A.I.D. Evaluation News

A Newsletter on Recent Evaluation Findings and Methods

1992 - Vol. 4 - No. 4

Development Experience Reviews

CDIE's Assessment of A.I.D.'s In-Country Presence

by Randal Thompson-Dorman,
Center for Development Information and Evaluation

In the view of many inside and outside of A.I.D., the Agency's most distinctive and valuable asset is its in-country presence. A.I.D. currently has 82 bilateral country organizations of which 48 are Missions, 33 are offices, and 1 is an embassy section. Counting regional support offices and housing offices, A.I.D.'s total overseas organizations number 108.

In recent years changing global commitments and continuing budgetary pressures have led A.I.D. to take a closer look at its in-country presence. Consequently, in spring 1991, the Center for Development Information and Evaluation (CDIE) began an assessment of A.I.D.'s presence abroad.

A 13-member team comprising 11 U.S. direct-hire employees, a retired A.I.D. Mission Director, and a consultant specializing in organizational development conducted the assessment. The purpose of the study was to answer three major questions: (1) What are the essential advantages of A.I.D.'s in-country presence? (2) Are the functions being performed overseas the ones most suited for realizing these advantages? and (3) What are the Agency's options for increasing the cost-effectiveness of its overseas presence while retaining the chief advantages of that presence? The assessment team also examined recent evidence concerning the distribution of U.S. direct-hire staff levels in Missions to determine the degree of correlation of U.S. staffing with program variables.

The team conducted more than 400 interviews, using both open-ended and structured questionnaires. The team began by first interviewing a broad

Inside this issue

spectrum of senior A.I.D./Washington managers to understand Washington's perception of the benefits of A.I.D.'s in-country presence and to discover which functions Washington considers best performed overseas. Members of the team then visited Costa Rica, Bolivia, Guinea, Honduras, Indonesia, Kenya, Morocco, Pakistan, Senegal, and Uganda to conduct further extensive interviews.

In each country, team members interviewed all U.S. direct-hire staff and senior foreign service nationals and some of the private consultants, as well as representatives from the host government, other donors, and the private sector. In addition, the team conducted a statistical regression analysis to determine the degree of association across Missions of program and country variables with factors that most directly correlated to U.S. direct-hire staffing levels.

The Major Advantages of A.I.D. Presence

The assessment team identified several advantages of A.I.D.'s overseas presence that increased the Agency's effectiveness. Chief among these were two factors that the team felt could best be obtained through in-country presence: influence and program accountability. The interviewees agreed that A.I.D.'s presence strengthened the Agency's influence on several levels, for example, in

- Shaping the country program to best enhance the recipient country's development prospects and ensuring that overarching U.S. policy objectives are realized
- Convincing the recipient government to make the policy changes necessary to overcome obstacles to its development programs
- Building consensus among donors on the country's development problems and the appropriate measures for addressing them
- Persuading Washington decision-makers to accept and support policy objectives specific to and appropriate for each country assisted

Interviewees noted that A.I.D.'s influence resulted from the specific benefits derived from having U.S. staff living in the host country. The personal working relationships developed through daily contact contributed to a better understanding of local conditions and greater political and cultural sensitivity. Interviewees said that U.S. staff were able to gain access to host country officials, maintain daily involvement in the process of policy reform, keep development on the U.S. Government agenda, and promote sensitive issues, such as family planning and environmental protection.

A.I.D. presence also promoted greater accountability for program effectiveness. For example, respondents noted that the presence of U.S. staff in the country improved the quality of program and project implementation, allowed for regular review of projects and quick response to new priorities, permitted mid-course corrections during implementation, generated better inside information about projects, led to prompt decision-making by A.I.D., facilitated resolution of misunderstandings and miscommunication, and provided institutional continuity.

The Overseas Functions

That A.I.D. presence provided important advantages was clear, but were A.I.D. Missions and offices staffed appropriately to optimize those advantages? Moreover, was A.I.D.'s approach to its overseas functions as cost-effective and efficient as it could be?

The study concluded that some functions, particularly those relating to policy dialogue, strategy formulation, and program and project negotiations, were so intimately linked to the two key advantages of influence and program accountability that they clearly could best be performed by U.S. direct-hire employees assigned to the country. Other functions,

A.I.D. Evaluation News

Farah Ebrahimi—Managing Editor
Abbe Fessenden and Steven Gale—Contributing Editors
Pamela McDade—Publication Coordinator

Produced by the Center for Development Information and Evaluation (CDIE), with contract assistance from Conwal Incorporated for editing and desktop production. The views and interpretations expressed herein are those of the authors and should not be attributed to the Agency for International Development. Comments, articles, and inquiries may be sent to Farah Ebrahimi, Managing Editor, *A.I.D. Evaluation News*, CDIE, Agency for International Development, Washington, DC 20523-1802. Articles should be six double-spaced typed pages or fewer.

however, could be delegated to foreign service nationals or contract staff or be performed elsewhere.

Overarching Concerns

Given these findings, how can A.I.D. preserve the essential advantages and yet increase the cost-effectiveness of its overseas presence? The assessment team expressed two overarching concerns related to this challenge: (1) inadequate planning for transition to self-reliance by countries in the management of their development efforts and (2) widespread underutilization of foreign service national staff.

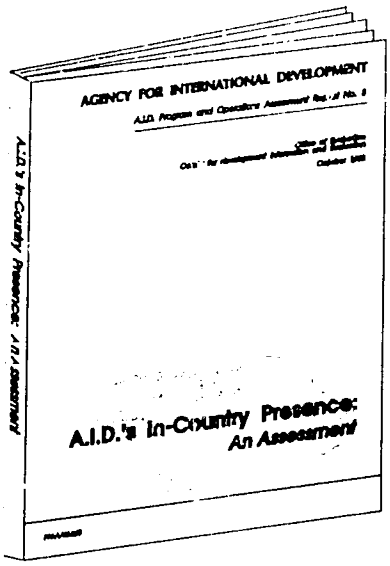
A.I.D. is in the business of assisting recipient nations to achieve their development objectives. A.I.D. Missions therefore should be specifically planning, which they are not, for the transition of these nations to self-reliance. This means that Missions should be developing transition management strategies at program and project levels as an essential part of their strategic planning. These plans should indicate time horizons and benchmarks for increasing Mission dependence on non-U.S. direct-hire staff and for eventually turning over to the host country the management responsibility of A.I.D.-supported development activities. Currently, Missions have no incentive for such planning. In fact, Mission staff are rewarded for helping their Mission grow, not contract. Moreover, the ever-increasing contracts, reporting requirements, and financial accountability rules militate against U.S. direct-hire staff relinquishing control.

Foreign service nationals, who could be relied on to perform many of A.I.D.'s overseas functions, appear to be underutilized in a number of Missions. Much of this underutilization can be attributed to confusion about the legal constraints on using foreign personnel for key functions. Another reason is reluctance to rely on such staff. The existing legal scope for fuller utilization of foreign service nationals needs to be clarified and widely promulgated.

Beyond these broad considerations, the team clustered specific proposals into short-, medium-, and long-term actions.

A.I.D.'s In-Country Presence

An Assessment



Order this report from:

A.I.D. Development Information Services Clearinghouse
1500 Wilson Boulevard, Suite 1010
Arlington, VA 22209-2404
telephone: (703) 351-4006

Please give the following number when ordering: PN-AAX-260.

Short-Term Opportunities

Adjust current U.S. direct-hire staffing imbalances. According to staff interviews, many factors entered into overseas staffing decisions. Nevertheless, the regression analysis of data covering the last 4 years identified two statistically significant variables—program obligation levels and total number of projects in a Mission's portfolio—that explain 90 percent of the variation in U.S. direct-hire staffing levels among Missions. The regression analysis also identified staffing levels at several Missions that departed substantially from those predicted based on program and project levels. An examination of the particular circumstances of these Missions could identify other reasons for their staffing levels and might suggest reallocation of staff.

Identify countries for immediate transition. A number of countries, for example Costa Rica, India, Thailand, and Tunisia, are capable of managing their own development programs with a minimum of A.I.D. presence. For these countries, in particular, transition plans should be in place.

Modify programs in unfavorable development environments. In some countries, A.I.D. may have a sizable aid program, but faces an environment unsuitable for a traditional, staff-intensive aid program. Missions in these countries must cope with such situations as political instability in the country, poor policy environments, corruption, or uncooperative governments. An option in such cases is to limit the A.I.D. program to less staff-intensive approaches, such as an exclusive focus on participant training or private voluntary organizations (PVOs).

Decrease documentation and reporting requirements. The interviews highlighted the inefficiency caused by the in. dinate reporting requirements that drive much of staff time. Staff believe they spend too much time preparing reports and performing project analyses and design. Although attempts are being made to reduce this burden, A.I.D. should follow up to ensure that the efforts are successful.

Lessen Bureau/Mission competition. A.I.D.'s decentralized structure breeds competition among Missions and between Missions and bureaus for scarce human and financial resources. Indeed, Mission U.S. direct-hire staff spend much time attempting to convince A.I.D./Washington to provide them with resources. A solution would be to send a joint A.I.D./Washington and Mission team to the field to assess the Mission's needs and propose a strategy for meeting them.

Medium-Term Opportunities

Medium-term options tend to involve more extensive policy changes and restructuring. A few of these are discussed below.

Establish and expand regional and shared services. A.I.D. already has an organizational model for providing regional technical and support services—the Regional Economic Development Services Offices (REDSOs) in Africa. This model can be explored for wider use, incorporating among its tasks many of the support services and even some of the technical services being offered by individual Missions. USAID/Bangkok is beginning to provide such services to other Missions in the region. Such a model could be particularly successful in Central America. There, A.I.D. could phase out its individual bilateral Missions in favor of a single regional Mission headquartered in Costa Rica with branch offices in neighboring countries staffed by a much reduced U.S. direct-hire presence.

Focus on what A.I.D. does best and on addressing fewer development problems. A.I.D. could reallocate staff and increase impact if it reduced the number of units of management and focused on areas in which it had the greatest comparative advantage in the country.

