BIBLIOGRAPHIC DATA SHEET	1. CONTROL NUMBER	2. SUBJECT CLASSIFICATION (695)
	PN-AAK-702	DA00-0000-G750

3. TITLE AND SUBTITLE (240) Evaluation of impact assessment techniques in Thailand

4. PERSONAL AUTHORS (100)

5. CORPORATE AUTHORS (101)

Richard J. Barber Associates, Inc.

6. DOCUMENT DATE (110)	7. NUMBER OF PAGES (120)	8. ARC NUMBER (170)
1974	127p.	

9. REFERENCE ORGANIZATION (130)

Barber

10. SUPPLEMENTARY NOTES (500)

11. ABSTRACT (950)

Evaluation

Assessments

Models

12. DESCRIPTORS (920)

Project impact evaluation

~~Rural development~~

AID

Thailand

~~Insurgency~~

Investments

~~Community development~~

~~Villages~~

~~Community development~~

Program planning

~~Economic development~~

Methodologies

13. PROJECT NUMBER (150)		
	14. CONTRACT NO.(140) AID/Asia-C-1110	15. CONTRACT TYPE (140)
	16. TYPE OF DOCUMENT (160)	

AID 590-7 (10-79)

EVALUATION OF
IMPACT ASSESSMENT TECHNIQUES
IN THAILAND

A Report

prepared for the

Agency for International Development

under

Contract No. AID/Asia-C-1110

by

Richard J. Barber Associates, Inc.
1000 Connecticut Avenue, N. W.
Washington, D. C. 20036

October 25, 1974

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont'd)

PREFACE

This evaluation was conducted during a period of approximately six weeks in July and August 1974. A team composed of Lee W. Huff, Dail K. Phillips, Thomas C. Thayer, and W. Cody Wilson was assembled by Richard J. Barber Associates, Inc. (RJBA) to undertake field research in Thailand for half that period.* Dr. Jacques Amyot, Director of the Chulalongkorn University Social Science Research Institute, Professor Raymond Tanter of the University of Michigan, and Richard Sharp of RJBA were also involved in parts of the work.

Much of the effort consisted of an intensive interview and discussion program in the United States and Thailand and the report relies heavily on these sources. Scores of people were consulted, but RJBA is especially grateful for the willing and essential cooperation of the Secretary-General of the Office of Accelerated Rural Development (ARD) and the Director-General of the Community Development

---

* Dr. Huff, Vice President of RJBA, served in Thailand for about four years as an advisor on socioeconomic development to the Office of Accelerated Rural Development (1966-67) and to the Mobile Development Unit program (1962-64). Mr. Phillips, who is fluent in Thai, served for extended periods in the 1960's in advisory capacities with ARD and the Community Development Department and handled other development planning and programming tasks for USOM/Thailand. Mr. Thayer is a systems analyst who has specialized in Southeast Asia research and data reporting systems for over ten years, in the region and in the U. S. Dr. Wilson has taught graduate courses in social psychology, social research and statistics at Harvard, Texas, GWU, and CUNY; he was formerly Director for Behavioral Sciences in the Office of the Secretary of Defense (OSD/ARPA) and presently teaches and directs the social research program at Adelphi University.

Department (CD) and their staffs in Thailand, and the American Institutes for Research (AIR).

In attempting to cover such complex subject matter in such a short period of time, it would be surprising if errors of fact and/or interpretation were avoided completely. In order to minimize this possibility, it is recommended that AID/W circulate the report to informed parties in the U. S. and Thailand for comment.

It is assumed that readers have a general familiarity with the administrative structure in Thailand, with Royal Thai Government (RTG) and U. S. development activities conducted there in recent years, and with the U. S.-sponsored contractor research program on impact assessment techniques carried on from 1968 through June 1974.

SUMMARY OF PRINCIPAL FINDINGS AND CONCLUSIONS

1. No Village Impact Assessment "System" is in use in RTG agencies today. Certain impact assessment techniques have been or are in use under specified circumstances.

2. Where confusion arises pertaining to the status and value of the impact assessment techniques, it often is due to problems of semantics and rhetoric. A great deal of the language that has been used by various parties to describe the assessment techniques creates impressions of scope and power that are not matched either by their state of development or their actual use in the field. Similarly, words and phrases such as "investment," "development planning" and "polity strength" are especially value-loaded and tend to invite a host of implicit assumptions on the part of observers that are often quite misleading when attempting to understand the assessment techniques.

