short-term, highly specialized technical assistance; or more automated data processing equipment. Category Three - medium to high capability missions combined with low to medium capability host countries - is where AID could have a significant impact on improving data related activities. These missions have resources, albeit limited, for improving data collection and analysis and should try to build a basic capability in host country ministries in key sectors consistent with the mission's program. If these missions are successful, the host country will move from an extremely limited capability (or none at all) for data related activities to one which is rudimentary but adequate (or nearly so). This advance can be comparatively greater than that made by

further refining an established capability. In fact, many of the information needs of LDC's can be met with the simple, bare-bones systems which can be developed with the support of medium to high capability missions. To reiterate a basic point of this report, AID should be very wary of trying to transfer statistical techniques and sophisticated information systems to countries which are inappropriate from the standpoint of host country needs and capabilities. In short, the greatest potential for improvement from institution-building exists where the USAID mission has a medium to high in-house capability. This also means that the Agency has valid reasons to increase mission capability for data related activities. To do this the Agency must first come to grips with the institutional constraints discussed in Part Two of this report. At the operational level, the factors affecting mission capability also identify areas where in-house improvements could be made (e.g., encouraging greater senior management support, capitalizing on existing staff skills, providing worktime for using those skills for the benefit of mission operations). Mission capability could also be augmented if AID/Washington established a division responsible for assisting the missions with data related activities. Regional data support centers operating out of selected missions are another possibility. These and related issues will be discussed in subsequent parts of this report.

### 3.4 Mission Strategies for Improving Host Country Capability for Data Related Activities

This section concerns how USAID missions are currently attempting to assist host countries/develop their institutional capacities for data related activities. Some additional options missions might also consider are suggested in connection with each of the key factors affecting host country capability.

#### 1 & 2) A Tradition of Using Data for Decision-making and Current Interest in Improving Data Related Activities.

The lack of an established tradition in the host country of using data for administrative and planning purposes and the lack of genuine interest in increasing its capacity to do so pose a major obstacle to improving available data. None of the six missions selected for this study provides an example of effectively dealing with this problem. On the one hand, indifference on the part of the host country places the mission in the position of insisting that the host country improve its capability. In effect, the mission is unilaterally stating that the Western perspective concerning what constitutes rational decision-making will be used in connection with AID funded activities. This can cast the mission in a very negative light (e.g., arrogant and demanding) which runs contrary to other foreign policy objectives and the Agency's own code of working cooperatively with the host country. On the other hand, the mission, as a U.S. government entity, must

use its financial resources as efficiently and responsibly as possible. Clearly sound project design, monitoring and evaluation - all of which can require quantitative data - are essential for such fiscal responsibility. From this perspective, the mission is obligated to obtain adequate data, the primary source of which is the host country.

It is quite easy to say that missions in this situation are engaged in changing the behavior of the host country so that data figure more prominently in the decision-making process. The next step is to admonish the missions to target on host country identified problems, provide positive examples of data use, and so forth. In fact, mission staff seem perfectly well aware of the nature of the problem and the logical options available to them. The reality is that even if the mission succeeds, it will be years before noticeable improvements occur in many countries. The question is what to do in the meantime to meet pressing information needs.

In lieu of host country capability and interest, the mission has no alternative than bearing a disproportionate share of the responsibility for supporting the collection and analysis of data in the sectors central to its program. This will probably involve greater direct involvement by mission staff or more outside technical assistance to assure that data are collected and analyzed as planned. Otherwise, the mission will continue to confront inadequate information bases. Greater support and involvement with data collection and analysis, however, is possible

only for high capability missions and perhaps some medium capability missions. This is one area where data support services to augment mission capability are clearly needed.

As a general strategy for dealing with host country indifference toward better data use, AID must maintain a very consistent perspective about the importance of data collection and analysis. It makes little sense to design data collection components into projects, provide any number of justifications for the necessity of these activities and then later, jettison them because of project implementation problems or host country disinterest. Similarly, evaluations which bemoan the poor performance of the host country concerning data collection and then recommend continued funding only evince AID's own indifferent commitment to these activities. If AID fails to show strong support for the data related components of its projects, the message that these activities can be ignored will not be lost on the host country. Mission staff also need to reinforce AID's position that data collection and analysis are important project objectives. For example, data collection and analysis should be considered very important for projects which are implemented in stages where data from an earlier phase are to be used to help design follow-on activities. Data related activities should also be accorded high priority in projects designed to improve ministry operations or establish a planning and evaluation unit. Third, projects which can provide data useful for assessing sector-

wide development should be given special emphasis. Additional examples could be cited, but the point is that the missions need to demonstrate a commitment to accomplishing the data related objectives of such projects. The means for doing so is funding. If the host country continues to shirk its responsibilities for data collection and analysis and these activities figure prominently in the project (i.e., inadequate data will undermine project effectiveness or success), then AID should consider reducing or stopping funding. No clearer, more definitive statement could be made to the host country. Furthermore, if a project were terminated for these reasons, given the pressure on missions to keep the money moving, senior management would give far more attention to accomplishing data related activities as planned.

Taking a harder line against compromising project objectives for data collection and analysis might lead to greater compliance by the host country, but a more persuasive approach is necessary for generating interest in better data use. The strategy missions currently use is to provide positive examples of how improved data use facilitates ministry operations (along the lines of "seeing is believing"). The typical outputs of these projects are data capture and analysis systems which contribute to better service delivery, improved cost effectiveness and increased planning capability. The key element is designing these projects around existing ministry (or other host country agencies)

tasks to demonstrate how better data use contributes to, rather than handicaps, their performance. These projects usually involve substantial staff training and often include construction of new facilities or acquisition of computer systems. Technical assistance is provided by either long-term advisors or through a series of TDY's. (There is considerable disagreement among AID staff about the effectiveness of long-term versus TDY technical assistance. Probably the best guide to determining which is preferable is the host country's existing capability. Long-term advisors might be better for low capability countries; TDY's might be better suited to higher capability countries.)

Not all efforts to promote better data use by the host country need be large scale undertakings. In many LDC's, significant amounts of data are collected; however, they lack sufficient numbers of analytically skilled personnel to fully utilize the data. In such situations, AID could sponsor short-term analytic support projects directed to better analysis of available data. Outside analysts are also accorded greater credibility than local analysts in many LDC's which will facilitate use of the results by the host country government.

The success of these efforts is largely dependent on the final products: analyses interpreted in terms of alternative options with their associated pros and cons in a form which decision-makers, who do not have statistical skills, will readily understand. The project also has to demonstrate to senior officials that 1) the techniques and systems are affordable and cost-effective and 2) ministry or agency staff are capable of performing the work independently (i.e., that the

If these requirements can be met, there is a much greater chance of generating interest among upper and middle level staff necessary to sustain improved data use after the project is completed. But as mentioned earlier, the mission should be prepared to make a long-term commitment to this process because there are no quick fixes for generating host country interest in data related activities.

### 3) Financial and Human Resources

USAID missions are making a concerted effort to help LDC's develop the skilled manpower needed for better data collection and analysis. Depending on the type and level of skill training needed, host country staff are trained abroad, in-country or on the job. Missions are assisting with the construction and staffing of technical colleges and institutes, particularly in the area of agriculture, which will offer instruction in analytic techniques and research methodology. AID projects which include a data related component, in most cases, involve host country staff directly in the process to provide them with training and field experience. In short, USAID missions are doing a very good job in this area.

Mission staff frequently cited the fact that a significant number of host country staff who receive technical training leave government service for jobs in the private sector. This problem varies from country to country depending on the availability of employment outside of government, but it seems to be a very common pattern. The effect is that the host country does not improve its capability for data collection and analysis. There is no solution to this situation other than to keep training new staff. Perhaps the jobs in the private sector

will become fewer in number as they are filled by former government employees. In any case, as long as those leaving government take jobs in the country and not abroad, AID support for training is still serving a useful function - it is strengthening the private sector, which is a perfectly legitimate goal. As for coping with the chronic shortage of skilled staff, the USAID mission might consider turning to the private sector to obtain the services needed. Most countries have local market research groups or other consulting firms capable of conducting sample surveys and performing simple analyses. With some specialized technical assistance, perhaps on a TDY basis, these firms could be used for data collection for development projects. They would also be interested in such work if AID provided the funding, hence, overcoming the problem of low government salaries (which is a key reason why skilled staff left government in the first place). In time these firms could work independently or with minimum technical assistance. If AID were to encourage contracting for data related services using local firms, a new or expanded job market would be created; a local capability would be established; and host country and mission information needs would be better met as a result.

The recurrent costs resulting from increased data collection and analysis is a serious problem for many countries. There are several courses of action USAID missions can follow to minimize the burden of recurrent costs for data related activities. First, the mission can provide the necessary funds for establishing and maintaining data bases for the sectors and regions of the country where its program

is concentrated and where it plans to remain active. Second, in a number of countries, demand for data and analysis is donor created and, in some instances, far exceeds the host country's capability to comply. USAID missions should avoid contributing to this problem. Requests for data, special studies, etc. should be supported by the funds and/or technical assistance to perform the necessary work. Third, projects which have an institution-building component should be carefully reviewed to guarantee that the technology involved is appropriate for the host country in terms of staff skill and cost requirements. A rule of thumb might be that such institution-building efforts are to develop systems of improved data use which will be no more expensive than current operations.

#### 4) Automated Data Processing Equipment

Additional funding for computer equipment and software development is the only solution to meeting host country needs for better systems. A rule of thumb for missions purchasing ADP for host country operations using project funds is to "think medium". That is, carefully identify the amount and type of data processing to be done and then acquire a system which meets those needs and then some (actually most missions will work through a consultant with expertise in computer hardware, but the same principle applies in the mission's dealings with such consultants). A little excess capacity allows for meeting new or unanticipated uses which arise. However, cost should be a primary

consideration not only in terms of the initial outlay for equipment and installation, but also for the support staff required for daily operation and maintenance. In general, buying systems which have capacity far in excess of what is needed (e.g., a minicomputer when a micro would suffice) should be avoided.

The usual means of funding computer equipment purchases for host country operations is through project funds. An alternative is available to missions which have a Commodity Import Program. USAID/Zimbabwe, for example, specified that a percentage of CIP funds would be used for purchasing computer equipment for the public and private sectors.

#### 5) Political Context

As mentioned earlier, political factors can have an adverse effect on improving data related activities. Limited access to data held by the host country can impede mission data use. A very sound strategy USAID/Egypt follows is to share whatever data and analyses the mission acquires with their Egyptian counterparts. For example, the program office provides copies of data and studies to the GOE in part to assist them with their planning, but also to foster a better exchange of information.