A.I.D. Missions should be specifically planning, which they are not, for the transition of these [recipient] nations to self-reliance.

Create incentives that promote efficiency and effectiveness. The team proposed a reorientation of A.I.D.'s incentive system so that Mission Directors are encouraged and rewarded for running lean operations, consistent with the management objectives of program performance and host country self-reliance in management.

Establish Mission antennae in A.I.D./Washington. The team suggested that consideration be given to relocating to Washington all Mission and regional personnel whose functions do not require daily project monitoring or dialogue with recipient-country officials. Such personnel would work in Mission "antennae" offices in Washington. They would be responsible to the Mission Director under the direct supervision of a manager designated as assistant director for Washington-based operations.

Introduce the concept of limited or appropriate financial accountability. Justification for U.S. direct-hire staff is often based more on A.I.D./Washington and Mission concern for financial accountability and a pervasive fear of audits than the requirements of basic program management. There is a general perception in A.I.D. that the costs associated with the Agency's approach to financial accountability exceed the benefits it is accruing. The assessment recommended that the Agency, in cooperation with Congress and the Inspector General, investigate the concept of "limited accountability" for overseas operations in the developing countries as practiced or being explored by other donors.

Long-Term Opportunities

Limit staff to those essential to achieving the Mission's strategic objectives. In the long term, A.I.D. could introduce a core Mission concept, which would limit U.S. direct-hire staff to those essential to achieving the Mission's strategic objectives. This approach is similar to zero-based budgeting, with the U.S. direct-hire principal officer as the point of departure. Any additional direct-hire program managers would be tied to the number and type of strategic objectives and the availability of qualified local staff.

Each program manager would have broad managerial responsibilities and would be responsible for all phases of activities under his or her strategic objective. All other functions would be performed by foreign service nationals or contract staff, supplemented by U.S. direct-hire staff through telecommunication links and field visits from Washington or from regional support offices or other Missions. The principal officer (Mission Director or A.I.D. representative) would be required to justify requests for additional U.S. staff on such grounds as nonavailability of qualified local staff, inadequacy of telecommunication links, and scale and complexities of program.

For more information on this study, contact Ronald D. Levin, Center for Development Information and Evaluation, Agency for International Development, SA-18, Room 219B, Washington, D.C. 20523-1802 or call (703)875-4852.

In-Country Presence: The Experience of Other Donors

by Ronald D. Levin, Labat-Anderson, Incorporated

In the majority of countries where A.I.D. maintains Missions, the Agency's presence is, or appears to be, much larger than that of other donors. Does the experience of other donors offer any lessons for A.I.D.? Are other donors even pursuing the same objectives as A.I.D.? These are two of the questions addressed in a recent evaluation of A.I.D.'s in-country presence, conducted by A.I.D. Policy Directorate's Center for Development Information and Evaluation (CDIE). (See article by Randal Thompson-Dorman in this issue of *A.I.D. Evaluation News.*)

In carrying out the evaluation, the assessment team conducted a literature search in Washington followed by extensive interviews in 10 countries on 3 continents. Team members interviewed representatives of the major bilateral and multilateral organizations, as well as officials of the host countries. From the study several conclusions emerge.

Although most donors require some kind of in-country presence, none expects its overseas staff to do as much as A.I.D. requires. For example, no other donor requires its staff to have much of a role in preparing development strategies, designing projects, procuring goods and services, conducting a policy dialogue, supervising project or program performance, or performing financial management, accounting, or evaluation. Few authorities, if any, are vested in the field, and issues are routinely decided at headquarters.

Thus it is not surprising that the overseas staffs of other donors are not large. But to count only the personnel who sit in a bilateral donor's aid office can be deceptive; often they represent only a small part of their country's aid presence.

Of the donors the team surveyed, Germany's Ministry of Economic Cooperation has the thinnest presence overseas—with only about 20 officers, seconded to the Foreign Ministry for assignment to the largest of the 120 German aid recipient countries. This sparsity is somewhat compensated by Germany's reliance on parastatal institutions and contract staff to help the recipient countries implement capital and technical assistance programs.

Japan's foreign aid program is thinly staffed overseas as well. In Indonesia in the late 1980s, for example, a staff of five administered Japan's aid budget of well over $1 billion. For reasons of economy, some countries, including Canada, the Netherlands, and Sweden, limit the number of recipient countries where aid staffs—of 2 to 10 officers each—are assigned. Only countries with aid staff in place to handle the work load are eligible to receive the full panoply of the donors' development assistance. These donors provide assistance through their embassies to other countries as well, but using modes of assistance that do not require the presence of an in-country professional staff. Some donors, notably the Scandinavians, reach a broader group of recipient countries by channeling a proportionally higher level of resources through multilateral organizations, such as United Nations Development Programme, using, in effect, the staff capacity of those organizations as a substitute for their own.

Both Britain and France maintain assistance relationships with many countries but concentrate on former colonies, where they have aid staff of 5 to 15 officers attached to their diplomatic Missions. Britain also has five regional development divisions (in the Caribbean, the Pacific, Southeast Asia, Southern Africa, and East Africa) that reinforce its diplomatic Missions in carrying out Britain's development aid program. In many places, these offices are an alternative to Britain's in-country presence. In essence, they form a compromise between the need for technical expertise in each recipient country and the expense of maintaining field staff in Britain's more than 120 recipient countries. The British aid program also enjoys the support of other official bodies, such as the Crown Agents, the Commonwealth Development Corporation, and the British Council. Although these organizations do not appear to be part of the British aid structure, they take direct responsibility for parts of the aid activity.

More significant for the French and the British, however, are their networks of thousands of expatriate technical experts assigned within the operational institutions of recipient governments. These experts technically function as civil servants of the host-country ministries to which they are assigned. In many ways, however, they act as extensions of the donor's official aid presence by providing constant administrative support, planning, and monitoring for bilateral projects.

In addition to contributions to the World Bank, the specialized agencies of the United Nations, and the regional development banks worldwide, all European Economic Community donors make substantial contributions to the European Development Fund (EDF). The EDF is, in many ways, a surrogate for European bilateral aid. It manages large multilateral assistance programs in the Africa, Caribbean, and Pacific regions. In 1990, the European countries channeled more than $3 billion in development and economic assistance through the EDF, which typically maintains an in-country presence of 5 to 12 career officers, who are aided by frequent visits from teams of specialists.

For A.I.D. to reduce its overseas presence to a size commensurate with...other donors would require a substantial recentralization of authority in A.I.D./Washington and a dramatic reduction in scope...

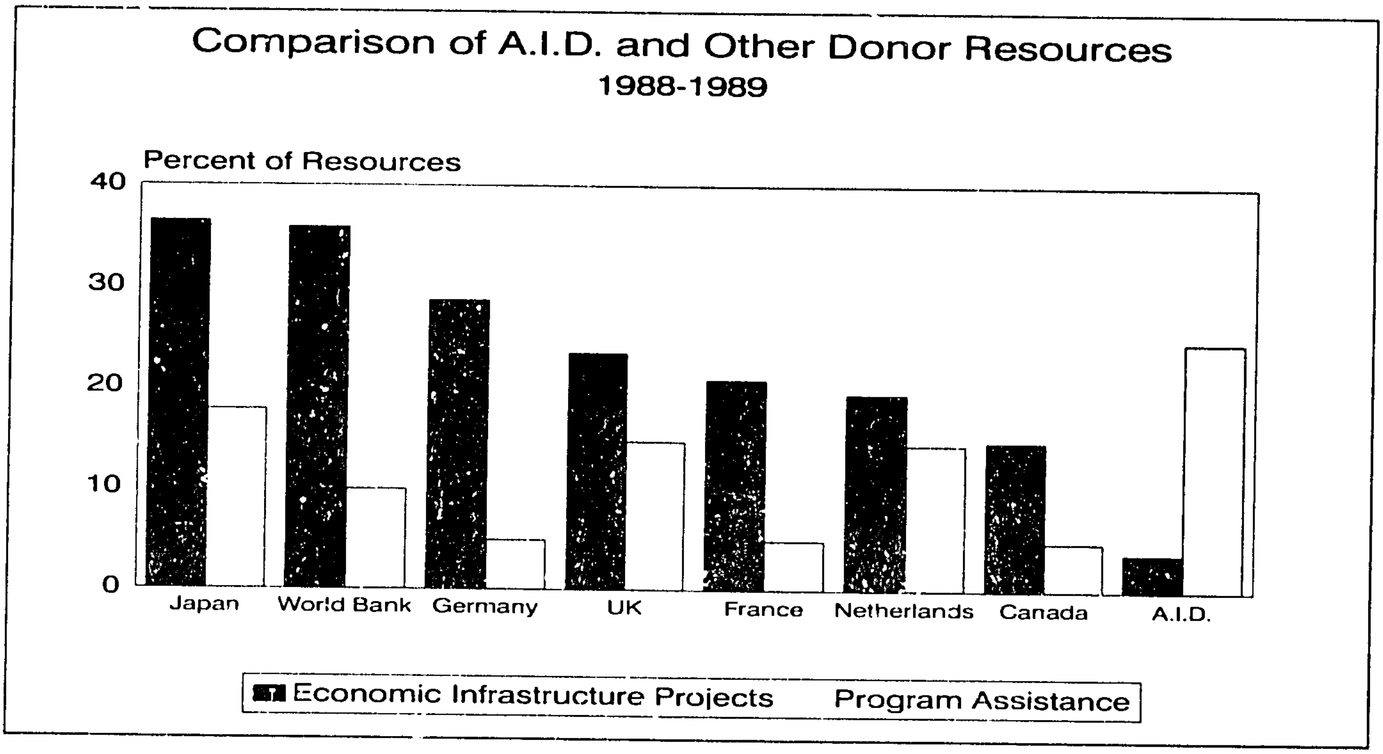
Although most major bilateral donors have tried to get along with only limited overseas presence, some are finding they need more. The Japanese, in particular, have found themselves increasingly unable to program and disburse the greatly increasing amounts of aid they are providing. Japan is urgently pursuing expansion of their in-country presence in some of the countries they are assisting.