3. The basic conceptual model from which the assessment techniques spring -- Development Inputs create Opportunities which promote Villager Investment Behaviors which serve to increase Polity Strength -- is plausible, but it remains essentially an assumption or research hypothesis. Almost no evidence exists today either to confirm or to deny the validity of the model.

4. The conceptual model retains a strong counter-insurgency orientation, thus raising questions for AID to answer as to: (a) the presumption that development has anything to do with insurgency

or "contest situations," and (b) AID's interest in attempting to measure development impact in terms of allegiance to governments or governmental systems. In brief, is this model the most significant or relevant model of the development process that AID would now like to see tested further?

5. Several significant methodological problems affect the impact assessment techniques currently in use in RTG agencies. In essence, the degrees of technical "reliability" and "validity" achieved with respect to the instruments developed to try to measure important constructs of the model vary from unacceptable to satisfactory, even for research purposes. They are not satisfactory for making routine operational decisions, where more stringent standards of reliability and validity are warranted.

6. Use of impact assessment techniques in the Office of Accelerated Rural Development and the Community Development Department is peripheral to the main planning and programming systems in those agencies and affects a very small proportion of the resources allocated by either agency.

7. ARD has used an application of the impact assessment techniques as an aid to planning the location of infrastructure projects in villages, when called upon to participate in security support operations. (There have been three such occasions: one completed, one in process, and one pending.) ARD has not used them for program evaluation. CD has applied a version of the impact assessment techniques, on one occasion, for monitoring or scorekeeping purposes.

8. Given their present state of development, and the technical capabilities extant in RTG agencies, this evaluation recommends against continued use of the impact assessment techniques for routine operational purposes in Thailand.

9. Published information concerning the impact assessment techniques is difficult to use and subject to misinterpretation.

10. The applications of impact assessment techniques now in use in Thailand should not be transferred to other LDC planning agencies for operational use.

11. The model of the development process which was used in the Thailand research; the idea of measuring intermediate variables that intervene between Development Inputs and Ultimate Outcomes; and the method used to develop instruments (indexes) to measure concepts in the model, are transferable to other LDC's, should AID decide that they are of sufficient substantive interest to warrant doing so. Specific instruments, such as the Opportunity Index and Investment Index, are not transferable. A research and development program would be required to develop an "impact assessment system," in Thailand and in other LDC's.

I

STATEMENT OF THE PROBLEM

AID/Washington asked Richard J. Barber Associates, Inc. (RJBA) to assess the current status of what it called the Village Impact Assessment System (VIAS). This system had been developed by American behavioral scientists in the course of a six-year research program in Thailand and was intended "for use by Thai rural development agencies when planning the location of project activities and evaluating the impact scored by those projects."* AID had questions about the extent to which the system had actually been adopted and applied. As the PIO/T put it:

> This situation needs to be assessed. Has the system been adopted by Thai administrators in greater detail than appears on the surface? If it has not been adopted or is only partially adopted what are the explanations?
>
> Questions of this character require an answer because there are development assistance agencies, as well as contractors, who insist that the system ... is now replicable and useful for other countries, perhaps in some modified form.

Accordingly, AID/W directed RJBA to apply two tests in assessing the VIAS: (1) the system's ability to enhance the capacity of LDC rural development agencies to determine the suitability of alternative rural area locations for "development," i.e., to improve resource allocation decisions, and (2) the system's ability to enhance the capacity of LDC rural development agencies to

---

* PIO/T, p. 6.

measure the results or impacts of project activities, in terms of behavioral changes in the target population. (It was intended that the impact information obtained by the VIAS be of sufficient quality to feed back into and further improve the next generation of resource allocation decisions.) In applying these tests, RJBA was given three principal categories of issues to address: conceptual and methodological soundness, actual use of the techniques in question by Royal Thai Government (RTG) agencies, and the potential for transferability to other LDC's.