Coping with the problems created by political instability might require greater direct involvement on the part of the mission with data collection and management. To assure the maintenance of essential

data bases (i.e., those central to the mission's program), an in-house data bank might be the only practical solution. Properly documented copies of data sets could be acquired by the mission and stored in a format suitable for use on its computer system (mini or micro). In a sense, the mission would serve as a surrogate for the host government to protect valuable sources of information. The data would always be available to the host country and analyses based on it would routinely be provided to the current government. An in-house data bank would make the data more accessible to mission staff and contractors while at the same time providing the continuity necessary to develop good data bases. The limiting factor here is, of course, mission capability for data management. Again, this suggests a genuine need for a data support service within the Agency.

#### 6) Geography, Climate and Infrastructure

Data related activities are subject to the constraints imposed by geography, climate and the lack of infrastructure and have to be worked as possible. around these factors as best / For example, data collection should be scheduled for the dry season while data processing and analysis should be done during the rainy season, or the reverse, as in the case of agricultural data. However the problem is dealt with, even very difficult conditions are no excuse for abandoning the objective of better data use. The proof of that claim is that if data can be collected in Nepal (as USAID/Nepal's population and agriculture projects have done with considerable success), then data can be collected anywhere.

### 3.5 Conclusions

The main conclusions to be drawn from the preceding analysis of USAID mission capacity for data related activities are as follows:

- 1) The fundamental problem of improving USAID mission data use and support for data collection and analysis is the Agency's present ambiguous commitment to meeting information needs with the appropriate data. Until the Agency recognizes that demands for analytically sounder work will only be met if the resources to do so are made available to the missions, improvements in data use will continue to be the exception rather than the rule.
- 2) With AID/Washington support, there are a number of in-house improvements which could be made with a relatively small investment of time and funds. The most important appears to be providing the worktime necessary for staff who have quantitative, analytic skills to assist the mission with its data related activities. Inadequate computer support can easily be resolved by purchasing microcomputers and appropriate software for mission operations.
- 3) Though in-house improvements are possible, there are definite limits on what individual missions can accomplish without additional support from AID/Washington. As pointed out in connection with the ranking of the missions selected for this study, the designations of low, medium or high capacity are comparative assessments. A high capability does not necessarily mean high in some absolute sense but only in relation

to the other five missions. Therefore, if high capability missions are limited as to what they can accomplish in terms of improving their data related activities, low and medium capability missions are even more restricted.

4) Given the limitations of the missions for in-house improvements, AID/Washington needs to develop data support services to assist missions improve their data related activities. A key justification for increasing mission capacity is that the findings of this study indicate that higher capacity missions are better able to assist host countries develop their own capabilities for data related activities.

5) Developing the capability of the host country for data collection and analysis is the only logical course for improving available data. A basic capability for data related activities is fundamental to improving government operations and planning. Institution-building for better data use is consistent with AID's overall development goals. Without such a capability, improvements in the quality and quantity of available data are unlikely. One short-term solution might be for USAID missions capable of assuming certain responsibilities for establishing and maintaining essential data bases to do so until the host country develops the necessary capacity.

4. A Pragmatic Approach to Data Sources, Quality, Representativeness,  
Amount and Frequency of Collection

This report discusses the utility and, in many instances, the necessity of taking a pragmatic approach in AID's support for data collection and analysis. A pragmatic approach entails an acceptance of compromising statistical standards of data quality (in particular, validity and reliability) to accommodate data collection to the conditions under which AID operates (e.g., the general difficulty of data collection in LDC's, AID budget and time constraints). How much statistical rigor has to be compromised to make data collection possible in the context of an AID project depends on the specifics of the situation - in some cases, perhaps none; in others, perhaps a lot.

This is clearly not a view most professional statisticians would advocate. But maintaining rigorous standards for all of AID's data collection efforts is a luxury the Agency cannot afford. In a number of instances, statistical rigor is unrealistic and even counter-productive given the information needs and resources at hand. As one AID staffer who has observed numerous data collection efforts supported by AID described the problem, the ideal (statistical rigor) becomes the enemy of the good (adequate information). On the other hand, a pragmatic approach to AID's data related activities is not a call to throw all standards to the wind for the sake of expediency. The goal of obtaining data of as high a quality as is practical and necessary ought to be central to AID's support for data collection and analysis.

But this does not mean that data quality must be uniformly high for all of AID's activities. Nor should this goal be treated with such overriding importance that it prevents data collection when circumstances or resources will not permit obtaining data which meet statistically rigorous standards. Many AID projects are, for one reason or another, unable to generate high quality data. From the pragmatic perspective of imperfect data are still better than no data, flawed data do have utility if the limitations of the data are recognized and not exceeded. In short, AID's position should be to collect the best data circumstances allow, yet it should set its standards for what constitutes usable data far lower than those adhered to by professional statisticians.

During the course of the interviews conducted for this study, staff in each of the missions expressed concern about fundamental aspects of data collection and analysis in projects they were designing or managing. The most frequently cited areas of concern were the following:

- 1) the sources of data available to the project and methods of collection required to obtain it;
- 2) the level of precision or quality necessary to meet project information needs and other objectives;
- 3) the representativeness of the data needed to accurately reflect conditions in the project area;

4) the amount of detail the data should provide about the topic being investigated; and

5) the frequency of data collection during the course of the project.

First, it should be noted that in the context of an AID project, there are no single, uniform answers to these questions nor hard and fast rules which apply without exception to every situation. Second, any number of additional technical decisions have to be made before a survey or other data collection effort could actually be conducted. But mission staff are quite right to be concerned about the above planning and design questions because the success of the data related component of the project depends heavily on how they are answered. Even when the mission relies in technical advisors to resolve these questions, USAID staff still need to determine whether their advice is consistent with the goals, objectives and resources of the project. Similarly, host country staff frequently have strong opinions about the scale and scope of data collection and analysis. USAID staff need to evaluate their views before agreeing to them. This study has also found instances of mission staff having primary responsibility for dealing with these questions. A pragmatic approach to data collection and analysis could provide useful guidance in each of these situations.

Pragmatic strategies have guided some of AID's more successful efforts to support data collection and analysis. The following sections briefly discuss basic options and considerations pertinent to the questions cited above.

#### 4.1 Data Sources

When data are not available, inadequate or inappropriate for a specific need, USAID missions have only two main options for obtaining quantitative data. The mission either supports a sample survey or works through host country ministries to obtain data from their service or operations statistics. Both have their advantages and disadvantages which experienced mission staff seem to recognize. However, there are some aspects of these two alternatives which warrant greater attention.

##### 4.1.1 Sample Surveys

AID supports numerous sample surveys for project and program purposes. The scale of these surveys ranges from the national level (e.g., nationwide household surveys) to small, geographically focused studies (e.g., crop yield surveys of fifty to one hundred farms located in a specific district). Obviously the information needs at hand dictate what the appropriate scale of the survey must be. Where AID seems to have some difficulty is in determining the size of the survey; that is, how many regions or districts ought to be included, how many respondents ought to be selected, etc. If anything, AID errs in the direction of too big rather than too small. This is definitely a bad thing to do. As size increases, cost, time, technical expertise, manpower and data processing requirements increase commensurately. Furthermore, there is a tendency to invest in one big survey rather than spread those funds across two or smaller surveys. This is done

despite evidence that small scale surveys have been very useful to USAID missions and host countries.

A good example of the utility of small scale surveys is the Seed-Production and Storage Program supported by the Agriculture and Resource Conservation Office of USAID/Nepal. A major output of the project is the construction of mini seed storage houses in remote areas where agricultural potential is high but transportation is difficult. The program will sponsor local production of improved seed stocks to overcome the transportation problem. The project needed information on suitable sites for the seed houses and in particular, selecting locations where the farmers were interested in producing seed stock. The International Agricultural Development Service (IADS) developed a small survey to meet the projects information needs. The consultant hired for the assignment constructed a very simple questionnaire to collect data on the size of landholding, crop production, crop mix, current source of seed stock, available extension services and other questions pertaining to the location and management of the seed houses. Farmers from two panchayats were selected for the interviews. From the analysis of the data (which consisted of just simple percentages), the survey results will provide useful guidance for selecting the best construction sites for seed houses.

Additional examples of effective, small scale surveys could be cited from the other missions selected for this study. As a group, these surveys share certain common characteristics. Most important,

they tend to be highly focussed in terms of specific data to be collected and the area to be surveyed. Second, they tend to use very simple sampling designs (e.g., how respondents are selected). As a result of their simplicity, these small scale surveys can be completed quickly, require minimal technical assistance, facilitate analysis and are fairly inexpensive. Furthermore, such surveys can also provide host country staff or local contracting firms with the opportunity for additional field experience at tasks which are commensurate (or nearly so) to their skill level. Because small scale surveys can be completed rapidly (sometimes design through analysis within four weeks), they can generate limited amounts of timely data essential for better planning by the USAID mission and the host country. In short, this type of survey represents a very useful tool within the capability of many USAID missions.

As useful as small surveys can be, the data they generate have very definite limitations. The highly focussed content and restricted geographic coverage of such surveys limits the user's ability to generalize to larger segments of the population. Second, the quality of the data is sometimes lower than that produced by larger, more elaborate surveys. This is not something intrinsic to small scale surveys; rather, it seems to be an artifact of funding and time constraints under which these surveys are conducted. Third, small scale surveys done on a one-shot basis (i.e., they are not replicated at a later date)

are consistent with, but not the best mechanism for the institution-building objectives of the USAID mission. Small scale surveys are frequently conducted because they can quickly obtain basic data for a specific task or purpose. In that context, there is little room for extensive staff training as compared to what is possible in large surveys. However, a successfully completed small scale survey which provides information sufficient for the limited needs at hand demonstrates the utility of less than perfect data to host country staff who have advanced statistical and methodological skills. In short, there are many instances when much larger surveys are needed to obtain the amount or type of data required. But even in these cases, small scale surveys can serve as a stop-gap measure until what is actually needed can be done. The deciding factor in opting for a small scale survey is whether it will provide data sufficient for the particular need at hand and whether it will be better than no data at all. When information needs are highly focussed, a survey guided by the rule of thumb "the smaller, the better" is an effective method for obtaining necessary data.