The Canadians, having concluded that their system needed to be made more responsive and timely, embarked on a decentralization effort in 1988. Their effort included a considerably greater in-country presence, based on what they term the "USAID model." Even the World Bank has expanded its in-country presence since 1985, adding substantial numbers of technical specialists. Indeed, interviews revealed that many donors feel that they need more people working for them, whether hired locally or sent from their headquarters, but not as many as A.I.D. requires.

Other donors do not seek the same objectives as A.I.D. In many cases, the advantages A.I.D. seeks to realize through its in-country presence are less relevant to other donors. The quality of development performance, for example, is of far less institutional concern to most bilateral donors than it is to A.I.D. (although it is of great personal interest to their officials in the field). Donors recognize, however, that without extensive in-country staff they are unable to monitor closely the implementation of their activities. To resolve their dilemma, they rationalize in some cases that effective implementation is the responsibility of the recipient. In other cases, where development assistance is meeting the commercial objectives set for it, the quality of performance is clearly a marginal issue.

Illustrative of this point is the emphasis different donors put on their individual assistance portfolios. For example, during the 1988–1989 period, a significant percentage of the resource flows of other donors was for economic infrastructure projects (e.g., transportation and energy), which are normally turnkey operations implemented by contractors from the donor country (see figure). During this period, the percentage of development assistance budgets that donors devoted to this sector was as follows: Japan (36.4 percent), World Bank (35.8 percent), Germany (28.6 percent), United Kingdom (23.4 percent), France (20.8 percent), Netherlands (19.4 percent), and Canada (14.7 percent). By contrast, the equivalent percentage for A.I.D. was 3.7 percent. However, during the same period, program assistance, as opposed to project assistance and more closely related to development performance, presented a different picture. The share of program budgets devoted to this mode was as follows: A.I.D. (24.6 percent), Japan (17.8 percent), United Kingdom (14.7 percent), Netherlands (14.4 percent), World Bank (9.9 percent), France (4.9 percent), Germany (4.9 percent), and Canada (4.8 percent).

Bilateral donors—except for A.I.D.—also typically do not become deeply involved in influencing macroeconomic adjustment or policy reform through their economic assistance programs. To the extent that their governments have an interest in exerting that kind of influence, it is done at the political level through their embassies. Indeed, inter-



view comments suggest that A.I.D. Missions are viewed by some of the other donors as intrusive, insensitive, and paternalistic.

The World Bank, as a multilateral organization, cares both about the quality of project implementation and about policy change. But the Bank recognizes that it faces problems in both. It has just completed a study of why so many of its projects are far behind schedule and remain far from accomplishing their objectives. Field interviews repeatedly pointed to the fact that teams sent by the Bank to study these problems cannot assess the true state of affairs because their visits are "managed" by the host country and are too brief to produce reliable results.

No donor accepts the idea of diversion of project funds with equanimity or likes the idea of waste and mismanagement. The approach of most other donors, however, seems to be one of accepting some degree of resource erosion as a smaller price to pay than the cost of extensive controls. Routine audits, and the corollary requirement to respond to audit recommendations, are unheard of. Scandal, of course, gives rise to corrective action when it surfaces and sometimes leads to the withdrawal of all development assistance from a recipient country. But the donors interviewed uniformly believed that A.I.D.'s requirements for fund accountability were, in third world environments, unachievable, excessive, and counterproductive.

Other donors are subject to far less oversight. It is fair to say that no other donor, bilateral or multilateral, responds to oversight requirements that are as deep or as complex as A.I.D.'s. In parliamentary systems, the legislature has a limited role in the aid program. Some donors, such as Japan and France, have no aid policy legislation at all. Legislatures in other donor countries, while establishing general policies, review the programs only at very general levels and appropriate funds in a single-line item, leaving the country allocations and project selection entirely to the responsible executive entities. Thus, no overseas staff ever needs to prepare documentation for submission to its legislature or to supply information in response to legislative inquiries. Policies and procedures, too, are vastly simpler and applied with more uniformity. They thus do not require the large numbers of legal advisers, contract specialists, controllers, and project development and program officers to interpret changing objectives and complex requirements, or to monitor conformance with rules and regulations.

The multilateral donors, although responsible to governing boards, control their own decisions largely through their bureaucracies. There are no other bureaucracies, legislative bodies, or constituencies to be satisfied.

Relevance of the experience of other donors. Comparison of A.I.D.'s overseas staffing with that of other donors does, indeed, demonstrate that staff of other donors are much smaller and, thus, less costly than

A.I.D.'s. Even when other donors rely on large numbers of contractors or technical experts for some aid functions, there is no added cost for the services because the individuals are already assigned to the country for other purposes. For A.I.D. to reduce its overseas presence to a size commensurate with that of the other principal donors would require a substantial recentralization of authority in A.I.D./Washington and a dramatic reduction in the scope of the assistance effort and the audits, controls, and legislative oversight to which the program is subjected.

For more information on this subject, contact Ronald D. Levin, A.I.D./Center for Development Information and Evaluation, SA-18, Room 219B, Washington, DC 20523-1802 or call (703) 875-4852.

Evaluation Design and Methodology

On August 4-5, 1992, the Office of Evaluation in CDIE held a workshop on the theory and practice of evaluation. A number of the presentations given at the workshop had broad relevance to evaluators outside of CDIE. Two of these have been summarized for this issue of A.I.D. Evaluation News. *The first article, drawn by Farah Ebrahimi from the presentation given by Dr. Peter H. Rossi, summarizes 12 points for conducting evaluations. The second article summarizes Gary Hansen's presentation on classical versus neoclassical approach to evaluation.*

Twelve Points for Conducting Evaluations

by Peter H. Rossi,
University of Massachusetts, Amherst

In its broadest, most inclusive meaning, to evaluate means to place a value on some object, action, or thought. In this context, all humans must be regarded as evaluators because they are constantly assessing, judging, and hence evaluating. But in the narrow context of the word, to evaluate means to make a judgment in a particular way about a particular class of objects. Within the framework of modern evaluation, the class in question consists of public programs and policies; the particular way is that of social science. So, modern evaluation may be defined as the application of social science research methods to the design and improvement of policies and programs and to the assessment of their efficiency and effectiveness.

Historically, the purpose of evaluation research was to provide ways of assessing whether or not programs were succeeding in achieving their stated goals. It soon became clear, however, that there was also a great need for the products of social science and social research in the design of programs. Consequently, two major evaluation modes evolved through the years: (1) formative evaluation to improve programs at the design stage and to monitor the implementation of programs and (2) summative evaluation to assess their success by developing indicators of net effects. By their nature, operating agencies or units are usually more interested in formative evaluation, whereas policymakers and oversight agencies, such as the Executive and the Congress, are more interested in summative evaluation.

But whether formative or summative, evaluation practice can be helped by "laws," some of which are applicable to all types of evaluation, others only to a specific type. From these, I have culled 12 of the more useful, fully aware that they by no means exhaust the corpus of laws that could be written. I have grouped these laws into three sets: general (Law G), formative (Law F), and summative (Law S).

General Laws

Law G1: There is no such thing as a free evaluation.

Every evaluation effo' entails a cost. In addition, there is a rough proportionality between quality and price. Indeed one of the unfortunate side effects of the widespread acceptance of evaluation has been the tendency of legislatures to tack on, without appropriating additional funds, an evaluation requirement to legislation authorizing a program. Such actions force agencies to undertake evaluations for which they have neither the qualifications nor the resources. As a result, agencies often take a pro forma approach to conducting evaluation, producing work that falls short of its intent.

Law G2: Evaluations should not cost more than the program being evaluated.

Evaluation is not an end in itself; it must be subservient to the program to which it is applied. Consequently, trivial programs do not merit elaborate evaluation, but important programs do.

Law G3: Evaluation design and program design should get under way simultaneously.

Most experts agree that evaluation activities should be built into every program at the outset. Even the best ex post evaluations can never attain the same degree of validity as evaluations that have been planned at the outset and systematically pursued through the planning, design, and implementation stages of the program.