In evaluating the conceptual aspects of the VIAS, "as it presently stands," for LDC development agency planning purposes, RJBA was asked to:

1. Review the basic concepts relevant to operational use of the system.

2. Examine methodological issues associated with implementation of those concepts.

3. Assess the extent to which the system has been tested and validated.

The description of the uses of the VIAS or variants of it in RTG agencies was to deal with the following:

1. The role of the VIAS in RTG agencies' planning and evaluation systems.

2. Constraints which might condition the VIAS as a planning and evaluation technique.

3. Problems encountered by RTG agencies with the VIAS, and steps needed to resolve those problems.

4. The potential for further development and/or broader use of VIAS in Thai operating agencies.

If the findings regarding conceptual matters and RTG use demonstrated that VIAS could meet either or both of AID's basic tests, we were to "assess the extent to which the relevant concept today and its operationalization as a technique of planning and measuring impact in rural development may be transferable to other LDC's."

II

GENERAL OBSERVATIONS

Before proceeding to the Team's specific findings, some general observations are in order. First, the phrase Village Impact Assessment System is something of a misnomer. "VIAS" is not used by any of the principals, Thai or U. S. Nor do they use the phrase Impact Assessment Methodologies (IAM), which appears in Michael Dwyre's informative May 1974 USOM paper entitled The ARD Blue Book. There was general agreement among those with whom we spoke that no "system" or integrated set of methodologies is in use. The RJBA Team concurs in this judgment. AIR concluded, "Although it cannot be said that such a system exists it is possible to consider the current impact assessment applications as predecessors of a coordinated system yet to be developed."* Communication about the assessment techniques normally focuses on particular techniques or measuring instruments which are of interest to one group or another, with the most common descriptor being the Opportunity Index or the Investment Index, or simply the O-I Indexes. The terms PAM-I and RIAF are also used within ARD. To repeat, the assessment techniques are not viewed in Thailand, conceptually or practically, as a system. In this report, the phrases "impact assessment techniques" and "assessment techniques" are used instead of VIAS.

---

* AIR, Impact Assessment Handbook, June 1974, p. VI-6 (hereinafter referred to as Handbook).

Second, there is a significant lack of explicit documentation concerning the basic ingredients of the assessment techniques, decisions taken during their design and evolution, and their usage in RTG agencies. Virtually no written material is available from Thai or U. S. agencies. The U. S. contractor group published a number of reports covering six years of R&D activity, but they do not include all the information needed to answer the questions raised in AID/W's PIO/T. The Final Report, especially, is inadequate in its documentation of crucial issues. Apparently none of the sponsoring U. S. agencies -- OSD/ARPA, Embassy Bangkok, or USOM -- undertook serious in-house technical reviews of the research while it was in progress. This finding, however, may in itself be of some value to AID. The lack of documentation has hampered the Team's evaluation efforts.

Finally, we cannot overstress a strong impression that much of the confusion surrounding the assessment techniques is rooted in semantics and rhetoric. When words and phrases such as "systems," "economic development planning," "investment," "resource allocation," "counter-insurgency," "polity strength," and "the development process" -- to name only a few -- are bandied about, perceptions of what they mean vary considerably in the eye of the beholder. Much of the language used by various parties to try to describe the assessment techniques and their use in Thailand, in writing, does tend to conjure up (especially in Washington) visions of scope, power and influence which simply are not matched by their actual state of development

or their use in the field. The areas of relevance to which these techniques can be applied are much more restricted than is often asserted or implied by the language used to describe them. Part of the problem is probably due to the tendency common in government-sponsored research to discuss and focus on a given research project more or less in isolation, without relating it too closely to the broader set of conditions to which it must ultimately relate; in this instance the status of economic and social development in general in rural Thailand is especially pertinent.

We wish to observe at the outset that signs and evidence of the behavioral research work pioneered in Thailand by AIR are clearly identifiable within Thai government agencies, a claim which few, if any, foreign social science research groups can make. While this report is not addressed directly to the serious problems of undertaking research in a Thai administrative setting which historically has been resistant or simply indifferent to the concept of research, and especially social science research, it must be noted that AIR's achievement in "making a dent" is unique.