Another use of small scale surveys is to use a series of them to substitute for one large survey. Where a mission and host country might not be able to carry out a single large scale survey, the project can be divided into smaller segments which are amenable to the limited resources and manpower available for data collection. A series of small scale surveys can be spread out over two or more years and in the end, they will provide fairly comprehensive coverage. The surveys can be conducted in one region or district at a time so that host country staff can be used. As a result, the staff obtain

valuable field experience. Data processing can begin as soon as the first survey is completed and continue on as the series proceeds. This reduces the substantial demand created by one large survey. The trade-off is timeliness. The complete data set will not be available until the end of the series which could be one or more years after the initial start-up. Also the topic being studied must have a certain degree of stability or continuity so that conditions do not radically change between surveys. Where these factors are not problematic, a series of small scale surveys should be carefully considered by USAID missions planning a large scale data collection effort.

Another type of survey which could be of considerable value to AID is data collection at the community level. This type of survey has been conducted primarily for academic research; however, it could be readily adapted to more applied purposes. What distinguishes these surveys from others is that the community is the unit of analysis. This contrasts with more common surveys which select individual respondents to report on personal behavior or on the activities and conditions of a unit of organization to which they belong (e.g., a farm household, a small business, etc.). In community level studies, data are collected which pertain to the community at large, such as the availability of potable water, public health facilities, electricity, the distance to paved or all-weather roads, typical construction materials used for housing, the number and type of local business establishments, and

the presence of government offices and services. Much of this data can be obtained by direct observation and from interviews with local leaders (e.g., mayors, village headmen). A minimum of approximately fifty communities must be selected so that at least simple analyses of the data can be made. Beyond that minimum number, sample size is contingent on the desired coverage of the survey.

Community level data collection is generally cheaper to conduct, requires less manpower (i.e., fewer interviewers) and can be completed more quickly than individual based sample surveys. Moreover, they are readily amenable to tracking change over time, a common AID project objective. Sample surveys based on individual respondents encounter serious problems in locating the same people at two or more points in time. Communities, in contrast, are far more stable - they occasionally disappear entirely, but they do not move around unpredictably like individuals. Community level data collection also corresponds closely to a common objective of many AID projects. A phrase frequently found in project papers, particularly in reference to evaluation plans, is that "a process will be set in motion" whereby a positive impact on the well-being of the beneficiary population will be achieved. Measuring socio-economic processes and their development is one of the most difficult topics for quantitative investigation. Yet such processes are central to AID's development objectives. Often times the most that can be accomplished is to infer the process from the observable effects

assumed to result from it. A major impact of development processes is the improvement of the quality of life for an entire community. Therefore, the community is a very appropriate level for data collection and analysis to evaluate processes and their development impact. This does not mean that community level studies are "magic bullets" which will solve all of AID's information requirements. For one thing, they lack the very close detail necessary for certain types of studies, such as farming systems research or nutrition and health investigations. The purpose of emphasizing community level data collection is to draw attention to this approach in AID's work and encourage increased use of it where appropriate.

#### 4.1.2 Operations Data

Service or operations data collected and managed by the host country constitute an important alternative to sample surveys. Operations data pertain to the financial, managerial and service delivery functions of a central ministry and its field offices. At the field level, the data sometimes include extensive client histories. In many instances, individual field staff (e.g., a health worker, an extension agent) reporting on their performance of duties are the primary collectors of operations data. The data are typically reported by the lowest level unit (e.g., a rural health clinic, an agricultural extension office) to a district or regional office which, in turn, submits periodic reports based on this data to the central ministry.

AID projects designed to establish or improve ministry information systems usually have two main components. First, the reporting system is upgraded to improve the quality, coverage and timeliness of the data. Second, the analytic and planning uses of the data are strengthened to upgrade the monitoring and evaluation of ministry performance. USAID/Nepal and USAID/Zimbabwe have supported the introduction of microcomputers with necessary software into the central ministries. The computers have been instrumental in achieving significant improvements in the ministries where the machines were placed. USAID/Zimbabwe's Basic Education and Skills Training Project will try to further refine the system by decentralizing data use. That is, provincial offices will at some future date have the capacity to do budget modelling like the central office in the Ministry of Education and Culture. To monitor and evaluate service delivery and overall performance, methods of data analysis used for operations research are employed. The analyses typically focus on differentials in service delivery by geographic location, type of field unit, organizational or managerial system (e.g., integrated versus non-integrated health delivery) and staffing. The goal is to improve service delivery in a cost-effective manner.

There is much to recommend this approach to meeting mission and host country information needs. Such projects build the institutional capacity for improved operations and planning. They develop essential

administrative and analytic skills. In fact, there is considerable overlap between the methods and skills used for operations research and standard sample survey work. The data bases which are produced by these projects can be extremely useful. They can be used, for example, to identify where services should be improved and to monitor subsequent attempts to make necessary changes and the costs entailed. On the other hand, there are a number of reasons why more AID projects do not use operations data. First, many countries lack the financial and human resources to maintain such data bases. Consequently, existing systems are typically at a rudimentary stage of development or do not exist for certain sectors. Data collected through existing systems are often of very low quality and limited coverage (both in terms of content and geographic range). Consequently, it can take several years before data are available and adequate for analytic purposes. Clearly, if data are needed earlier on in the project, operations data will not be suitable and instead, a survey will probably be necessary.

One means of coping with the lack of data during the time it takes to get a ministry's information system up to speed is to conduct small scale sample surveys in the interim. In the initial two or three years, data needs could be met through a series of focussed surveys. These would later be phased out or integrated into the information system as it becomes operational.

## 4.2 Data Quality

A basic fact of life concerning the data related activities of USAID missions is that in most countries, available data and even data which are currently being collected does not meet rigorous standards of statistical validity. Moreover, this situation will continue in many LDC's for the foreseeable future. If a purist position concerning data quality were adopted, the majority of data collection efforts AID supports (with the exception of national population censuses conducted with direct assistance from BuCen staff) would probably never be attempted, or, alternatively, attempted at such a level of sophistication that they would be incommensurate with host country capabilities and/or impractical in the context of a development project. In other words, I am firmly convinced that strict adherence to purist standards for data quality would, in most instances, have a totally stifling, paralytic impact on the Agency's data related activities. The basis for this assertion is that this study has identified only a handful of information needs which must be met by statistically rigorous data and, therefore, warrant the expense and time entailed. Far more common are information needs which could be met with data of lower quality and more limited utility. This view is widely supported by information obtained from AID/Washington and USAID mission staff and reflects opinions expressed by mission staffers.

As mentioned earlier, flawed data can have utility for AID if the limitations and weaknesses of the data are recognized and not grossly exceeded. The internal validity of data - i.e., does a variable actually measure the specific thing it purports to measure - and reliability - i.e., does the variable provide an accurate measurement so that re-measurement would obtain the same value or response - are the primary considerations in assessing the quality of data. The answers are a matter of degree. The point at which data quality is so low that it should be used only with heavy qualifications or not at all depends on the purpose at hand. As unpalatable as this type of pragmatic, junkman's approach to data quality is to some, it has already guided, quite effectively I think, some of AID's data collection and analysis activities. The utility and necessity of this approach in large part is due to having <sup>to</sup> /choose between using flawed data or no data. In most instances the former is far preferable to the latter. The following suggests how flawed data should be used conservatively:

- 1) Identify the weaknesses and limitations of available data. If this is not possible, then the data should probably not be used.
- 2) Identify the causes of those limitations.
- 3) Use the available data as much or as far as possible without exceeding known limitations.
- 4) Explicitly state the qualifications which must be placed on findings or generalizations whenever the data are used.

5) In current and future data collection efforts, try to avoid repeating the problems identified above.

Trying to decide what level of precision or quality is necessary in upcoming projects is always difficult. What needs to be considered is the trade-off between cost and time on the one hand (meaning that quality can be bought), and what will be achieved or possible with higher quality data on the other (i.e., the payoff). Immediate utility - i.e., for what purposes are the data needed in the short-term - versus potential future uses - e.g., will the data collection effort be replicated at a later date, or will another project or a later evaluation use or re-analyze the data - is another key consideration. One point which should be kept in mind (and even treated as an eternal verity of data collection) is that a false sense of precision inevitably leads to too much data being collected. The project simply becomes bogged down with excessively long questionnaires, too many respondents, too frequent multiple round interviews, etc. False precision and too much data are the kiss of death to many of AID's data related activities.

#### 4.3 Representativeness

How representative data are of the true conditions of a population (e.g., the total country, a region, small businesses, farm workers, etc.) is closely tied to data quality considerations. The external validity of data - i.e., the degree of confidence which can be placed in inferences drawn from sample data about the study population - can determine whether available data are sufficient for specific information

needs. The central issues are 1) coverage - does the sample (or other method used to collect data) include key elements (e.g., geographic areas, ethnic groups, farming systems, age cohorts) pertinent to the problem at hand, and 2) selection - were cases selected according to a valid sampling technique and if not, what limitations does this impose on the data. Again, the answers to these questions should be viewed as a matter of degree. In general, AID should exercise greater flexibility toward using data which have a lower level of representativeness than what is required to meet scientific, statistical standards. Sample data which are not perfectly representative (they might be accurate and of high quality but not representative of the study population) should be used in the same fashion as lower quality data - conservatively without exceeding known limitations and, if that is not possible, then not at all.

One place where AID must always work with data of limited representativeness is in pilot studies. USAID/Mali provides a good example. The mission is currently developing a farming systems research project. The Ministry of Agriculture has proposed to purposively (as opposed to randomly) select two villages in the project area for pilot study sites. The common strategy followed for non-random selection of sites is to pick ones which are considered to be most indicative, typical or illustrative of the topic being investigated. By purposively choosing two or more villages, basic differences can be emphasized

out  
or brought/by the study. This is essentially an attempt to use a matched-pair design; cases are selected which are identical or very similar on important criteria (e.g., population size, distance to market centers, etc.), yet differ on one key variable (e.g., ethnic composition, size of land holding). The matching is intended to control for extraneous factors thereby strengthening the argument that observable differences (e.g., crop yield) are attributable to a key variable (e.g., ethnicity). In the Mali farming systems study, if it were thought that cultural systems strongly influenced farming practices and productivity, it would be important to pick two villages which allowed making comparisons between major cultural groups in the area. Alternatively, family structure is often a determinant of farm management practices. Again villages should be selected to make comparisons between traditional extended family structures and single adult generation households possible.

From a strict methodological point of view, purposive selection even with matching is a flawed design. First, there is no guarantee that the selected villages are indeed representative of any larger population as can be estimated when a truly random selection process is used. Matching is always problematic. Simply finding cases as complex as villages which are very similar except for particular key variables can be extremely difficult. Furthermore, there is no way of ascertaining whether all the important criteria for similarity

have in fact been included (setting aside the entire issue of interaction effects between these key criteria and the variables on which the cases differ). In short, under the best of circumstances, matching and carefully planned, purposive selection can improve the general validity of the effort but they are still a poor substitute for random selection of cases.