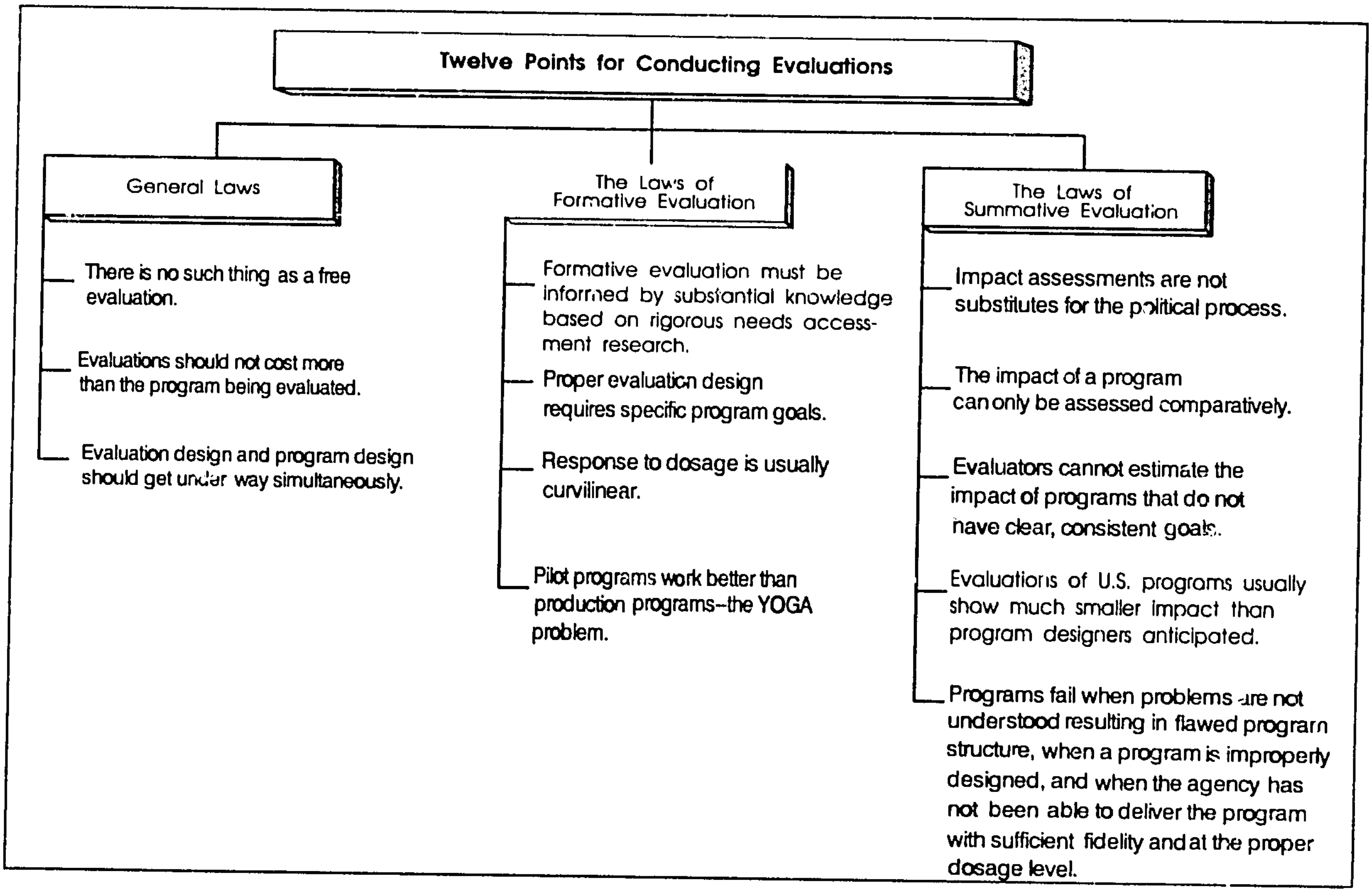
The Laws of Formative Evaluation

Law F1: Formative evaluation must be informed by substantive knowledge based on rigorous needs assessment research.

A program design can only be successful if it is based on the best intelligence available on the nature, extent, and location of the problem in question. Similarly, formative evaluation must also be informed by substantive knowledge based on rigorous needs assessment research. In this context, a general knowledge of social science theory and practice is essential for understanding the processes that give rise to social problems and the processes influenced by changes in policy. In short, without a thorough needs assessment informed by current social science theory and empirical knowledge, the chances of inventing an effective program are remote.

Law F2: Proper evaluation design requires specific program goals.

If you don't know where you're going, you can't figure out how to get there—an elementary but often ignored rule in program design. There are all too



many examples of legislation that provides funds for a program without specifying what that program is to accomplish, resulting in diffused and unfocused activities. "To improve well being," for example, is not a sufficiently precise goal for developing an effective program. Program designers must know specifically what, for their purpose, constitutes an effective program.

... the art of impact assessment lies in evaluators' skill in defining and using the best possible comparison groups or situations.

Law F3: Response to dosage is usually curvilinear.

Reducing the dosage of a treatment does not usually produce a proportional reduction in response. Often, the response is reduced much more, at times even to zero. Accordingly, the effectiveness of a 2-week educational campaign is usually not half that of a 4-week campaign, rather much less.

Of course the reverse is also true in that additional doses do not proportionately increase results. Program designers and managers therefore must always attempt to identify an optimal level of treatment. But in lean times, it is always tempting to reduce the program's resources to save money, thus running the risk of diminishing the program's effectiveness to zero and dissipating the funding allocated.

Law F4: Pilot programs work better than production programs—the YOGA problem.

It is relatively easy to run pilot programs effectively when they are under the control of those who designed them. The critical issue is fashioning full-scale programs that continue to run effectively when turned over to others—particularly, when turned over to "your ordinary government agency" (YOGA). The YOGA problem is not unique to the United States, nor does it apply only to Federal agencies. It is universal, affecting private as well as government agencies.

The Laws of Summative Evaluation

Summative evaluations estimate impact—that is, effects that have occurred over and beyond those that would have occurred without the intervention. Summative evaluations are difficult to carry out well and require evaluators with technical expertise.

Summative evaluations or impact assessments are usually commissioned by agencies or offices with oversight responsibility, for example, by Congress or by the central policymaking unit of an agency. As a result, program operators often consider summative evaluators as antagonists. In contrast, formative evaluations are carried out in close cooperation with program designers and operators. Formative evaluators are often regarded as the "good cops," whereas the "bad cops" are the summative evaluators.

Law S1: Impact assessments are not substitutes for the political process.

Policymaking is a political, not a technical, function. Consequently, whether a program is effective is not the sole determinant of its continuation. Other factors are considered as well, which is as it should be. There are many examples of weak or ineffective programs that nevertheless have continued (a prime example is job training programs); just as there are examples of effective programs that have never been implemented as social policy because of some feature that was considered politically unacceptable.

Law S2: The impact of a program can only be assessed comparatively.

The impact of a program can only be assessed when compared with the absence of that program: What would the situation have been like without this program? The second law of summative evaluation makes the use of comparison groups mandatory, and randomized controlled experiments the preferred method for making such comparisons. Indeed, the art of impact assessment lies in evaluators' skill in defining and using the best possible comparison groups or situations. Sometimes programs achieve such broad coverage that comparison groups cannot be found. For example, the coverage of the U.S. social security old age pension program is so broad that it might be difficult to find a meaningful comparison group of people not covered by the program. In such cases, evaluators might have to rely on the judgment of experts instead of control groups, which may not be as reliable.

Law S3: Evaluators cannot estimate the impact of programs that do not have clear, consistent goals.

Law S3 is a restatement of the second law of formative evaluation. It maintains that to design a successful impact assessment, program designers must have specified in advance the program's indicators of success—a process that involves translating program goals into concrete measures of success. For example, a program goal described as "improving the quality of living in urban areas" is too vague and impossible to assess properly.

Law S4: Evaluations of U.S. programs usually show much smaller impact than program designers anticipated.

Most program evaluations in the United States find that programs have been only marginally effective in reaching their stated goals. This law is often referred to as the "iron law of evaluation." It is based on two findings. First, the United States is a developed society that has moved a long way toward improving the level of living of most of its members. After reaching a certain level, however, achieving additional improvements becomes increasingly difficult. For example, it is relatively easy to move a society from 10 percent to 60 percent literacy. But moving the society from 80 to 90 percent literacy is much more difficult. Similarly, historically, much of the reduction in U.S. mortality rates was accomplished through such rudimentary and inexpensive health measures as installing sanitary sewage and drinkable water systems. To further reduce mortality, however, has required more resources at a higher level of effort with poorer prospects of immediate success. Second, the tendency of most officials, as it should be, is to require evaluations of programs whose effectiveness is not readily apparent. Thus, it is not surprising when findings show that the programs have indeed had shortcomings.

Law S5: Programs fail for three reasons: (1) when the problems have not been correctly understood resulting in a flawed program structure, (2) when a program is improperly designed, and (3) when the agency administration has not been able to deliver the program with sufficient fidelity and at the proper dosage level.

First, an impact assessment has to take into consideration the fit between a program and the existing knowledge concerning the problem in question. If there is no correspondence between the two, the program will fail. For example, program designers of a housing voucher experiment assumed that the families living in the substandard housing would agree with the project's definition of what constituted substandard housing and why they would welcome a change for better housing. In fact, many of the program's standards were irrelevant to the participants.

Second, programs often suffer from design flaws that weaken a program's effectiveness. For example, a famous California program providing group therapy for prisoners was designed to use prison guards as the therapists, an obvious design flaw that ensured the program's failure. Most design defects are not as apparent, however, and are more difficult to detect. Finally, all too often, the implementing unit or agency responsible for the program is unsuited for the task or is not given sufficient resources. The result is a fatally weakened version of the program.

Conclusion

Evaluations, whether formative or summative, are not just technical exercises. They require technical skill informed by substantive knowledge of the program and problem. Of the many principles guiding the science of evaluation, these 12 are central, linking evaluation research with the required substantive knowledge of the program and the problem it is designed to address.

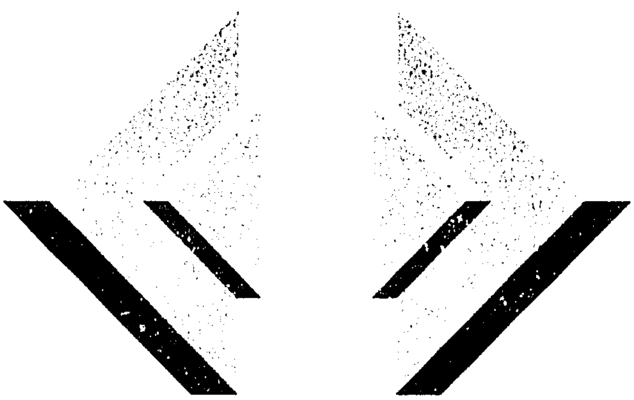
Peter H. Rossi is S.A. Rice Professor Emeritus of Sociology, University of Massachusetts. He is past-president of the American Sociological Association and the recipient of the Myrdal Award for contributions to evaluation research, awarded by the American Evaluation Association. His recent books include Evaluation: A Systematic Approach *(with Howard Freeman)*, Thinking About Evaluation *(with Richard Berk), and* Down and Out in America: The Origins of Homelessness.

Classicism, Neoclassicism, and Beyond: Fission and Fusion in the Praxis of Evaluation

by Gary Hansen,
Center for Development Information and Evaluation

The discipline of evaluation is in a state of paradigmatic flux, with new approaches rising to challenge and even supplant traditions and canons that have long held sway in the profession. Such change is a healthy sign that the field is dynamic, with many practitioners seeking to break out of the narrow epistemological boundaries that have dominated the discipline and tap into the new and fertile streams of thinking emerging from advances in a variety of new areas.