The remainder of the report is structured as follows. Section III presents some observations about the basic conceptual model which was designed for the research program and served as the guiding force for development of impact assessment techniques. Section IV is a summary of methodological issues. We have attempted to keep it as concise as possible, but the subject matter is very difficult to present in shorthand form. It will be seen that a

number of methodological problems serve as important constraints on effective use of the impact assessment techniques for operational purposes. Section V, supplemented by an Appendix, provides a description of the ways in which RTG agencies have used impact assessment techniques. Sections VI and VII discuss the prospects for such techniques in Thailand and other less developed countries, in line with the requirements of the PIO/T.

III

CONCEPTUAL SOUNDNESS

The assessment techniques relate directly to a concept that the contractor articulated in the form of a model or construct.* The technical relationships between the basic construct and the techniques used to measure it are discussed more fully in Section IV. Suffice it to say here that the model in essence postulates the following: "Development Programs" tend, along with natural resource endowments, to create "Opportunity," which in turn, mediated by certain "Disposing Conditions" such as village size or wealth, promotes "Villager Investment Behaviors" which ultimately serve to increase "Polity Strength." Increased investments, whatever their nature, are asserted to give villagers a stake in the system and a set of mutually rewarding dependencies is established between government and people. Polity Strength is the positive social condition which the development programs, over time, are expected to produce on a nationwide basis.

Since measurement of so comprehensive an "ultimate outcome" is an exceedingly long-term proposition, the model asserts that increased villager investment behavior can serve adequately as a shorter-term proxy or indicator of progress toward the final objective. Thus the investment behaviors -- defined as "the extent to

---

* The model has been described in many AIR research reports. Three of the most recent are: Systems for Evaluating the Impact of Rural Development Programs, Final Technical Report, Contract AID 493-037-T, June 1974; "Disposing Conditions for Development Impact in Rural Thailand," Technical Report, April 22, 1974; and "Measuring Village Commitment to Development," May 1973.

which the residents of a given village are investing their own resources in the lawful opportunities for economic, social, and political improvement that are available to them" -- and the opportunities from which these behaviors tend to spring, become the central elements in the model. And indeed, they have attracted far and away the most attention of Thai and U. S. practitioners who have been exposed to the model. The innovative characteristic at issue in pursuing this line of inquiry is the prospect that one can measure the actual "behaviors" of target populations in response to development programs and projects and thus obtain a more dynamic indicator of progress or impact than can be gleaned from the conventional procedure of cataloging physical measures such as wells dug, miles of road built, quantity of medicines dispensed, etc.

The model is entirely plausible and it has a certain intuitive appeal, especially to hard-pressed development planners and programmers in search of more discriminating clues about the success or failure of the actions they undertake. But it is crucial for the purposes of this evaluation to understand that the model is simply a research hypothesis or assumption to be tested empirically against reality. Despite six years of research and test, very little evidence presently exists to confirm or deny its validity. No evidence has been presented to show (or not show) that a relationship exists between Opportunity or Investment and the consequent Polity Strength which the model hypothesizes will result or, for that matter, with any other outcome. There is some very modest evidence that

Opportunity and Investment related to Development Inputs such as the number of ARD projects completed and the amount of RTG funding for development projects in a village, but the size of the relationship is such that only about 5 per cent of the variance in Investment is accounted for by these development inputs (indeed, size of village accounts for much more of the variance in Investment). In our judgment, this does not constitute sufficient confirmation of the theoretical model for use as a guide to program and resource allocation decisions by an operational agency.

The absence of confirming or denying evidence is hardly fatal in a research project, but it becomes crucial when consideration is being given to use of the model in an operational setting. We would like to re-state these points for emphasis because, despite their simplicity, they are often missed by persons who encounter the work on assessment techniques. As a result, incorrect assumptions are made implicitly and confusion results. To repeat, the research program did not include a substantial effort to list and match up directly ARD's development programs (or those in CD or any other agency), or even development inputs generally, with Opportunity (O) and Investment (I) measures. The connection between "programmed inputs" and O and I was, and remains, largely presumed. No effort has been directed to establishing an empirical connection between O/I and Polity Strength.

The point then is that the Thailand research has not settled the question of the basic validity of the model one way or the other.