Should AID support this type of study? Absolutely. The purpose of the pilot study is to identify central problems and provide guidance for future investigations. Systematic data collection is clearly needed, but a more sophisticated and elaborate effort is simply not warranted for these purposes. Purposive selection based on the best available information and careful consideration of the type of factors mentioned above is preferable to selecting villages because they just happen to be where some ministry official's family lives or/other bogus reasons. The point is that despite the limited representativeness of the data, very useful and important insights into the mechanics of the problems further research and development of farming technologies must solve come to light.

#### 4.4 Amount of Data

The sad but true fact is that many well intended and much needed data collection and analysis efforts have been far less effective than they ought to have been or failed entirely simply because too much data was collected. The resources available for data collection -

time, money and manpower - get disproportionately consumed as a result of plowing through questionnaires which are much longer than good judgment would dictate. Alternatively, the malaise of spurious precision strikes when multiple round interviews are conducted until the project is virtually awash in unprocessed and unanalyzed data. This is no exaggeration, for example, in one project, data concerning all energy use for a six hour period one day a week and retrospective data about energy use during the preceding three days were collected every week for two hundred households for an entire year. This is a staggering amount of data to collect especially for an experimental, pilot project.

With very few exceptions, AID's approach to the amount of data to be collected should be "the less, the better". The absolute minimum amount which will meet the specific information needs of the project should be the rule. A possible exception might be where two separate surveys can be logically combined to reduce redundancy. This might require expanding the content of the single survey somewhat to accommodate all pertinent information needs. Another might be when a survey could include a limited number of key indicators collected by an earlier survey to allow for measuring change over time. But this type of data collection is not frequently part of AID's activities. A "bare bones" approach should also guide the content of information systems developed for host country ministries. A practical solution to limiting the amount of data collected is to insist on a page or time limit to the data

collection instrument (e.g., questionnaires, reporting forms).

#### 4.5 Frequency of Data Collection

There are certain topics or areas of investigation which require serial data collection spaced at fairly close intervals. The maternal child health care and family planning programs AID supports exemplify this situation. Multiple round surveys (i.e., the same individuals interviewed two or more times on the same topic) are necessary for such studies. However, with the exception of these special areas, AID should avoid multiple round surveys whenever possible. Such surveys are very expensive and often encounter serious problems in tracking and locating individuals selected at the outset of the study. A high sample mortality rate (i.e., losing cases from the original sample) can seriously jeopardize the validity of the findings. Consequently, wherever there is high population mobility, multiple round designs should not be funded. When multiple round surveys are necessary, USAID missions should insist on the following points:

- 1) the interval between interviews should be as long as possible without undermining the utility of the data (e.g., every four months instead of two, bi-monthly instead of weekly), and
- 2) the time period for data collection should be as short as possible (e.g., twelve months instead of eighteen).

These two demands will reduce the total costs as well as the amount of data collected. In short, AID's strategy toward the frequency

and amount of data to be collected should emphasize quality over quantity while ensuring that essential information needs are covered.

## 5. Project Versus Sector Level Data Collection and Analysis

The practicality and necessity of baseline - follow-up designs for evaluation purposes is discussed in this part of the report. It is argued that sector level data collection and analysis is a viable alternative or adjunct to AID's present project by project emphasis and investment in data collection and analysis.

### 5.1 The Baseline - Follow-up Design: Good Intentions Gone Awry

The exact dollar amount the Agency spends annually on data collection and analysis for evaluation purposes is not available; however, its central importance is very evident in the data related activities of the USAID missions selected for this study. In a majority of the projects reviewed, their data related components were for evaluation purposes. Moreover, many of these projects originally planned a baseline - follow-up design to measure change over time which resulted from the project. It is, therefore, not surprising to find that one of the most common problems affecting the on-going data related activities of these missions involves the collection and analysis of data at two or more points in time for project evaluation purposes.

In principle, the baseline - follow-up design should provide the type of data needed to evaluate the success of a project in achieving

project goals and purposes. The basic requirement of baseline - follow-up designs is to measure the condition, status or wellbeing of a beneficiary population before or at the very start of the project and again near or soon after the completion of the project. Changes and improvements observed over time are then attributed, at least in part, to the effectiveness of the project. Sample surveys are typically used to collect the necessary data. To improve the credibility of the results, experimental research designs involving control groups are often employed. It should be noted that a causal linkage between the project and measurable improvements is assumed in this design. In short, the baseline - follow-up design represents a potentially useful means for documenting AID's successes and failures. In theory, the utility of such information should help improve the design of future projects.

In practice, baseline -follow-up designs for evaluation purposes have proven to be much more difficult to implement and complete and of less utility than might be expected. The experience of the six missions used for this study indicates that USAID projects have had a low completion rate for baseline - follow-up designs. For example, in several on-going projects, the baseline data are so seriously flawed that a follow-up survey seems pointless. In other projects, baseline data were collected, but a follow-up is unlikely because of host country disinterest due to a change in government administrations. This

finding is consistent with the views expressed by mission and AID/Washington staff concerning the questionable necessity and possible inappropriateness of baseline -follow-up designs frequently incorporated into AID projects. The problems USAID missions confront in undertaking baseline -follow-up surveys is also supported by a recent AID/Washington review of the feasibility of using statistically rigorous sample survey methods for project evaluation. A major conclusion of the authors was:

"Rigorous multi-round sample surveys ... have not proven to be cost effective for purposes of project evaluation." (Cooley and Mазzie:1983. PPC/E/ESDS)

They also found that in the few cases where the baseline - follow-up design was completed, neither the host country nor the USAID mission made extensive use of the data, thereby minimizing its utility for evaluation and future project design purposes. Given what appears to be a growing concern about the suitability of using the baseline - follow-up design as frequently as AID currently attempts to do, a brief discussion of the main problems this study has found and some suggestions about how missions might avoid them in the future is warranted.

#### 5.1.1 Planning

From discussions with mission staff and review of project papers, it appears that too little formal planning is devoted to the design and implementation of baseline - follow-up surveys. One gets the impression from the project papers that the project designers recog-

nized the need for empirical, systematic evidence of project effects and impacts; specified that baseline and follow-up data would be collected, and then gave little or no thought to the practical considerations of how such plans can be put into action. In fact, evaluation plans involving multiple round data collection are sometimes as brief as a couple of paragraphs in some project papers. Data sources, host country capability for data collection, the need for technical assistance, USAID mission capability and specifically what types of effects and impacts can realistically be measured are frequently omitted. Such open-ended vagueness has been/undoing of several projects. The data collection component became a general fishing expedition which failed to satisfy the particular information needs of the project in some cases. Project managers confirm the fact that such poor planning does adversely affect data collection components of their projects. For example, adequate technical assistance was not available in several projects when it was needed. Decisions were made on an ad hoc basis as problems occurred by project staff who apparently did not have sufficient expertise. Obviously, this can jeopardize the entire effort. In short, there is no clear plan of analysis to guide data collection in a number of AID projects.

In sharp contrast to the negative examples, the Basic Education and Skills Training Program (BEST) supported by USAID/Zimbabwe is a perfect example of the type of planning all AID projects ought to

give to data collection and analysis especially when multiple round surveys are to be conducted. The PAAD for BEST contains the clearest statement of information requirements and methods appropriate for meeting those needs found in any AID document reviewed for this study. A plan of analysis is presented which specifies the key variables needed to evaluate program effectiveness. This level of clarity is not extravagant or unrealistic to expect from other planning documents. Rather, it constitutes sound planning on AID's part to better assure the success of data related activities.

#### 5.1.2 Design

Some projects planning to use a baseline -follow-up design have no control group (i.e., a basis of comparison for the project beneficiary group). Without a control group, it is not possible to determine whether an observed improvement is due to project inputs or to extraneous factors which could produce the same type of change. Control groups are also necessary when two different projects could have the same overall effect. In short, the utility of the baseline -follow-up design is diminished , in some cases, substantially, without including an appropriate control group. On the other hand, a control group increases the total sample size and the amount of data collection. If the project cannot afford additional data collection, perhaps an alternative less rigorous approach to project evaluation would be a better investment of time and money than a baseline- follow-up design.

### 5.1.3 Host Country Capability

Baseline - follow-up designs have suffered from a lack of support by the host country. In some cases, the host country was never genuinely committed to the necessity of collecting data at two or more points in time as planned. In other projects, a change in government resulted in a turnover in senior ministry officials. The new administration was not interested in more data collection even though a baseline survey had been conducted. In such circumstances, the USAID mission is put in the difficult position of choosing between sticking slavishly to the goals of institution-building (i.e., if the host country refuses to participate, then no follow-up is undertaken) or proceeding according to plan with or without the host country (i.e., using contractors to carry out the survey). If the follow-up survey is not done, then much, if not all of the utility of the baseline is lost. To proceed with minimal or no participation by the host country means little institution building and, therefore, no improvement of host country capacity for data related activities.

Obviously there are no simple answers and each mission must deal with the problem on a case by case basis. The best solution is to avoid the problem at the outset. That is, when mission staff recognize a decided lack of host country interest in multiple round data collection and analysis, alternative designs should be considered or the mission should be prepared to carry out the surveys more or less independently

if they are essential for project or program purposes. Once the mission gets as far as completing the baseline in active projects, the follow-up should probably be conducted if: 1) the baseline survey provides reasonably valid data (if not, then lack of host country interest is a good reason for avoiding throwing more good money after bad); 2) the surveys would establish an important data base which previously did not exist; 3) data collected have considerable future utility for project design, program planning, evaluation, etc.; 4) host country interest on the data might re-emerge at some later date; 5) the data can be used to monitor an entire sector in which the mission will continue to be active; and 6) adequate funds are available to contract for the follow-up survey.