The intellectual divisions, which are rife within evaluation, will be presented here as two ends of a spectrum. One end is what I call the "classical" paradigm and the other is the "neoclassical" paradigm. The full spectrum is rich with issues for discussion, but for brevity, and at the risk of oversimplification, the focus will be on characterizing the differences at the two extremes.



Classical approaches to evaluation stand in the great tradition of logical positivism, meaning that there is a real world of facts, where social and economic interactions function with some degree of cause and effect regularity and where distinctions can be made between dependent and independent variables, which can then be reduced to verifiable and replicable laws.

The classical paradigm characterizes much of what has been and continues to be done in the field of project and program evaluation. Its trademark is embodied in the "impact" evaluation, when an effort is made (1) to determine whether the project or program achieved its prescribed purpose, (2) to identify the variables that accounted for the project's success and failure, and (3) to formulate prescriptions for improving the design and implementation for similar projects and programs.

The classical approach is particularly appealing to those who ask the proverbial "accountability" question, Did the donor and recipient carry out their mandates efficiently and effectively in expending public resources and thereby secure an adequate return on investment? The classical approach is especially suitable for measuring the outwardly physical properties of the development process, for example, the number of patients served in a health clinic, increases in crop yields, and the number of roads and dams built.

While the classical approach may be a convenient instrument for displaying the tangible indicators of progress or failure, it begins to manifest severe limitations when used to diagnose the conditions leading to success and failure—particularly when the implementing agencies of the host country are public or quasi-public, nonmarket organizations, such as education, research, judicial, health, and family service agencies (which usually are responsible for implementation). The makeup of such organizations is complex, and the organizations' behavior is frequently a product of ideological, political, and social forces that emanate from many sources and are not reducible to simple cause and effect equations.

The classical paradigm has other shortcomings. Its approach is retrospective and its 20-20 hindsight frequently, and unfairly, makes decisions and actions of managers and implementors seem irrational and shortsighted. And, because the classical paradigm is retrospective, it frequently lacks the tools for prescribing how to ensure and sustain program or project effectiveness into the future. In particular, its retrospective posture has little to offer in guidance in an environment undergoing significant change.

Finally, the classical paradigm is, one hopes, a value-neutral methodology in the service of theory. One hopes because often where theory is lacking, the paradigm ceases to be a means and instead becomes an end in its own right. This danger is apparent in the not-so-seldom rush to develop questionnaires and indicators for measuring variables that are not well grounded in theory. When used this way, the classical paradigm cannot ask the larger purposive questions, such as, Is a program or organization under evaluation addressing the appropriate problems within its environment?

In the end the classical paradigm will please the auditor and the project officer who want to know if their money has been well spent, but the programmer and policymaker will still feel a sense of unease, if not dissatisfaction, with the kind of static, reductionistic truisms and mechanistic recommendations that this approach is prone to offer as a guide for future investments. It is this particular inadequacy, along with the other deficiencies mentioned, that is driving some evaluators to either abandon or supplement the classical approach with the neoclassical paradigm. The neoclassical paradigm is fed from a multitude of intellectual tributaries, the most prominent of which are the organizational and management sciences, systems theory, learning theory, and action research. Its emphasis is on applying measurement methodologies that help organizations think, learn, and plan strategically. Many neoclassicists are somewhat skeptical about how much guidance for the future the study of the past can provide. Indeed, past success in program performance frequently inhibits an organization from discerning important changes within its environment and even limits its capacity to comprehend the potential for decay within its own operations.

From the neoclassical perspective, the primary emphasis is on investments in building a recipient organization's capacity for learning. Such an approach provides a view on how organizations and the programs they undertake should be evaluated that is entirely different from the classical paradigm. Using the neoclassical approach, programs and projects become both a means and an end; a means in that a program or project is an instrument for the organization to use to learn from its environment and to adapt goals and interventions accordingly. In brief, this paradigm is very dynamic, with the emphasis on understanding and supporting measures that allow, if not induce, organizations to evolve strategically.

The neoclassical paradigm should be appealing to policymakers and programmers, as evaluations generated from this approach will provide a sense of whether a donor-supported organization is on track in developing learning capacities and is thus able to cope and adapt to improve its performance. However, the art and craft of neoclassical evaluation is still incipient and thereby suffers from many of the problems that characterize the emergence of a new paradigm.

In some instances application of the neoclassical paradigm fails to meet the expectations of program designers and policymakers who assume that all evaluations should render concrete and tangible recommendations about what does and does not work. The neoclassicists sometimes resist reducing their work to this level of analysis on grounds that complex and messy development processes cannot be blueprinted into simple formulaic approaches. However, it is frequently the case that neoclassicists have yet to learn how to convey their conclusions and recommendations so that consumers can understand and appreciate them.

Over time, practitioners of neoclassicism should become more experienced in demonstrating the value of their work to the donor and host government consumer. Until then, the paradigmatic upheavals in the evaluation profession will continue as a source of considerable, but, one hopes, healthy debate. Some resolution may eventually be achieved, perhaps as a new synthesis and greater tolerance of paradigmatic pluralism. Then, evaluation designers could see the value of drawing on a variety of approaches and combining them in a form tailored to the task at hand.

More information about the classical and neoclassical evaluation approaches and the application of the neoclassical approach to A.I.D. projects can be obtained from the author, A.I.D., CDIE, SA-18, Room 220B, Washington, D.C., or by calling (703) 875-4853, Fax (703) 875-5269.

A New Framework for Evaluating Development Finance Programs

*by Elisabeth Rhyne**

In recent years, A.I.D. and other development institutions have adopted new views on credit and finance programs, views that emphasize building financial intermediation systems offering both savings and credit services on a commercially sustainable basis. Projects with these aims cannot be easily evaluated using the traditional indicators of direct borrower impact that have been associated with targeted credit programs. This article outlines a framework for evaluation that is in keeping with the new views of finance.

Do Financial Services Benefit the Economy and Clients?

Evaluation of finance projects must begin with well-specified goals and objectives drawn from an understanding of what financial services do. Economists agree that financial institutions contribute to economic growth by performing three crucial functions:

- Mobilizing a society's resources and allocating these resources to efficient uses (intermediation)
- Helping the economy manage risk by diversifying it
- Facilitating transactions, that is, providing axle grease to the economy

Financial services do not directly create economic opportunities; rather, they help people and enterprises position themselves to take advantage of opportunities. In general, goals for finance programs should be specified in terms of improving the ability of financial systems to perform or extend these crucial functions to new areas or client groups. Evaluation should focus on indicators that reveal how well the assisted institutions perform these basic financial functions.

Turning the perspective to the client, the focus on extending or improving financial services for customers continues, particularly for poorer customers, such as small farmers, small business owners, and their households. Financial services allow clients to

- Protect themselves against hardship by building a stock of assets or by borrowing during hard times
- Manage their enterprises more efficiently, purchasing inventory or inputs at advantageous times and prices
- Obtain capital for productive investments

From these possibilities, several important conclusions emerge. First, clients can use savings, as well as credit, for any of these purposes. Many people prefer to save rather than borrow. Second, any given transaction is only one event in an ongoing series of financial decisions that the client manages. Thus, a focus on an individual transaction will miss much of the real story. Third, financial services are used for important functions other than growth-oriented investment. All these considerations set the context for specifying client-level goals for finance projects.

From the point of view of both the economy overall and the client, finance projects should be evaluated based on how well they perform as ongoing parts of the financial system, providing valued financial services. Let us examine how such an evaluation might be organized.

A Framework for Evaluating Finance Projects

Evaluations of finance programs should proceed on two levels—that of the financial institution and that of the client (see box). Each level is associated with a strong commercial test. The client-service relationship is best measured by a market test of demand or willingness of clients to pay. If people pay full cost for a service, then evaluators can be reasonably sure the service is providing something valued at least as highly as its price. By their actions, clients reveal information about the value of benefits, which is more credible than verbal responses to questionnaires, although market tests are less valid when services are underpriced or subsidized.

At the level of the institution, the strong test is financial self-sufficiency, the ability to cover all costs from program revenues. Together, these two tests make finance programs "self-evaluating." The tests are easily verified and give assurance that a program is successful. In the case of donor-supported financial programs, however, these two tests, although important, need to be supplemented.

The fundamental purpose for donor involvement in such programs is to extend the frontier of the financial sector of the developing country through innovation and demonstration effects. Such an orientation requires that evaluations concern themselves with the nature of the innovation, and innovation, most often, involves new client groups and new services.

The following section discusses each element that an evaluation must address, using the framework shown in the box.

Client-Service Relationships

To assess client-service relationships, evaluations should basically resemble market research. The main point of the research is understanding the needs and preferences of, and the alternatives available to, the clients as these relate to clients use of financial services in general and the services offered in particular.

Client Outreach. Indicators of outreach include number of clients reached and basic information on client profile, for example, gender, location, or type of business. Although assessment of outreach is in many ways the most straightforward aspect of an evaluation, the assessment can be controversial because it touches on the highly charged issue of whether targeting a client group is ever acceptable in a finance program. Such concerns can be allayed by adopting a marketing stand that stresses the importance of knowing the customers.

An appropriate targeting strategy centers within the financial institution, not within the donor, and does not include the imposition of stringent eligibility requirements from the outside. Targeting even the very poor and disadvantaged portions of the population is legitimate, provided these groups are treated as potential market niches for the financial institution to serve.

Outreach indicators can be incorporated into ongoing monitoring information systems or gathered periodically through sample surveys. Indicators should be compared with information about the general population of the area covered by the program, so that the function of the program within its context can be assessed. Market penetration ratios, for example, indicate both whether a service is well accepted and how much growth potential it has. For example, estimates of the percentage of female clients relative to the percentage of women in the service area can indicate whether a program needs to make a minor adjustment or start a completely new service to reach women.

Quality of Service. Evaluators should judge finance programs on how effective they are at reaching clients with suitable financial services. Programs should be judged with respect for the clients' preferences and decision-making positions. The strongest and simplest test of service quality, as stated above, is willingness of clients to pay.