The research work to date has focussed instead on the methodological questions involved in measuring a particular conception of something called Opportunity and something called Investment and in attempting to make such measurements (discussed below). Consequently "off the shelf" transference of the assessment techniques to LDC planning and operating agencies for routine use could not be done on grounds that the "development process model" involved has been proven. The case for transference would have to be based on a belief that additional R&D was warranted, and use of the techniques would most properly be vested in agencies equipped with a capability to conduct R&D activity.

This raises a more fundamental issue for AID decision-makers: Is "polity strength" the most significant "ultimate outcome" of the development process insofar as AID is concerned? The model being used in Thailand was conceptualized in the "preventive counter-insurgency" atmosphere of the late 1960's. U. S. programs assisting the Thais then were intended to reduce or remove socioeconomic conditions thought to be conducive to insurgency. The dominant management issue of the time was determining how well one was doing via social and economic development action programs to convince citizens, normally in rural areas, that they had a future within the existing polity. Indeed the central concern was to have "impact," measured in terms of answers to such questions as does the villager know that activities are being undertaken for his benefit; does he like what is being done for him; is he aware that the "authorities" (whomever they may be) are genuinely concerned about his welfare and that

"government" (however expressed) or the "system of government" is responsible for undertaking action programs that affect him? Ultimately, of course, one wanted to know if the villager would resist the blandishments of the insurgents more effectively as a result.

The contractor group commenced its research work in Thailand in this environment. Although it concluded in the winter of 1970 that "countering insurgency" was too negative and too restrictive an objective, and substituted instead the achievement of "pro-polity strength" (achievement of which, it was asserted, would also serve as an antidote to insurgency if such a condition were to arise), it is the RJBA Team's judgment that the model and its workings still retain a strong strain of this "hearts and minds" flavor. A recent (April 1974) technical paper, for example, states the assumption in explaining the model that "rural development projects can have a positive effect on the degree to which rural people are willing to support the existing system of government in a contest situation."* Similarly, the Impact Assessment Handbook (June 1974) speaks of "the potential of well done development projects to 'innoculate' neutral or already-loyal villagers against potential loyalty to the other side," and states that "It is assumed that continuous villager investment behaviors if rewarded tend over time to create satisfactions, dependencies and habits, which in turn influence the villagers to support the existing system of government and actively resist

---

* AIR, "Disposing Conditions for Development Impact in Rural Thailand," Technical Report, April 22, 1974, p. 2. Emphasis in original.

attempts by an outside force to destroy it."* Coincidentally or otherwise, ARD currently only uses the O/I Indexes for planning in support of counter-insurgency operations in security-sensitive areas; it does not use them for planning ARD programs in its so-called "economic development Growth Areas" or in entering new provinces with its traditional panoply of activities.

This in turn raises two questions. The first concerns the validity of the assertion that development has anything to do with "contest situations," specifically insurgencies. Some of our Thai informants, in and out of government, were prepared to assert that it did not. In their view development might or might not be worth pursuing, but it should not be assumed to have much to do with the propensity of villagers to embrace or ignore insurgent appeals or reach judgments about the value of their system of government. The latter questions were felt to be more matters of the style and quality of government-citizen interactions on a wide range of matters. It was further asserted that development programs, and public relations campaigns proclaiming the government's willingness and/or obligation to carry out development activities, tended, if anything, to have a negative effect because they stifled natural tendencies toward self-help and exacerbated tensions by building up false hopes which most LDC governments simply could not satisfy quickly enough. In this view, the development of villager-government "dependencies"

---

* Handbook, pp. III-1 and I-4.

is seen as potentially destabilizing. It is not our concern to comment on the validity of this set of views, but it is important to note that it exists in Thailand. If it were accepted as valid, its proponents would argue that measurements of O and I have little or no meaning.

The second question, and a more important one for AID, is whether AID, as a development assistance agency today, is particularly concerned with "contest situations" at all. Does AID, in its basic development strategies and programs, and those it would recommend to LDC's, most want to measure development impact in terms of allegiance to governments or governmental systems? Or is it seeking to measure "impacts" that are somewhat less concerned with direct or indirect expressions of loyalty, popularity, etc.? Is "Polity Strength" an adequate analogue for currently fashionable qualitative development objectives such as income redistribution, reduction in unemployment through promotion of labor-intensive technologies, alleviation of poverty, etc., or even for more traditional goals such as increased rates of GNP and the hoped-for "trickle-down" benefits which are presumed to result therefrom? If a different "ultimate outcome" were postulated, would the Opportunity and Investment Indexes be the most suitable measure for it?