#### 5.1.4 USAID Management

Baseline - follow-up designs have also encountered problems due to the administrative procedures and organization USAID missions must follow. First, the circulation of U.S. direct hire staff usually means that few ever see one project completely through from start to finish. The strengths and weaknesses of the baseline data, therefore, can be unknown to project managers overseeing the follow-up survey. The simplest solution to this is better documentation of AID funded data sets. Second, a common project goal is to "set in motion a process" which leads to a general improvement in the economic and social wellbeing of the beneficiary population. Even if projects are very

successful, such processes and the global impacts they are thought to produce might not emerge until years after the conclusion of the project. However, the follow-up survey has to be completed within the limited time-frame of the project. It is, therefore, very questionable as to what precisely can be measured so soon after project outputs have been completed. Third, some baseline - follow-up studies are not completed because funds were not available for the second survey. With shrinking AID/Washington and mission budgets, the considerable expense of statistically rigorous baseline - follow-up designs (costing \$250,000 to more than \$1 million depending on the scale and sophistication of the survey) constitutes a substantial investment for the Agency. It is likely that a number of forthcoming follow-up surveys will not be conducted as a result of insufficient funds. In one project, the funds for the follow-up were not available because the USAID office simply forgot to include the costs in its annual budget. These examples illustrate how the Agency's own internal organization and operation can interfere with the completion of baseline -follow-up designs. But the organizational structure and administrative procedures of the Agency are the "givens" in this situation to which designs for data collection and analysis must be accommodated and not vice versa. Therefore, in addition to other problems affecting baseline -follow-up studies in AID projects, this design might also be inappropriate from an administrative or managerial point of view.

### 5.1.5 Integrated Rural Development Projects

Meeting the substantial information needs of integrated rural development projects is very problematic for USAID missions. This study has not found one case where the baseline - follow-up design incorporated into these projects has been implemented and completed successfully. Because integrated development projects are very complex in terms of the range of activities included in them, various types of data from a number of different sectors and at different levels of socio-economic organization (e.g., regional, district, village) are required. Collecting and analyzing this data, in turn, involves a degree of coordination among the various host country ministries participating in the project which can be very difficult to achieve. Though this complexity is not intrinsic to baseline - follow-up designs, the need for longitudinal data to track various project effects typically necessitates the use of multiple round data collection. The situation is further complicated when the project is designed to be implemented in phases over four or five years where subsequent stages are to be guided by data generated in a preceding stage. Invariably, delays and implementation problems result in data not being available on time as planned. The subsequent phases proceed without the type of information project designers had anticipated. The amount of data collected for integrated rural development projects can be substantial especially where a series of surveys is to be conducted. The data processing

demands and skilled manpower requirements this creates can easily exceed host country capabilities. The problems USAID missions are encountering with integrated rural development projects using a baseline - follow-up design suggest that the amount of data planned to be collected is unrealistic given mission and host country capabilities.

#### 5.1.6 Short-term Effects Versus Long-term Impacts

A question which plagues AID's longitudinal data collection for project evaluation purposes is identifying appropriate indicators of project success. The problem is selecting indicators which should show improvements or changes logically resulting from project outputs and yet, at the same time, can reasonably be expected to occur within the time-frame of the project. This issue is central to scheduling the follow-up survey as well as the actual content of the questionnaires used. In the terminology of AID's Logical Framework, the question is whether it is reasonable to attempt to measure goal achievement within the life of a project. The alternative is to focus baseline - follow-up designs on project purposes and leave the issue of goal attainment to impact evaluations or other special types of studies. Based on the information obtained for this study, it is clear that the latter approach - focussing on project purposes and not goal attainment - should be AID's general strategy guiding evaluation oriented data collection and analysis for the majority of projects. If a later impact evaluation using quantitative data is anticipated, then special, and perhaps separate, plans for collecting this additional goal level

data should be made which will not overburden the project's more immediate and limited data collection and analysis activities.

Two water and sanitation projects illustrate the logic of focussing baseline -follow-up designs on short-term project effects rather than on project impacts. USAID/Egypt currently funds the Basic Village Services Project which supports any number of community implemented development activities. In discussing the data related component of BVS, the project manager cited a visit he made to a village where a piped water system was installed. Finding a typical "head of household" the project manager asked whether the family liked their new water system and why. The gentleman replied that he liked it very much because now his wife does not have to go to the public pipestand where she would be scrutinized by the men of the village. She could now remain indoors at home. Asked whether they used to water for washing, the project manager was told certainly not; the water is "too good" for that; washing is done in the irrigation canal. The villager took the project manager to the canal where his wife and other women were doing the wash. A short distance upstream was a dead donkey carcass someone had disposed of in the canal. As the project manager pointed out, baseline - follow-up surveys would certainly find that the number of households with piped water had radically increased as a result of the project. But any assumption that this indicated that the project had "set in motion a process" where family health practices and health status had been improved would overstate the project's impact at this point in time.

USAID/Honduras also funds a water and sanitation project which will provide improved sanitary facilities and relatively clean, but not potable, water supplies. In addition to the construction of facilities (which the villagers help build), a simple educational program is given which explains germ theory in layman's terms. This is premised on the idea that if the villagers understand the nature of water-borne diseases and how better sanitary practices will improve their own health as well as the health of the community at large, then it is much more likely that the new facilities will be used and properly maintained. The problem the project manager confronted was that a baseline - follow-up design was to be used to measure actual health improvements resulting from the project. The complexity and cost of such a study did not seem warranted for this project. Measuring morbidity carefully enough to distinguish between diseases which are symptomatically similar or identical would have been very difficult. Separating out the health impact of this project from other projects which would also affect health status (e.g., an infants' oral rehydration program) further complicated matters. In short, the project manager argued that health improvements would become measurable only at some later date after the project was completed. For evaluation purposes, data on project outputs and short-term effects, such as the type of facilities constructed, whether the education program reached adequate numbers of villagers, and whether the villagers were

using the facilities and maintaining them would be more realistic and within the capabilities of the Honduran ministry involved with the project. If AID were not willing to assume that access to and use of improved water supplies and sanitation facilities contributed to improved health status, then a special study should be conducted because such issues go beyond the limited scope and capabilities of the project.

Other examples of trying to force projects to collect data on development processes and their long-term impacts could be cited. I contend that these two water and sanitation projects illustrate the problem USAID missions currently confront concerning the type of data to be collected during the course of a project. Therefore, it appears that the Agency, perhaps unwittingly, is over-burdening projects and mission staff with demands for baseline - follow-up surveys for process or goal level measures of success which are unrealistic given the time limitations, and would be better obtained through some mechanism other than project evaluations.

## 5.2 Improving AID's Use of Baseline - Follow-up Designs

Given the difficulties USAID missions are having in implementing and completing baseline and follow-up surveys for project evaluation purposes, the Agency needs to re-consider its use of this design. First the design should be used more judiciously. Missions with medium or large programs can easily have several active projects which plan baseline - followup data collection. By reducing the use of this de-

sign, the demand placed on the missions and host countries for such data will be decreased. This should allow missions to concentrate their financial and managerial resources for data related activities on those projects which will collect baseline and follow-up data.

The experience of the missions selected for this study suggests that projects which should continue to use baseline - follow-up designs include:

- 1) projects specifically designed to test alternative service delivery systems and/or ministry operations and management systems;
- 2) experimental or pilot projects which test a new technology or technique which can logically be expected to produce measurable short-term effects;
- 3) institution-building projects which are to increase the data collection and analysis capacity of host country ministries for better management and planning (e.g., through establishing a monitoring and evaluation unit); and
- 4) a limited number of projects which can readily produce data pertaining to an entire sector central to the mission's program.

The third and fourth categories warrant particular attention for limiting the use of baseline - follow-up designs. Most of AID's data collection and analysis activities include an institution-building element which is, of course, a good thing. But it should be noted that the capacity of many LDC's for data collection and analysis

can be increased without developing longitudinal survey capabilities at the outset. An alternative approach would be to view baseline - follow-up surveys as a more sophisticated technique which can be incorporated after the host country develops basic methodological and statistical competency. In regard to program evaluation and information needs, most USAID missions will only have a few projects which can generate longitudinal, sector-wide data. In some countries, this will be the only comprehensive data base available for that sector. Given the potential utility of such data for a variety of purposes, the opportunity of obtaining such data should not be missed. In short, with the exception of the types of projects cited above, the Agency might be able to reduce its use of baseline - follow-up surveys without seriously handicapping project evaluation. A general strategy AID might employ to accomplish this reduction is to 1) require stronger justification for including baseline - follow-up designs in projects and 2) seriously question any assertion that it will be possible to measure goal attainment involving broad economic or social processes producing a long-term impact within the time-frame of the project.

### 5.3 Sector Level Analyses and Evaluation as an Alternative or Adjunct to AID's Project by Project Approach

The Agency has a very definite need for information about the long-term impact of the missions' programs and projects. If anything, this particular need is increasing as pressure on the Agency to

demonstrate its effectiveness has grown in recent years. The preceding section noted the difficulties USAID missions are encountering with trying to collect goal level, impact data through baseline - follow-up designs on a project by project basis. It was suggested that project level data collection focus more on the purpose level and attempt to measure near-term effects which can reasonably <sup>be</sup>/to expected to occur by the conclusion (or shortly thereafter) of the project. What AID needs to assess its success at setting in motion development processes which produce long-term improvements in the wellbeing of the beneficiary population is a data collection and analysis strategy better suited to measuring change at this level.

One possible strategy AID should consider is supporting periodic ~~sector~~ level evaluations. Each mission would periodically (e.g., every three to seven years) evaluate progress made in overcoming key constraints to sector development in the geographic areas where its program has concentrated. The criteria for assessing progress would be quantifiable objectives identified by the mission as central to its program and project goals. Sector level evaluations would be like AID's current sector assessments in that they would be useful to the missions for planning and strategy development. They would differ in that the primary focus of the evaluations would be on the effectiveness of the missions program and projects and, therefore, be much briefer and narrower in scope than sector assessments.

A sector level approach to data collection and analysis is already used for evaluation of sector funded programs. In discussing the implications of sector funding for USAID/Egypt's information requirements, Richard Seifman noted that quantifiable objectives of sector-wide improvements will be necessary:

"In energy, a target might be identified based on the proportion of consumers utilizing less than one kilowatt per month, increased kerosene sales; in industry, the percentage of industrial production jobs or investment in the private sector compared to the public sector might be an indicator."

(from A Report on Implementation Management, 1981:14)

Comparable indicators could be suggested for agriculture, nutrition, education - in short, for any subject area where the mission's program is concentrated. Mr. Seifman notes a major advantage of the sector approach when he states:

"...if we took those sectors where we direct most of our resources and have policy objectives and looked at appropriate indicators, it would give a better picture of what is happening in development terms" (ibid.)