The market test for quality should be supplemented with assessments of specific service features. This line of investigation is again very close to market research and should be of similar value. The terms of the service—loan maturity, collateral, group guarantees, liquidity of deposits, for example—should suit the needs of the clients. These terms should be examined in light of the needs associated with clients' economic activities. Evaluators should ask whether the program has developed the right product or whether other products might have greater demand.

Another aspect of service quality concerns transactions costs borne by clients. Because these costs can be a high proportion of total costs of financial services, particularly for poor clients and informal

---

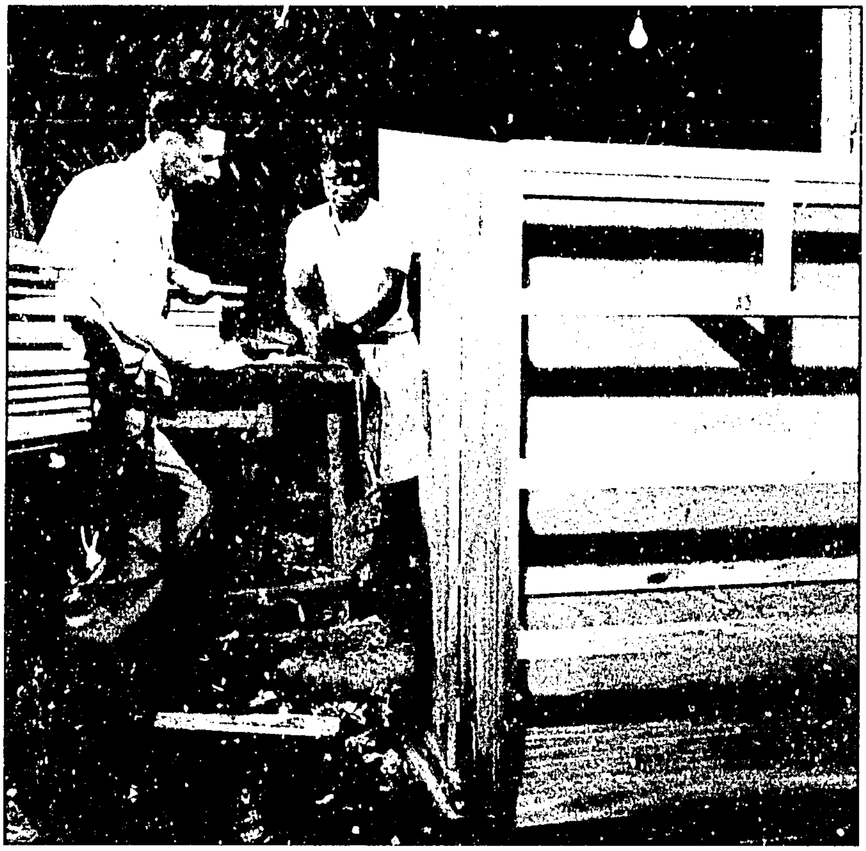
A Framework for Evaluating Finance Programs

The Client-Service Relationship

- Outreach
  - Number of clients; market penetration indicators
  - Characteristics of clients: for example, gender, location, income status, sector of enterprise
- Quality of service versus alternatives
  - Market test: willingness to pay
  - Client transaction costs: for example, convenience and timeliness
  - Service terms: for example, price, loan sizes, maturity, collateral, access to deposits, and eligibility requirements
- Enlarging clients' decision-making options (how the service fits into the client's financial management process: liquidity, consumption smoothing, and investment)

Institutional Viability

- Financial self-sufficiency of service
- Financial condition of institution
  - Profitability or ability to break even
  - Portfolio quality
  - Liquidity
  - Capital adequacy
- Institutional strength and context



Typical clients of development finance institutions are small business owners, small farmers, and their families.

enterprises, convenience and timeliness are crucial features of good services.

The assessment process should always be carried out with a clear understanding of the context provided by other available services. Attributes of a service can be interpreted only when they can be compared with the characteristics of the alternatives. Many standard finance project evaluations assume that clients had no other source of finance. In some cases, what appears at face value to be an extremely successful service may be only marginally better than the next best alternative.

The requirement to assess alternatives adds a significant burden for information gathering, but the comparison is essential if conclusions are to be reliable. In general, information from clients regarding their views on service attributes and alternative services will be credible, especially when compared with standard impact information, such as changes in income. Clients have little incentive to distort their responses to such questions.

Effects on Clients' Range of Decisions. Evaluators should recognize that clients will view the project-provided service as simply another option for consideration during their ongoing financial management. Evaluators should therefore look at how the service changed the strategies available to the client. Questions should stay close to the direct and immediate uses of the service and should examine how the stated use was filled before the service was available. The focus should be on the three main functions of finance: protecting against bad times (consumption smoothing), facilitating efficient business operations (liquidity), and financing investment. Questions should not focus on whether the clients' business has become more profitable, whether more workers have been hired, or whether the family has more to eat. Such questions tend to assume a direct causal link between particular transactions and a final outcome, ignoring the overwhelming effects of the wider economic setting.

Institutional Viability

Evaluations of the financial institution rely mainly on standard techniques of financial analysis. However, such evaluations need a shift in perspective, away from the donor-centric examination of the use of donor funds. The center must be the service and the institution, particularly their financial self-sufficiency.

Self-Sufficiency of the Service. The basic test of service self-sufficiency is whether the service really works from a commercial point of view. Does it have the potential to survive and expand? The answer is yes only if the income from the service covers its associated costs. It is important to answer this question separately from the question of viability of the sponsoring institution (except when the service is the only activity of the institution). The key indicators at this level are profitability of the service and, for credit, portfolio quality.

One problem facing evaluators, particularly for nonbank programs, is a lack of standards for services aimed at marginal groups. Programs tend to be evaluated without reference to industry norms or the achievements of well-performing institutions. Such indicators as arrears, charge offs, and operating margins vary widely from the best institutions to the mediocre, and evaluators have no guidance to determine what levels might be considered standard. Moreover, information is often kept in a form that prevents accurate assessment, especially in the area of portfolio quality.

Financial Condition of the Institution. Just as important as knowing whether the service is viable is knowing whether it has a secure institutional base, and that requires a look at the financial condition of the institution as a whole. Methods for assessing financial condition are well established and widely known. They involve analysis of income statement and balance sheet information. The indicators of financial condition used most often are profitability, portfolio quality, liquidity, and capital adequacy.

Profitability must be assessed from both an operating standpoint (actual costs) and one that takes subsidies into account. When grants or cheap capital make an operation appear more profitable, the subsidy must be corrected for in an assessment of financial self-sufficiency. In addition, the equity of an institution (or quasi-equity for a nonprofit one) must bear a cost at least equal to the rate of inflation to assess whether the operation can maintain its value. Evaluations must recognize stages in the pursuit of self-sufficiency, principally including a start-up stage during which the goal is for income to cover operating costs, followed by a second stage in which income covers the real costs of capital. A program may be acceptable even if it has not reached viability if it shows a convincing trend toward smaller subsidies and eventual commercial operations.

Institutional Strength and Context. Although strong financial performance indicates a strong institution, qualitative assessment must supplement the quantitative indicators, covering internal issues such as leadership, vision, and management, and external issues such as political forces, regulatory issues, and the relation of the institution to the financial system. Examination of these qualitative questions informs decisions about future directions for the institution and its services.

Evaluations based on this two-level framework should provide a clear picture of program achievements and a useful market and service analysis for the assisted financial institution.

**Elisabeth Rhyne was the former A.I.D. Coordinator for the GEMINI Project. Readers interested in the longer report on which this article is based should contact the GEMINI Project at (301) 718-8294.*

Field Perspective

Evaluating the Impacts of Selected Investments in Agricultural Technology in Sub-Saharan Africa

by James F. Oehmke and Eric W. Crawford, Michigan State University

Background

Over the past 15 years, A.I.D. and other donors made considerable investments in agricultural technology development and transfer (TDT) activities in Africa, most prominently in research and extension. However, support for TDT declined steadily from $55 million in 1986 to $30 million in 1991, reflecting a decrease in the proportion of allocation to agricultural TDT from 34 percent for all agricultural activities in 1986 to 14 percent in 1991.*

This dramatic decrease in A.I.D. funding stems in part from the perception that agricultural TDT has had little impact on Africa's development. Such a view has been perpetuated by Africa's stagnant per capita food production, exacerbated by high population growth rates, war, drought, and other factors. Further reinforcing this view are some real problems facing agricultural research and extension organizations.

In order to guide future funding decisions, A.I.D. commissioned Michigan State University (MSU) to conduct a set of studies to measure the impact of TDT on people in Sub-Saharan Africa, as well to measure TDT's success at the sector and national levels. Results from these studies and other relevant assessments were presented at the Symposium on the Impact of Technology in Sub-Saharan Africa held from October 14 to 16, 1992, in Washington, D.C.

A primary purpose of the Symposium was to determine whether accomplishments of TDT were sufficient to justify continued funding. A secondary purpose was to consider the adequacy of available analytical tools for conducting the impact assessment. This article summarizes the evidence presented at the Symposium. It particularly focuses on the analysis of rates of return to TDT investments in Sub-Saharan Africa.

Measuring Impacts

Assessing Rate of Return

The TDT process has four sequential stages: (1) creation of the institutional capacity to develop improved techniques of production; (2) expansion of the technology frontier; (3) transfer of technology to the users; (4) and sustainable changes in long-term productivity.

The rate of return—the most commonly used measure of economic profitability of investments in TDT—summarizes the benefits, costs, and time frame of an activity. Investments with positive rates of return provide benefits that more than cover costs. Investments with rates of return that exceed the return of alternative investments or the cost of obtaining funds are considered economically profitable. Benefits in assessments of TDT investments are usually those that affect people, such as changes in income or other measures of household welfare.

Several other TDT accomplishments are not often considered benefits because of difficulties in quantifying their impact. For example, improvements in the status of women within the household, in the environment, in the sustainability of agricultural production, in the human and institutional capacity for research, and in equity (income distribution) fall within this category. Studies on rate of return presented at the Symposium did not account directly for these other benefits, although there was reported evidence of progress in these areas.