In brief, given the Thailand model's conceptual roots in the counter-insurgency context, and the development communities' highly fluid outlook on the nature of the development process, AID should examine the substantive merit of the research hypothesis used in

Thailand, placing special emphasis on the current policy relevance of the asserted ultimate outcome, namely, Polity Strength. The first level of discussion and debate within AID should address the elementary question of whether the concept hypothesized is in fact the most significant or relevant model of the developmentprocess that AID would now like to see tested.* Ideally, of course, the research design would have included some rival hypotheses for test.

In Thailand today, the CD Department has discarded the "Polity Strength" outcome and ARD has simply ignored it. As a practical matter we sense that many persons involved with or exposed to the assessment techniques, in Thailand and in the U. S., are fascinated with the O and I Indexes quite independent of any presumed links that these measures may have with something called Polity Strength. Most people definitely convey the feeling that measuring investment behavior in particular must somehow be "good," or valuable, in and of itself.** Investment connotes progress, something obviously very desirable in the minds of development planners and advocates of social change. In fact, it is practically impossible for anybody

---

* It is noteworthy that apparently none of the government sponsors of the research in Thailand, in six years, proposed alternative "ultimate outcomes" for test.

** Those who do not look deeply into the components of the Investment Index read whatever "investment" means to them normally (usually with a heavy or exclusive economic content) into it and probably embrace it too uncritically; many who do look carefully become disenchanted too quickly, largely through a failure to appreciate what an Index is supposed to do. These questions are discussed more fully below in the section on methodology.

to be against "Investment," or for that matter "Opportunity." These are value-loaded words and most people appear to have a difficult time questioning their validity. In any event, the measurement of something called Investment Behavior has de facto become the "ultimate objective" in the minds of many.

Even if the measurement of something called Investment Behavior alone is accepted as the ultimate objective, the work done in Thailand does not automatically provide an "off the shelf" assessment technique. There are still issues of the technical adequacy of the measurement techniques -- in terms of reliability and validity -- and the question of the "culture-boundness" of the product produced in Thailand. These issues are discussed more fully in Section IV.

The fundamental question, of course, remains: Does Villager Investment Behavior, however measured, tell you what you most want to know about the impact of economic and social development programs on rural populations? The answer to this question should not be left to researchers. Indeed such a polity decision is properly the province of development policy-makers, planners and operators.

It may provide some perspective to address the point mentioned in Section II above concerning the problem of relating a specific research project to the world around it. It is common for economic planners in Thailand to argue that "nothing statistically reliable is known about rural Thailand." This is a sweeping indictment -- perhaps overdone -- and very likely reflective of the current feeling of hopelessness which many of them feel when pondering the

sobering fact that the relative shares of income and production of the rural regions outside Bangkok and the Central Plains areas have actually decreased over the last decade and a half. Despite the impressive national growth rate, and a concentration of planning effort and U. S. aid in rural areas (especially the Northeast), regional disparities are felt to have actually increased over the period 1960-1972.

The Northeast, for example, is estimated to have a growth rate less than one-half the national average. In 1960 its per capita income was about 56 per cent of the national average and about 36 per cent of the Central Plains' share. By 1972 the figures had fallen to an estimated 46 per cent and 26 per cent, respectively.* In 1960 the Northeast accounted for almost 20 per cent of total GDP; by 1972 it produced about 15 per cent. Northeast residents are also alleged to receive less than half the public services provided to the "average" Thai citizen.

Economic planners in this depressing environment, who also feel that they lack even basic descriptive data of any reliability regarding such essential matters as migration, farm size, yield by farm size, income levels, land tenure, etc., thus tend to view efforts to document relatively minute, highly discriminative changes in Villager Investment Behaviors, at the individual village level, as pointless, if not wasteful of resources. They argue that higher priority, albeit elementary, development research and planning

---

* All figures cited in this paragraph are drawn from recent NESBD materials.