There would be additional advantages to sector level analyses of program and project effectiveness. First, sector level evaluations

correspond closely to the general goals of a mission's development activities. AID's current heavy investment in data collection at the project level permits only a piecemeal assessment of mission accomplishments. Large-scale or long-term impacts resulting from overcoming sector constraints cannot be adequately evaluated within the context of most projects and at best can only be inferred through AID's project by project approach. Moreover, projects within a given sector often have similar and reinforcing goals. It would be far more parsimonious and logically consistent to attempt to measure their collective impact on sector development than do so separately one at a time. Also projects have effects which transcend the artificial sectoral boundaries we impose on the real world. For example, agricultural projects frequently increase household food supply which, in turn, has beneficial effects on family health. Sector evaluations would capture such spin-off effects whereas the individual project approach is prone to missing them.

requirement for sector  
A / level analyses would also provide greater impetus to collect data. That is, missions would have to obtain the types of data needed for empirically valid evaluations of program impact. For the most part, the project level focus of AID's data collection efforts is presently not providing the necessary data for such assessments. As the authors of a USAID/Egypt document prepared for the Decentralization Support Fund note:

"...despite the fact that project agreements often call for development of the baseline data and analysis necessary to advance development of larger sector-related policies and procedures, such efforts consistently take a back seat to other, more immediate pressing demands. Regrettably, they are often treated by GOE counterparts and contractors as peripheral to or ranging beyond the practical limits and prime purposes of the discrete project with which they are working..."

Unfortunately, this problem is not restricted to Egypt and is quite common in many other countries. But if missions were charged with demonstrating their success in advancing "sector-related policies and procedures", the back seat status of data collection for these purposes would quickly end.

Sector level evaluations would not replace completely data collection and analysis for all projects. Quantitative data to evaluate the accomplishment of project purposes would still be necessary, but this might be possible using one end-of-project survey instead of a more rigorous baseline - follow-up design. As noted earlier, certain types of projects will continue to require longitudinal data collection regardless of whether the mission conducts sector level analyses or not. However, some reduction in data collection

for individual projects is very likely. Sector analyses could substitute for extensive data collection in some projects and serve as an adjunct to the data related components of other projects.

Any reduction of data collection and analysis within projects resulting from sector level evaluations would be viewed positively by project staff. The majority of USAID mission staff - both U.S. direct hire and foreign nationals - have substantive, managerial skills, but not technical, analytic skills pertinent to data related activities. Yet project staff frequently confront technical, analytic problems they are ill-equipped to resolve. Non-technicians dealing with technical problems is nothing short of a disaster looking for someplace to happen. This study has found examples of data related activities which have failed or suffered because of poor decisions made by USAID staff who lacked analytic or data management expertise. Sector level evaluations might avoid some of these problems by reducing the responsibility or involvement of project staff in data collection and analysis.

One possible objection to sector level evaluations concerns the fundamental logic of evaluating project effectiveness on the basis of sector level data. For example, assume that data show a significant improvement in agricultural productivity, health status and household income over a ten year period during which the USAID mission supported agricultural and health projects. If that were all

the information available, from a strict methodological point of view, the conclusion that the observed improvements are totally or partially the result of the USAID projects could not be sustained. The type of information necessary to support such a conclusion would include:

- 1) longitudinal data for a ten year period or so with annual measurements which show a significant improvement occurring only after the implementation; and/or
- 2) regional breakdowns which show that in the regions where the projects were implemented, improvements occurred more rapidly, continuously and/or for a longer period of time compared to regions which were not project areas.

For the majority of LDC's, this type of data simply does not exist, particularly extended time series data, and would have to be collected over time. But even with such data, from a strict methodological point of view, there is still something of a leap of faith in concluding observed improvements are actually the result of the projects. Further refinements and more data would be needed to eliminate additional alternative explanations of the improvement.

Such possible criticisms are indeed valid in terms of rigorous textbook standards for demonstrating the effects of an intervention (e.g., a development project). But what would the hardline purists have AID do? Continue to spend even greater amounts, project by

project on baseline - follow-up designs? And continue to do so even though there is growing evidence that these designs are unworkable and ill-suited for AID's mode of operation? The Agency needs to take a hard-nosed, pragmatic approach to the issue. First, the possibility of erroneously assuming a linkage between project inputs and an observed improvement at some later date is no greater than that involved with alternative, project level approaches to goal and impact achievement. At least sector level analyses would be logically consistent with the types of development improvements AID is trying to effect. Second, USAID missions have neither the time nor the resources to generate data with sufficient precision and comprehensiveness to satisfy unrealistic, purist standards for rigorous evidence of cause and effect. Sector level evaluations which provide some evidence, albeit limited, linking observable improvements to USAID efforts in those areas would be far preferable to muddling through without such empirical support or, worse yet, ignoring the issue of goal attainment and development impact entirely.

The final points to be made concern practical considerations of conducting sector level analyses. In general, a flexible view should be taken toward the format, content and timing of the evaluations and analyses. Obviously, the sector goals of the mission's program should guide the selection of indicators to monitor program and project effectiveness. Data collection through selected projects,

host country operations and special studies could become an on-going process tied to other mission and host country development activities. The frequency of the analyses should be negotiated among AID/Washington, the mission and the host country. Most if not all missions will require technical assistance. The host country would, of course, participate according to its capabilities and interests. Sector analyses could also serve as an important vehicle for the mission's institution-building efforts to improve host country data collection and analysis capabilities. The data bases would serve as a common ground for continuing mission dialogues with host country counterparts concerning policy reform as well.

A major purpose of sector level analyses is to improve the return on AID's investment in data related activities. As mentioned above, sector level analyses could in the long run reduce the amount of data that AID currently plans to collect in coming years. If such reductions are made, funds which would have been spent for project level data collection could be used for sector level evaluations. Some missions might even fund and manage sector level analyses as separate projects. In other words, missions should design and undertake sector level analyses which are best suited to their particular programs, resources, capabilities, etc. as long as the final product is an empirically based assessment of progress toward goal attainment. The data for these analyses should be as high a quality as is

possible given available resources. Obtaining valid and reliable data for sector level evaluations should, therefore, be given high priority in the mission's data related activities.

#### 6. Recommendations for Improving AID's Current Data Related Activities

The report now turns from the larger issues affecting AID's data related activities to specific actions for AID/Washington and USAID missions for better in-house data use and for more effective institution-building for host country data collection and analysis. AID/Washington should develop data support services to assist the missions and secondarily, AID/Washington offices. In AID/Washington, a data support division within one of the central bureaus - PPC or S&T - should be established to provide these services. The regional bureaus should develop regional data support centers based in selected USAID missions to provide technical assistance to missions for data related activities. The following actions could contribute to improving the data related activities of USAID missions:

- 1) support stand-alone institution-building projects to expand host country capability for data collection and analysis in each sector where its program is concentrated;
- 2) use local contracting firms as much as possible, even as a substitute for government staff, to build an in-country capability for data collection and analysis in development projects;

- 3) in general, data collection which has utility for project design warrants greater emphasis than it receives at present;
- 4) if possible, use qualified mission staff as an in-house source of technical assistance for data collection, analysis and management;
- 5) strive for better coordination among projects within the same sector to reduce redundant data collection as well as to generate data which will facilitate sector level evaluations; and
- 6) coordinate the mission's data collection and institution-building efforts as much as possible with similar activities funded by other international development agencies active in the country.

In a preceding report which discussed how AID/Washington could improve its management and use of quantitative data, I noted that analytic support services had fallen through the cracks of AID's organization. Almost two years later, little if any improvement in this situation has occurred. There is no single office for AID/Washington or USAID missions to turn to for assistance or advice about data collection and analysis. A most laudable improvement has been made concerning support to the missions for acquiring microcomputers with the establishment of SER/IRM. But their responsibilities only include hardware and software acquisition for AID use. For some unfathomable reason, what goes into the machines - i.e., the data the missions must work with - and how to maximize the Agency's investment in data related activities continues to be

ignored. PPC/E/ESDS has taken steps toward developing a sub-national data bank and improving contract language to guarantee that AID receives copies of properly documented data sets. However, the extremely limited budget and staff capability of ESDS has kept these efforts at a very rudimentary stage. In short, the Agency has done too little and moved too slowly in this area. Lack of funding and lack of staff is perfectly indicative of the apparent disinterest senior management continues to display toward developing systems for adequate information use in AID/Washington and the missions. To correct this situation, the Agency should develop data support services in Washington and in the field.

## 6.1 AID/Washington

### 6.1.1 A Data Support Division

The Agency should establish a division within one of the central bureaus, either PPC or S&T, charged with the responsibility of supporting AID's data related activities. Staffed with technically competent personnel, the division's primary focus should be on assisting USAID missions with data related activities. This assistance could be provided through TDY's to the missions and by backstopping services in AID/Washington. The missions could request assistance directly or through the Technical Resource Offices of the regional bureaus. The services the division should provide include the following:

- 1) Review all proposals for data collection and analysis to assure that proposed activities are appropriate for the projects and the host country, adequately designed and sound in regard to funding, staffing, data processing requirements and technical assistance.
- 2) Through TDY's to the missions, assist with the design of data collection and analysis components of future projects, and help to trouble-shoot problems which arise in on-going data related activities. If computer equipment or programmers are needed, the division would work collaboratively with SER/IRM.
- 3) Assist regional bureaus and missions with the writing of scopes of work for data related activities and the selection of suppliers of technical assistance to better assure competent individuals are giving sound, practical advice to the missions. (See Annex C for a further discussion of selecting sources of technical assistance for data related activities,)
- 4) Oversee the management of a sub-national data bank containing data sets produced through AID funded projects. The initial steps toward this objective have been started by PPC/E/ESDS and should be expanded and expedited. The division should assist missions and host countries at the contracting stage of projects to guarantee that they as well as AID/Washington receive usable properly documented data sets from contractors on a timely basis. The division would have the capability to perform special analyses of the data on re-

quest from the missions or AID/Washington offices. In this capacity, the division would constitute AID's in-house analytic unit.

5) Develop standards and oversee their use for data related activities USAID missions frequently fund. Such standards would be developed with the direct participation of AID/Washington offices in each of the functional areas (e.g., agriculture, population, education, etc.) and with outside assistance from appropriate U.S. government agencies (e.g., BuCen, USDA, BLS).

6) The division should be responsible for monitoring and evaluating AID's institution-building efforts designed to expand host country capability for data related activities. The division would then serve as a repository of expertise on this type of institution-building and make this expertise available to any part of the Agency on request. A monitoring system which tracks AID projects which include institution-building for better data use would be very useful for capturing valuable "lessons learned" from past projects to be incorporated into future projects.

This list is only indicative of the type of services the Agency needs to develop to provide adequate support for data related activities.