The studies were undertaken from the perspective of national research systems. The reported rates of return, therefore, represent the return to investment in national research programs, taking as given the contribution of the international agricultural research centers.

Rate of Return Assessment Results

Table 1 shows rate of return assessments from ex post studies that analyzed agricultural TDT benefits achieved to date. In general, these rates are not only positive but many are high enough to indicate economic profitability. These striking findings are in direct contrast to the negative views of the impact of African agricultural research permeating recent discussions. In the commissioned studies, only Niger and Uganda show negative ex post rates of return; the remainder show positive returns, ranging from 3 percent for cowpea in Cameroon to 135 percent for maize in Mali. Examined as a group, the estimated rates of return support the proposition that the impact of agricultural research on people in Sub-Saharan Africa is sufficient to justify the investment.

The Symposium examined the alternative interpretations of the consistently high rates of return. For example, if rate of return studies had focused primarily on success stories the results would have been upwardly biased. The MSU study examined cases of likely TDT successes (e.g., Kenyan maize) as well as those considered less successful (e.g., Niger) to avoid such bias. The countries included in the MSU study constitute a stratified random sample, although choice of crop was based on the importance of the crop to the food system or the priorities of the A.I.D. Missions, national agricultural research systems, and ministries of agriculture. Regional evidence relating measures of productivity to measures of all research funding, including successes and failures, also shows considerable positive rates of return. Thus, a biased selection procedure is not a likely explanation of the rate of return results.

A Comparison of Methodologies Used

Although each study used the same conceptual background in assessing benefits and costs, the evaluators' decisions about data collection, the scope of the study, and other critical variables varied somewhat from country to country.

Table 1. Summary of Rate of Return Studies for African Agricultural TDT

Author	Year	Country	Commodity	Time Period	ROR in %
Abidogun	1982	Nigeria	Cocoa	-	42
Makau	1984	Kenya	Wheat	1924-1974	33
Evenson	1987	Africa (Regional)	Maize & Staple Crops	1962-1980	30-40
Karanja	1990	Kenya	Maize	1955-1988	40-60
Mazzucato[a]	1991	Kenya	Maize	[b]	58-60
Mazzucato and Ly[b]	1992	Niger	Cowpea, Millet & Sorghum	1975-1991	< 0
Schwartz, Sterns & Oehmke	1992	Senegal	Cowpea	1981-1986	1-92
Sterns & Bernsten[a]	1992	Cameroon	Cowpea	1979-1992	3
Howard et al.[a]	1992	Zambia	Maize	1979-1991	21
Laker-Ojok[b]	1992	Uganda	Sunflower, Cowpea & Soybean	1986-1991	< 0
Boughton[a] and de Frahan	1992	Mali	Maize	1969-1991	135

Note: ROR is rate of return.
Source: Oehmke, J.F. "Technology impact and agricultural transformation: Lessons learned from impact assessments." Paper presented at the Symposium on the Impact of Technology on Agricultural Transformation in Africa, October 14-16, 1992, Washington, D.C.
[a]A.I.D.-commissioned studies.
[b]Parameter estimation using 1955-1988 data. ROR for research undertaken in 1978 as an example.

First, impact assessments can vary depending on the evaluator's choice of starting and ending points of the evaluation. For example, if a new research program is just beginning to have impact, an assessment of its current impact will not show benefits of the future use of the improved crop or livestock varieties or management practices. Ignoring likely future benefits may lead to an unprofitable or negative rate of return, as in the Niger and Uganda cases. Similarly, if the starting point of the analysis is extended backward, perhaps because the project under evaluation is the second phase of an earlier activity, additional costs would probably be included, reducing the estimated rate of return.

Second, in most cases it is difficult or impossible to identify the impact of given components of TDT, such as research or extension. Table 2 summarizes the components of TDT considered by more recent studies.

Third, some benefits are difficult to quantify. For example, most of the studies discuss, but do not estimate monetarily, the benefits of institution building and improvements in seasonal food security (see Table 2, last column). Often, any reductions in consumer prices resulting from increased production or marketing efficiency are not quantified. These benefits to consumers are particularly important for the poorest farmers, who are often net purchasers of food. Lower prices are expected to have positive implications for equity and result in higher estimated rates of return. A summary of the benefits by study is found in the last column of Table 2.

Factors Influencing Impact

Impact assessments provide insights on factors influencing the success or failure of TDT efforts. Such insights are important for improving the de-

Table 2. Components of TDT by Study

Study	Research Costs	Extension Costs	Other Costs	Other Outputs
Cameroon	Yes	Yes	No	Yes[a]
Kenya	Yes	No	No	No
Mali	Yes	Yes	Yes[b]	No
Niger	Yes	Yes	Yes[b]	Yes[c]
Senegal	Yes	Yes	Yes[b,d]	Yes[a]
Uganda	Yes	Yes	No	Yes
Zambia	Yes	Yes	Yes[b]	Yes[c]

[a]Sensitivity analysis included food security.
bCosts of providing farm-level inputs.
[c]Institution bu·lding.
[d]Costs of degree training.
Source: Oehmke (see Table 1).

sign and implementation of future TDT programs. From Symposium discussions five major observations emerged.

- *Policies affecting the supply and price of agricultural inputs, and the market for agricultural outputs, clearly influence the impact of improved technology.* A dramatic example is Zambia, where government policy stimulated the adoption of an improved maize variety to a degree that appears to have been economically unprofitable and budgetarily unsustainable. In Sudan, the adoption of improved sorghum suffered a setback when government pricing policy changed.
- *Input supplies (including seed and credit) and output markets play a key role in supporting or restraining adoption of productivity-increasing agricultural technology.* Lack of improved seed multiplication and distribution was a critical constraint in Niger and Uganda, as was lack of fertilizer in Zambia. Limited markets for output hindered TDT efforts in Mali and Uganda. By contrast, relatively effective input and output markets in Zambia encouraged the wide use of improved maize hybrids in that country.
- *Many of the TDT programs addressed in the impact studies were implemented in zones with difficult agro-climatic conditions.* The Cameroon and Niger TDT programs, for example, had to develop improved cereal and cowpea technology for areas with low and variable rainfall. In Niger recent droughts reduced impact. Within the targeted zones agro-climatic conditions were diverse; thus an improved technology that was successful in one area produced lower results in another. For example, the Zambia maize study shows that two-thirds of farmers in the best maize zone had adopted improved hybrids or varieties, but only one-third of farmers in the less favorable (low-rainfall) zone had adopted the improved variety. Also, farmers in the best maize zone planted three-quarters of their land in improved maize, compared with one-quarter in the low-rainfall zone.
- *Research organizations and other institutions needed for effective TDT depend on a stable political environment.* The Uganda study illustrates the magnitude of the constraints posed by destruction of the institutional framework resulting from civil unrest.
- *Appropriate priorities, scientific leadership, favorable incentives, and adequate human and financial resources are needed if research systems are to be effective in generating improved technology.* Several country studies (e.g., Cameroon, Kenya, Zambia) showed that a combination of well-funded programs by national and international research centers and donor agencies resulted in the release of improved technology that was adopted by farmers.
- *Maintaining productive research systems with tighter budgets and reduced donor involvement requires more rigorous priority setting (i.e., maintaining adequate funding for fewer research programs) and changes in the incentive structure (e.g., salary, merit-based researcher evaluation procedures) within national agricultural research systems.*

Conclusion

The importance of improving productivity in agriculture as a step toward agricultural transformation merits continued investment in agricultural TDT. Evidence of the positive impact of previous investments, and the beneficial changes in the macroeconomic policy environment in many countries, provide the basis for expecting that future investments in TDT will pay off. The evaluation results indicate that A.I.D. may wish to maintain or gradually increase the amount of funding allocated to agricultural TDT activities in Sub-Saharan Africa.

**Data from 1986 and 1987 are from General Foreign Assistance Funds allocated to Africa, 1988; and later data are from Development Fund for Africa expenditures.*

Additional information on this subject is available from Dr. James F. Oehmke, Department of Agricultural Economics, Michigan State University, East Lansing, MI 48824-1039.

The Symposium on the Impact of Technology on Agricultural Transformation in Africa was held at the Ramada Renaissance Techworld Hotel, Washington, D.C., from October 14 to 16, 1992. The Symposium was organized by Michigan State University and A.I.D.'s Bureau for Africa and Bureau for Research and Development.

Assessment studies referred to in the article are as follows:

Abidogun, A. 1982. Cocoa Research in Nigeria: An Ex post Investment Analysis. *Nigerian J. of Econ. and Soc. Studies.* 21-35.

Boughton, Duncan, and Burno Henry de Frahan. 1992. "Agricultural Research Impact Assessment: The Case of Maize Technology Adoption in Southern Mali." Department of Agricultural Economics, Michigan State University, East Lansing, MI.

Evenson, Robert E. 1987. *The International Agricultural Research Center: Their Impact on Spending for National Agricultural Research and Extension.* CGIAR Study Paper No. 22. Washington, D.C.: World Bank.

Howard, Julie, George Chitalu, and Sylvester Kalonge. 1992. "The Impact of Investments in Maize Research and Dissemination in Zambia: Preliminary Results." Paper presented at the Symposium on the Impact of Technology on Agricultural Transformation in Africa, Washington, D.C., October 14-16, 1992.

Karanja, Daniel D. 1990. "The Rate of Return to Maize Research in Kenya: 1955-88." M.S. Thesis, Department of Agricultural Economics, Michigan State University, East Lansing, MI.

Laker-Ojok, Rita. 1992. The Rate of Return to Agricultural Research in Uganda: The Case of Oil-Seeds and Maize. Department of Economics, Michigan State University, East Lansing, MI.

Makau, B. F. 1984. "Measurement of Economic Returns to Wheat Research in Kenya." Unpublished M.A. Thesis, University of Nairobi, Kenya.

Mazzucato, Valentina, 1991. "Non-research Policy Effects on the Rate of Return to Maize Research in Kenya: 1955-88." M.S. Thesis, Department of Agricultural Economics, Michigan State University, East Lansing, MI.