tasks remain unfulfilled. Given the alleged gaps in simple basic economic information about rural Thailand, and the feelings of frustration over lack of knowledge about the forces of development at work in the rural areas, can one expect the impact assessment techniques to explain much about development in Thailand or to guide development planning and programming properly? Most of the economists would respond that one cannot, i.e., without a clear understanding of where you are, how can you attempt to document or interpret "change" at the micro-level? And, they add, are not the natural forces of development, expressed positively or negatively, so strong as to overwhelm or swamp any presumed causal connection between village project inputs, villagers' individual investment decisions, and ultimate Polity Strength? (The contractor's finding that the size of the village has a stronger relationship with the level of investment than do the development input variables studied is perhaps a case in point. Indeed, as noted, size of village accounted for several times the amount of variance in "Investment" that was accounted for by the development input variables.)

The connection, conceptual and otherwise, between the economic development planners' outlook and the point of view expressed by the behavioral scientists who developed the assessment techniques, has not been made. This simply reinforces the point that any assessment of the "conceptual soundness" of the techniques must start with an examination by AID of its conception of the development process and the outcome or outcomes assumed to result from it.

IV

METHODOLOGICAL ANALYSIS

Technical adequacy and soundness are crucial factors for any proposed impact assessment technique. In this instance, methodological issues loom large in reaching decisions about the use and transferability of the assessment techniques developed in Thailand. This section presents a relatively detailed discussion of these issues. The basic source document is the AIR Final Technical Report to AID dated June 1974, hereinafter cited as Final Report. Other technical reports and progress reports have been referred to when the documentation of the Final Report is incomplete.

The section begins with a brief discussion of the Conceptual Model that was the focus of the impact assessment research program and then briefly considers the measurement model that was the basis of the attempts to measure the intervening concepts postulated in the conceptual model. There follows a rather long analysis of the evidence available regarding the technical adequacy of the Investment Index and a shorter discussion dealing with the Opportunity Index. These analyses consider such topics as the reliability, validity, and scoring norms for both Indexes. The section also reviews the evidence produced by the research program that may be used empirically to test the conceptual model and addresses the degree to which the theoretical model has been verified. Finally, the section reviews the data produced by the impact assessment research program in terms of conclusions drawn by the impact assessment researchers and evaluates these conclusions in light of the assembled evidence.

Conceptual Model

As sketched out in Section III, a rather straight forward theoretical or rational model served as the basic assumption guiding the development of the impact assessment techniques: Development Programs lead to Opportunity which leads to Investment which in turn leads to Polity Strength. The model is a plausible, if over-simplified, starting point that has considerable intuitive appeal.

Three aspects of the model should be noted. First, it is a simplification of reality, i.e., things other than Development Programs are likely to affect Opportunity, things other than Opportunity are likely to affect Investment, and Investment is likely to have consequences other than or in addition to Polity Strength. Second, it postulates two intervening constructs -- Opportunity and Investment -- which may be relatively easy to observe and which are hypothesized to provide a connection between the constructs called Development Programs and Polity Strength that are of central interest to policy makers and administrators. And third, the model is an assumption and the relationships that it postulates should be tested against reality by empirical investigation before it is used to guide policy.

The testing of a conceptual model is accomplished by making empirical observations of the constructs using measuring instruments that have been demonstrated to be reliable and valid, and then comparing the relationships among the empirical observations with those postulated by the model.

At one time the contractor asserted with respect to the research program that its "basic achievement was conceptual; given the rationale, the instrumentation (development of the Index) was a relatively straight forward matter for which existing technology was quite adequate."* Nonetheless, AIR expended considerable effort in attempting "to develop valid, reliable, and efficient measures of the concepts represented in this rational model" and eventually concluded that the Index of Investment had indeed become "the central product of the project."** As noted above, this Index has drawn the lion's share of attention and interest from planners and operators, especially within the RTG. It is thus important to understand the nature, strengths, limitations, and outcomes of the effort to develop measures of the concepts represented in the model, especially the Investment and Opportunity Indexes.

The Measurement Model

In attempting to develop "valid, reliable, and efficient measures" the researchers used a widely accepted measurement model with a well defined methodology, namely, the "psychometric" or "summated ratings" model.