#### 6.1.2 Regional Support Centers

In all probability, if a data support division were established in AID/Washington, it would have limited capability for working directly with mission staff on a TDY basis. However, it is apparent

from this study that such assistance is needed by the missions. Moreover, with current budget constraints and cutbacks, it will be difficult to significantly upgrade the in-house capability of all USAID missions for data related activities. In particular not all missions could justify in terms of time, costs and workload having in-house expertise for data collection and analysis. Though in-house expertise is important to mission capability, a good alternative would be providing assistance through regional centers. The primary function of the centers would be to provide technical assistance for all aspects of data related activities directly to the missions. In each major geographic region, a centrally located mission would be selected as the home base for the regional center. Staff would be available to the missions located in the region. At least one U.S. direct hire staffer would probably be required for management purposes, but technically competent foreign nationals should make up the core of the centers' staff. This would lend continuity to the operations of the centers. Staff size for each center would depend on the number and size of the missions in the region. Costs would be shared among the missions. The regional bureaus could have oversight responsibilities and should contribute to funding. The centers, however, would be directly accountable to the missions in terms of the services provided and the evaluation of staff performance. It should be added that many staff interviewed for this study were

very supportive of this idea and saw a definite utility in the services regional centers could provide. An additional advantage of regional centers would be that their staff would soon know of all on-going data related activities in the region. Having seen what worked and what failed, their advice could be invaluable to missions in the region.

### 6.1.3 PPC

PPC's guidance provides important impetus for improving the data related activities of USAID missions. Accordingly, PPC should assist the missions produce analytically sound work as follows:

- 1) Identify the information needs associated with the basic tasks of program offices and develop replicable formats for analyzing and presenting such information. These models should be developed using off-the-shelf software for Apple and IBM microcomputers.
- 2) Clarify the types of sub-national, socio-economic analyses missions need to perform to improve program planning and evaluation for their CDSS's.
- 3) Require missions to include an information strategy statement (one or two pages as an annex) in their ABS which details current and upcoming information needs of the mission and plans for how these requirements will be met. The purpose of the statement is to force mission management to give more attention to the necessity of obtaining basic data. Strategy statements of this sort would also

provide a very convenient means for monitoring mission progress in improving data bases, in-house data use and support to the host country for data related activities.

4) Guidance concerning the development and approval of plans for data collection and analysis to meet project information needs should be added to Handbook Three. PID's should clearly specify data required for the project paper. A clear and workable plan of analysis which states what data will be collected, the method of data collection and/or data sources to be used, and an outline of analyses to be performed should be presented in the project paper.

5) PPC should develop an information policy determination which clarifies AID's commitment to matching its information needs to the appropriate data and the importance of developing host country capability for data collection and analysis through technology transfer appropriate for the country's specific needs and resources. The policy determination should involve all relevant offices in the regional and central bureaus. The purpose of the policy is to support improvement of data related activities throughout the Agency and to draw the attention of senior management to the necessity of providing adequate funding and staffing to carry out these activities.

## 6.2 USAID Missions

Improving the data related activities of the USAID missions will largely be a matter of working through the principal factors determining mission capability discussed in Part Three of this report. However, there are several strategies that managers could

employ as far as their resources permit which could improve their effectiveness in this area.

#### 6.2.1 Stand-alone Institution-Building Projects

The importance of developing host country capability for data collection and analysis within the limits of available human and financial resources has been stressed throughout this report. This objective would most likely be acknowledged as important by the vast majority of AID/Washington and USAID mission staff. This study has found that missions have incorporated this goal into many of their projects involving data collection and analysis. However, there is also evidence of the data related components being eliminated when project implementation problems arise. Obviously, this minimizes improvements made in host country capability.

A more direct and perhaps more effective approach to institution-building is to fund projects which have the sole purpose of expanding host country capacity for data related activities. Instead of embedding this objective in some larger project, missions should consider funding at least one project in each sector where its program is active which is designed to expand the institutional capacity of the host country ministries for data collection and analysis. Instead of diffusing this goal among several projects, one stand-alone project might be more effective in achieving improvements. Funding which would otherwise be used for comparable purposes in separate projects would be pooled to fund this single institution-

building project. Moreover, such projects are consistent with the Agency's emphasis on maintaining discussions with host countries concerning the need for policy reform. That is, the data and analyses generated could serve as a basis for continuing policy dialogues. In fact, the policy and planning projects AID currently funds have essentially the same institution-building effect in terms of strengthening host country capability for data related activities. The major difference between present policy and planning projects and what is proposed here as stand-alone institution-building projects is that the latter broadens the area of possible AID involvement in assisting LDC's with data collection and analysis beyond policy oriented objectives.

A perfect example of this type of project is the Data Collection and Analysis Project which USAID/Egypt is supporting to assist the Ministry of Agriculture to improve its planning capability. The project

addresses a set of problems common to many LDC's:

"...(1) insufficient economic data is being collected, analyzed and fed into the decision-making process; (2) the capacity to utilize whatever data and analysis are available is not sufficiently developed; and (3) links which integrate the research and analysis process into the decision-making process regarding resource allocation, are weak/missing"

(Project Paper, 1980:2-3)

Though this particular project is concerned with agricultural data, it illustrates the basic purposes and objectives of stand-alone institution-building projects:

"The proposed project will help the Ministry of Agriculture to overcome data collection and analysis problems that hinder rational decision-making in agricultural planning and policy formulation. This will be accomplished over a five year period through assistance to improve the collection, compilation and storage of basic agricultural data and to increase the amount of analytic work performed, as well as the use of such information in policy development and planning. In the area of agricultural statistics and data collection, improvement is to be achieved through the provision of

of short-term technical assistance on a regular basis, modest commodity inputs, considerable training and limited amounts of funding for special data collection activities" (ibid., 1)

This description could be used as a basic template; simply substitute the appropriate country name, sector and ministry.

A prerequisite for these projects is genuine support by senior ministry officials; otherwise, the effectiveness of the project will be seriously undercut. Where the data touch on politically sensitive matters, stating the results of analyses in terms of options and their consequences will communicate the main findings to ministry officials while leaving room for maneuver. The effectiveness of these projects also depends heavily on focussing data collection and analysis on problems or tasks identified as important by ministry staff. The output of the project should then demonstrate how better data use facilitates ministry operations. Finally, personnel changes in the technical assistance team should be kept to a minimum. This will contribute to continuity in project activities between TDY's and build upon any rapport established between the contract team and ministry staff.

More variable characteristics of these projects include the size of the technical assistance team, the amount of assistance provided, the number of host country staff trained and the amount of

training they receive, the type and amount of equipment, and outputs such as reports and special studies. There is considerable flexibility in all of these elements. The amount of funds available, current ministry capabilities, computer hardware needs and other aspects of the particular situation will determine what is necessary and possible.

#### 6.2.2 Local Contracting Firms

An alternative approach to expanding host country capability for data collection and analysis is to encourage local consulting firms to develop their capacity for such work.

AID's institution-building efforts in some countries are thwarted to a certain degree by the inability of the government to retain skilled staff. A common pattern is for ministry staff who receive technical training or advanced education to leave government service for better paying jobs in the private sector. Quantitative, computer oriented skills are definitely in high demand. Consequently, it is very difficult to develop the institutional capacity of some countries for better data collection and analysis.

Rather than interpreting this situation as a shortcoming of AID's institution-building efforts, an alternative view is that as long as the trained worker remains in the country, the objective

of building in-country capability, which is not necessarily tied to host government capability, has been accomplished. In countries where the government is incapable of retaining staff necessary to meet all of its information needs, USAID missions should make greater use of local contracting firms as a substitute for government staff. In many countries, private firms exist which are able to collect and analyze data. Their competency varies widely from country to country, but a number of these firms do carry out market research studies which require the same basic skills necessary for survey work in development projects. With technical assistance from U.S. contractors or U.S. government agencies, such as BuCen/SEU, the local firms could adapt and develop their methodological and statistical skills to the needs of USAID funded projects. The impetus for taking on this work would be that USAID funding would pay considerably more than if the government were the sole contracting agent. Though perhaps not as desirable as developing the capacity of the host government, working through local consulting firms at least creates an in-country capability and makes use of former government workers who have gone to work for these firms. At the same time, AID would be strengthening a local job market which is consistent with the Agency's private sector initiatives.

### 6.2.3 Data for Project Design

This study has found relatively few examples of data collection specifically for project design needs. Theoretically, a major reason AID supports data collection and analysis is to obtain information which will help missions design more effective projects. Apparently, there is room for such improvements according to some mission staffers. As one individual with considerable experience observed, the view that AID projects are well designed is directly contradicted by the number of projects which encounter serious implementation problems at their outset and which continue to occur throughout the life of the project. Regrettably, one finds support for this opinion in heavily funded projects which ran aground in part because they were built on incomplete and faulty information. No stronger argument to justify the costs of data collection and analysis can be offered than to compare those costs to what is wasted as a result of not having adequate information. The obvious solution is for missions to give greater emphasis to data collection which will have utility supporting for project design. This will require/pre-project, exploratory studies, limited in scale and highly focussed on design questions. Second, more careful attention should be given to data collection in on-going projects to obtain information from these activities which will be useful for designing upcoming projects in the same sector.

#### 6.2.4 In-house Support for Data Related Activities

The single most important means for improving a mission's data related activities is to have at least one staffer who has statistical, analytic skills and the worktime to allow this individual to function as an in-house source of technical assistance to the mission. The experience of USAID/Honduras illustrates this point. One staffer who had quantitative skills and a keen interest in improving the mission's information base contributed significantly to the progress USAID/Honduras has made in this area. The mission still benefits from the analyses which were completed years ago with the assistance of this person.

Even in large missions, however, it would be difficult to justify a staff person assigned exclusively to data related activities. Rather, assistance to other mission staff and host country counterparts for data collection and analysis could be made one of a number of duties assigned to a staffer. In other words, this person would be available to work on any of the mission's data related activities as needed. Project managers might not be the best choice even if they have skills because they tend to be so engrossed in their particular projects (however, in USAID/Honduras, the person was a project manager). Project design, evaluation or program office positions might be better suited for providing this type of assistance. In short, who provides the assistance or oversight for the mission's

data related activities is less important than somehow finding the worktime for qualified staff to perform the task.

#### 6.2.5 Coordination of Data Related Activities within the Mission

In the larger missions, special attention should be given to coordinating data collection and analysis among projects to reduce redundancy and to generate data bases which can be used to evaluate sector level goal attainment. In the extreme case, USAID/Egypt has had some difficulty coordinating data collection among projects within the same sector. The experience of the Agriculture Office illustrates the problem. The Office currently (as of 9/82) manages one dozen projects which constitute a total funding obligation of approximately \$225 million. Several of these projects are very large efforts to increase the productivity of Egyptian agriculture. In this regard, they share broad goals for agricultural development and collectively should produce mutually reinforcing benefits within the sector. This suggests that a data base which would monitor the overall impact of USAID/Egypt's agriculture program is needed. As obvious as this might seem, the current office director explained that the individual projects had not been designed to generate data which would pertain directly to general sector level issues. To correct this situation, the office is now incorporating such concerns into on-going and future projects.