Mazzucato, Valentina, and Samba Ly. 1992. An Economic Analysis of Research and Technology Transfer of Millet, Sorghum, and Cowpeas in Niger. The Hague, Netherlands: ISNAR/Michigan State University.

Schwartz, L.A., J.A. Sterns, and J.F. Oehmke. 1992. "The Economic Payoff to Agricultural Research and Extension: The Case of Cowpeas in Senegal." Forthcoming, *Agricultural Economics.*

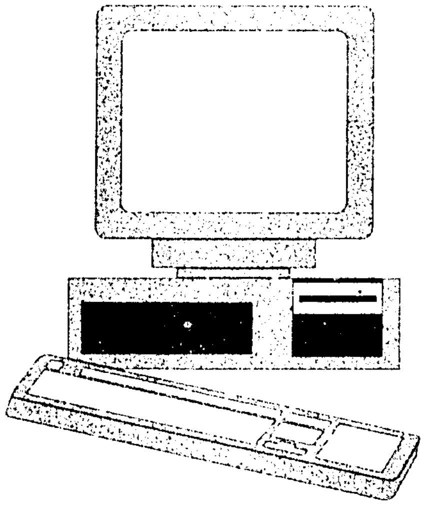
Sterns, J.A., and R. Bernsten. 1992. "Assessing the Impact of Cowpea and Sorghum Research and Extension: Lessons Learned in Northern Cameroon." Department of Agricultural Economics, Michigan State University, East Lansing, MI.

Evaluation News

CDIE's Office of Development Information: Turning Information into Knowledge

by Kaaren Christopherson,
Research and Reference Services,
Office of Development Information

The Center for Development Information and Evaluations's (CDIE) Office of Development Information (DI) has long been known for its work supporting project design and implementation, providing economic and social data, and maintaining the Agency's project and program memory bank. But the questions facing A.I.D. today require more than document delivery and raw data. In the 1990s, A.I.D. needs a thorough, analytical, and balanced base of knowledge on new sectors and countries. DI will help CDIE respond to these new and complex information needs and create new programs for outreach to A.I.D. in Washington and overseas.



Reaching Out to A.I.D Worldwide

In June 1982, the U.S. Government Accounting Office (GAO) published a report recommending that A.I.D. "...require and encourage the systematic and comprehensive identification, recording, and use of experience when designing and implementing projects." In response A.I.D. created the Development Information Unit of the S&T Bureau. Within a year, the new unit was merged with the Evaluation Division of the Bureau for Program and Policy Coordination (now the Directorate for Policy) to form CDIE, which became A.I.D.'s chief source of development knowledge.

DI's Services

A.I.D.'s growing need for a reliable and objective knowledge base has led to an expansion of services within DI. Where once A.I.D. requested primarily document retrieval and reporting services, now it needs much more. DI provides analyses of major issues, locates and synthesizes hard-to-find resources, and cultivates access to new resources, including individuals and organizations in development, to support A.I.D.'s decision-making.

In addition, outreach has become an important service at DI. Since 1988, DI has placed researchers and economists as liaisons with A.I.D.'s regional bureaus to ensure that bureaus continue to integrate current development information in their decision-making. Working closely with DI researchers, reference specialists, economists, and statisticians in Washington, the liaisons provide bureaus with onsite access to economic and social data, as well as information on the development activities of A.I.D., other donors, and private voluntary and nongovernment organizations. DI analysts onsite perform research and analysis on important development issues and help maintain a dialogue between DI main office staff and bureau staff, so that the bureau's information needs can be anticipated and new resources made available to meet them.

In providing economic and social data support to bureaus, DI's focus is on accuracy. To meet the challenge of providing accurate data, the office is continually expanding its network of multilateral organizations to collect, verify, and disseminate a broad range of data. The expertise of DI economists and statisticians in establishing and managing good data series has given the office flexibility to meet

A.I.D.'s changing data needs. Moreover, the bureaus depend on their onsite DI staff to guide them in effective use of data and to create such products as the *Selected Economic and Social Data*, a document produced for the Latin America and the Caribbean Bureau.

A.I.D.'s growing need for development information has required DI to provide increasingly specialized services. For example, in 1989 the Office of Women in Development (WID) asked DI to help organize WID materials and make them more accessible to those interested in WID issues. Today the DI WID coordinator provides the WID office with an array of services, including responding to information requests, writing articles on WID issues, and developing the WID materials collection.

In 1991, the Africa Bureau called on DI to help track activity in democracy and governance on the African continent and to serve as a coordination point for compiling information on other sector activity in Africa. In response, DI established the Africa Bureau Information Center, with its own research and editorial staff, to study and track these issues using a range of resources and to create mechanisms for disseminating information to African counterparts, as well as to A.I.D. staff.

Within the past year, the Capital Projects Office of the Private Enterprise Bureau also called on DI to provide special research support to help assess the capital projects needs of A.I.D.-assisted countries.

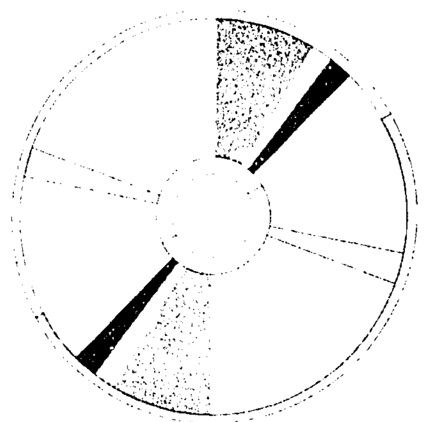
DI also routinely sends specialists in agricultural technical information to developing countries to assist agricultural development officers enhance their knowledge base and assess their information needs. In 1991, technical information specialists visited more than 10 A.I.D. Missions in Latin America, returning with more than 1,000 requests for information assistance. These information specialists maintain close contact with agricultural development officers in the Missions in an ongoing information dialogue.

In September 1992, the Bureau for Private Enterprise, with help from DI, launched the Center for Trade and Investment Services (CTIS). DI trade and investment information specialists at CTIS will offer U.S. businesses current information on investment opportunities and joint ventures in A.I.D.-assisted countries. CTIS staff will provide information about A.I.D.'s private sector programs overseas and A.I.D.-financed procurement opportunities, link U.S. firms with entrepreneurs overseas, and sponsor industry-specific transaction conferences.

Also this fall, the development information office began its first full-scale technical assistance effort to an A.I.D.-assisted region. The Regional Information Clearinghouse (RIC) in Guatemala City, sponsored by the Regional Office for Central American Programs (ROCAP) through the Regional Development Support project, will "marshall existing information resources, harness them to meet regional development information needs, and...emphasize the identification, exchange and dissemination of current topical information" (Project Paper Supplement No. 1). The RIC will serve a select group of regional institutions, key donors, participants in important regional policy fora, and U.S. Government development organizations working in Central America.

Other Services

Research and Reference Services



The Research and Reference Services (R&RS) unit of DI is a team of international development and information professionals, who analyze appropriate resources and synthesize information to provide A.I.D. with the knowledge base it needs for effective decision-making. R&RS facilitates the sharing of resources and networking among members of the development community and gives advice on how to use information in development programs effectively. R&RS researchers answer in-depth questions on development issues, and reference specialists operate the A.I.D. Development Information Center and provide quick reference access to an extensive collection of materials.

Development Information Services Clearinghouse

The Development Information Services Clearinghouse (DISC) provides technical information processing and support services to A.I.D. offices in Washington and overseas. DISC manages the Agency's program for acquiring and processing all A.I.D.-funded documents for inclusion in the A.I.D. Development Information System (DIS), a bibliographic database cataloging Agency projects and programs. The database is available to A.I.D. offices through a CD-ROM (compact disc-read only memory) version called CD-DIS, which was developed and produced by the clearinghouse and which contains more than 65,000 bibliographic records of A.I.D.-funded technical studies and project documents. In addition, DISC provides document dissemination, micrographics, and information management support to A.I.D.

Economic and Social Data Services

Economic, financial, trade, industrial, and social data on an extensive range of topics is available from DI through the Economic and Social Data Services (ESDS). ESDS has acquired more than 30 international-source databases, such as the World Bank's World Tables, the International Monetary Fund's International Finance Statistics, and the Food and Agricultural Organization's Commodity Production and Trade, which form the core resources for ESDS research and analytical services to the Agency.

Technical Inquiries Group

Information specialists of the Technical Inquiries Group (TIG) research and disseminate technical literature for A.I.D. staff working in agriculture and rural development projects. TIG specialists help solve specific agricultural problems, linking project officers with worldwide research results and current information in the agricultural sciences, agricultural economics, and natural resources management.

New Endeavors

DI is furthering the effective use and dissemination of development information worldwide through its participation in an Informal Study Group on the Exchange of Development Information. Composed of more than 90 member organizations, including all Development Assistance Committee members, United Nations organizations, development banks, and nongovernmental organizations, the study group works to initiate mechanisms for sharing development information. One of its recent achievements is the creation of a CD-ROM containing 70,000 records documenting the development activities of member organizations. DI helped spearhead this effort, which entailed identifying common data elements, designing a common format for the information, and producing the compact discs. The current format will be expanded to capture descriptive information and evaluative abstracts, with the goal of creating a complete picture of worldwide development activity in one resource.

DI also continues to refine and develop its own CD-ROM product, CD-DIS. This information tool, available free to A.I.D. offices in Washington and overseas and updated quarterly, can help Agency personnel become more self-sufficient in locating information on lessons learned from projects and programs and make technical information more widely available. More than 200 organizations are using this tool.

As A.I.D. encounters new development problems in the emerging nations of Eastern Europe and the former Soviet Union, DI is pursuing new avenues for acquiring information resources that will help the Agency expand its knowledge base. DI's mission is clear. Whatever endeavor A.I.D. undertakes, DI must anticipate the Agency's information needs and find new resources and create new programs to meet them.

For more information about DI's services, readers should contact Kaaren Christopherson at (703) 875-4994.