Certainly the size of USAID/Egypt's program makes it very difficult to coordinate data collection and analysis among projects managed by different offices even though there are some common information needs. But even in missions like USAID/Honduras (i.e., a comparatively large program for a small population), the problem of inter-office and inter-sector coordination also occurs.

Planning such coordination at the project design stage is the only practical solution to the problem. That will require identifying the common sector level goals the projects will jointly achieve and then remaining cognizant of on-going and upcoming projects which have data related components to dovetail or piggy-back data collection and analysis among these projects. Admittedly, this is easier said than done within the context of mission operations. But trying to impose broader sector issues on projects after the fact is a far less preferable route for creating this type of coordination.

#### 6.2.6 Coordination Between USAID Missions and Other International Development Agencies

This study has found very few examples of a USAID mission working cooperatively with another international development agency to collect and analyze quantitative data. The infrequency of such coordinated efforts is probably reflective of the independent operation necessary for USAID missions and to the specific information needs of USAID projects. But opportunities for data collection and analysis

do occur, particularly when AID and other donors are active in the same region and support related or reinforcing projects. Donors also share common information needs for data about general economic and social conditions. A proposed social indicators study for the Sahel in which AID would be one of the participants is one such example of overlapping information needs shared by the international donor community. A situation conducive to coordinated data collection is when a specific section of a country is the target of various projects funded by AID and other donors. The Rapti Zone in Nepal exemplifies this situation. A common project monitoring system was developed for the Government of Nepal to track the various activities being funded by USAID, ODA, CIDA and the German Government in the Rapti Zone.

A different approach to coordinating data related activities would be where each donor takes responsibility for developing data bases in different sectors and/or regions which are then made available to other international development agencies. This strategy could be very useful for the poorest countries, like Mali, which cannot afford to collect all the data they need and must rely on international development agencies to support data collection. For example, USAID/Mali, whose program is heavily concentrated in agricultural development in a limited number of regions, would be responsible for developing agricultural statistics over the next ten or twenty years,

while the French, WHO, etc. would have responsibility for areas designated to them according to the types of programs they support. Of course, this assumes a level of continuity and cooperation among development agencies which might be unrealistic in some countries. But if such coordination could be achieved, this might prove to be a very cost effective means for meeting basic information requirements.

## Annex A - Questions Used to Guide Staff Interviews

Staff interviews conducted in the six selected missions provided the basic information used for this study. A limited number of AID contractors and host country staff were also interviewed. The current data related activities of each mission were discussed in individual trip reports. This final report draws from and synthesizes much of the information contained in the trip reports. The basic procedure followed in each mission was to conduct as many interviews as time permitted, giving special attention to offices and projects most involved with data collection and analysis. The following questions were given to mission staff and others prior to or at the beginning of the interview. The purpose of the questions was to clarify what type of information was being sought and to give some direction to the course of the interview. However, the interviews were not highly structured in the sense that only these questions could be discussed. Moreover, all of the questions did not apply to everyone interviewed - e.g., some questions pertain to project activities while others concern program planning. The purpose of conducting loosely structured interviews was to cover a broad spectrum of opinion and experience concerning AID's data related activities. The following is the handout given to those interviewed.

PPC is interested in developing an information policy for the Agency. Current procedures and policies concerning quantitative data collection, analysis and management in AID/Washington have been reviewed. To be

effective, the information policy must reflect the needs and capabilities of USAID missions to use quantitative data effectively. It is necessary, therefore, to obtain information about current projects which will collect and analyze data and the mission's use of data and analysis as a guide to program and strategy development.

1. I would like a brief description of project(s) which involve the use of quantitative data for (a) project design, (b) project monitoring and (c) project evaluation. A copy of the project paper would be useful.
2. Does the project attempt to build the capacity of a government ministry or agency to use data for better planning? How successful do you expect this to be?
3. What problems has the project encountered in the collection and analysis of data?
4. In regard to analytic and technical skills necessary for data use:  
(a) how would you describe your own skills, (b) those of your co-workers and (c) those of your host country counterparts?
5. In regard to data obtained from the host country:  
(a) do you generally have easy access to available data?  
(b) are available data of reasonably good quality - valid, timely, and reliable?  
(c) are data exchanged on a cooperative basis between your office and host country ministries?  
(d) which data bases are strongest and which are weakest?

129  
~~A-3~~

6. What data bases are used for program purposes such as the CDSS and other mission level activities? Are the data sufficient and of acceptable quality for these purposes?

I want to shift the focus here to a discussion of possible improvements the Agency could make to improve data use by the missions.

7. How useful do you think an information strategy statement in the CDSS would be?

8. Inadequate data and limited staff skills are common problems affecting data use in the missions. But even analytically trained or experienced staff are hard pressed to find sufficient work time to engage in applied analysis even in a supervisory capacity. What do you think of developing regional centers (located in selected missions) which would be totally service oriented toward mission needs and whose staff would assist missions with data related activities?

10. Would it handicap or assist the mission in designing sounder projects which involve data collection and analysis if all such proposals were systematically and thoroughly reviewed by individuals competent in research methodology before the project was implemented?

What if these reviews were performed by AID/Washington?

11. What do you think of developing an in-house data bank (computer based on a Wang mini- or on a microcomputer) containing data sets generated by USAID projects or obtained from the host country which mission staff consider relevant to their information needs?

12. Would quality standards for data related activities help improve the performance of project contractors?
13. What other ways might AID/Washington assist the mission to improve and expand its use of quantitative data?
14. What would be the best and the worst things you could imagine resulting from an information policy for the Agency?

## Annex B - Ranking USAID Mission Capacity for Data Related Activities

The underlying assumptions of this ranking system should be made clear. First, each of the twelve factors is treated as though they are of equal importance both within and among missions and host countries. This assumption may or may not be valid. For example, one factor (e.g., senior management support) might be more important than the others, or the importance of a factor might vary among missions and countries. Whether such variation exists and if so, whether it is significant cannot be determined here. This possibility simply has to be accepted as a potential limitation of the analysis. Second, simply ranking missions and host countries says little if anything about the actual magnitude of differences between cases. To determine this, more refined measures would be needed. For example, mission size might be better expressed on a per capita basis - e.g., mission funding to host country population. Staff skill level could be expressed as the percentage of mission staffers with advanced training in fields which included quantitative analytic methods. The availability of ADP equipment could be reported in terms of the dollar value of current mission systems. Senior management support might be expressed in terms of the amount of work time specifically allocated to data related activities. The availability of data, however, is not amenable to other than ordinal measurement. Indicators of this sort could conceivably be constructed, but this too goes beyond the limits of this study.

It would also be useful to estimate actual performance or achievement concerning data related activities. This could be done retrospectively - i.e., projects which are now nearing completion or have been recently completed - or at a future date by reviewing the progress of the projects currently underway in the six missions. It would be important to determine whether the activity was completed as planned; whether the data contributed to mission and host country operations; whether the activity successfully expanded host country capability; and whether the techniques and systems developed by the project are likely to be institutionalized or otherwise sustained by the host country. This type of information and the factors contributing to success or failure could be very useful for planning future projects which build the institutional capacity of the host country for data collection and analysis.

## Annex C - Technical Assistance for Data Related Activities

For the majority of the data collection and analysis efforts AID supports, both substantive and technical expertise is needed. As obvious as this might be, several projects reviewed for this study had encountered major problems with data collection and analysis because the appropriate technical assistance was not available when it was needed. Consequently, the planning and/or implementation of the data related component of these projects were adversely affected. In at least one project, it appeared that as a result of inadequate technical assistance, the baseline survey which was conducted was so seriously flawed that the data are, for all intents and purposes, unusable. The obvious solution is to improve planning for data collection and analysis and to keep expectations for host country performance of data related activities within the bounds of the country's capability.

To improve planning and implementation, missions need to determine what type of technical assistance is needed and when it will be needed throughout the course of a project.. Though the specific requirements for technical assistance for data related activities varies from project to project, the data component can be divided into four general stages: 1) identification of information needs; 2) planning and design; 3) data collection and processing; and 4) analysis and interpretation. These stages are, of course, highly interrelated. However, substantive expertise is most important for stage 1 - identification of information needs - and stage 4 - analysis and inter-

pretation. Technical expertise, on the other hand, is essential for stage 2 - planning and design - and stage 3 - data collection and processing. Ideally, a single supplier of technical assistance can field staff who can provide both types of expertise. When that is not possible, the most effective means of procuring assistance for data related activities is to combine sources of substantive expertise with sources of technical expertise. For example, local consulting firms might be used for substantive requirements while BuCen/SEU provides technical support services. When different suppliers of assistance are combined in this way, responsibility for the data collection and analysis activity should be divided accordingly. Substantive experts should have primary responsibility for stages 1 and 4; technical experts should have primary responsibility for stages 2 and 3. USAID and host country staff would coordinate these activities and the staff involved as part of their overall project management responsibilities.

In regard to the procurement of substantive and technical services for data related activities, the Agency will probably obtain the best assistance when the specific needs of the project fit perfectly with the capabilities of an established supplier. That is, BuCen is the obvious source of assistance for census operations; USDA excels at area frame sampling; and private consulting firms, such as Westinghouse Health Systems, are the obvious choice for specialized surveys, such as contraceptive prevalence surveys. But where the needed service

cannot be easily linked to a specific supplier (such as identifying what information is needed, an appropriate method for obtaining it, etc.), technical assistance appears to be procured in a somewhat haphazard fashion. For example, AID's use of the Survey and Evaluation Unit of BuCen for technical assistance for sample surveys seems to be largely a function of past experience an AID/Washington office or mission has had with BuCen/SEU. As a result, SEU is used by one mission but not another eventhough they need essentially the same type of assistance. Perhaps more systematic use of BuCen/SEU would improve the rate of completed data related activities involving sample surveys. In other instances, a mission obtained very good services from someone who just happened to be known to mission staff (e.g., USAID/Nepal obtained excellent programming assistance from a freelance consultant who was looking for work in Kathmandu). As beneficial as these "fortuitous finds" can be, this is a very chancey means of procuring data support services. In short, a better system for identifying and obtaining assistance in this area is needed. Factors such as the quality of data needed, cost, time constraints and availability of staff to the mission could be used to select the most appropriate sources of assistance for a given project. A support service of this sort managed by AID/Washington and made available to the missions on request could better assure that the Agency obtains more cost-effective assistance for data related activities.