This model assumes that a number of empirical "indicators" of the construct of interest exist. Each of these empirical indicators is assumed to reflect partly that which it has in common with the

---

* AIR, Report of Progress, January 15-July 14, 1972, p. 2.
** Final Report, pp. 19 and 4.

other indicators -- presumably the construct of interest (specifically, in this instance, the constructs Investment and Opportunity) -- and partly something that is idiosyncratic to that particular indicator. If one sums across several indicators, the sum will reflect primarily that which is common to all the indicators -- presumably the construct of interest -- whereas that which is idiosyncratic to each indicator will "cancel out" and not be reflected in the summed score. Of course, if the several indicators have something else in common other than reflecting the construct of interest, then the "index" formed by summing across the several indicators may reflect that "something else" primarily rather than the construct which it is attempting to measure.

The measurement problem is to select a representative sample of the total population of indicators for the construct of interest, to ascertain that the sample of indicators has the construct of interest in common and is free of other extraneous common attributes, and to demonstrate that the summated score across the sample of indicators is reliable or consistent.

The Investment Index

Several different versions of an "Investment Index" were tried out during the research program. However, the Final Report focuses principally upon a 23 item "Investment Index" that was used in the later applications phase of the project and is the version now being used by ARD. The Community Development Department used a "similar"

23 item index in a recent study but reports having made some unspecified "modifications" in the questionnaire. The following discussion focuses principally on the 23 item "currently operational" version of the Investment Index.

Selection of Items in The Investment Index

Apparently a pool of 250 specific examples of "investment behavior" was compiled based on a review of the descriptive literature on rural Thailand and from interviews with knowledgeable persons.* This pool of 250 examples of investment behavior, which has not been published, presumably formed a definition of the construct investment. From this pool a set of 23 indicators was eventually selected to form the Investment Index. Each of these items is correlated moderately with the sum of the other 22 items and, therefore, each is reflecting something that it has in common with the others -- presumably the construct Investment.**

The basis for the selection of the final set of 23 indicators which comprise the Investment Index has not been explicitly described, but one can infer that they were selected primarily on the pragmatic basis of the ability of certain classes of informants (Village Headmen and CD Workers), to report on these particular indicators, rather than on the basis of the indicators necessarily being representative of the pool of potential indicators that defined the construct.***

---

* Ibid., p. 21.
** Ibid., Table 3.12, p. 37.
*** Ibid., pp. 22 and 29.

That is, the final set of items was not produced by a winnowing out on the basis of an explicit theoretical rationale.

Several researchers knowledgeable about rural Thailand have criticized a number of the specific individual items included in the Investment Index. Some of these criticisms reflect a failure to understand the measurement model being used, but often they are aimed at the degree of representativeness of the item or at the degree to which the item can possibly reflect the construct Investment.

These issues arise again in considering the "validity" of the index and are discussed more fully below. At this point we simply note that documentation regarding both the degree of representativeness of the items in the Index and the fact of the existence of criticism of the extent to which the items reflect the construct Investment, is lacking.

Reliability of the Investment Index

Reliability is concerned with the degree to which a measuring instrument provides a consistent result. If an index is not reliable -- i.e., if it is not consistent -- it is unlikely to be valid or useful. In evaluating a measuring instrument one usually examines three different kinds of reliability: (1) internal consistency reliability, or the degree to which different parts of the same measuring instrument provide similar results; (2) test-retest reliability, or the degree to which two measurements made with the same instrument at different times, relatively close together, produce similar results; and (3) alternate-form reliability, or the degree

to which two independent, but equivalent, versions of a measuring instrument produce similar results. Reliability is usually expressed in terms of a reliability coefficient -- a two place decimal fraction -- that is a coefficient of correlation.

General guidelines exist in the fields of sociometrics and psychometrics regarding the adequacy of reliability of measurements, but these guidelines are modified for certain uses. The general criteria of acceptability for reliability coefficients are: .90 and above is good (i.e., over 80% of the variance in the two sets of scores produced by the measurement process is common or shared variance); .80 to .89 is acceptable (roughly two-thirds to four-fifths of the variance in the two sets of measurements is shared); .70 to .79 is marginal (less than two-thirds, but more than one-half of the variance is common); and less than .70 is not acceptable (less than half of the variance in the two sets of measurements is common variance).

These general guidelines may not be strictly appropriate for particular uses of a measuring instrument, however, and they are modified upward or downward. In general, the more important the decision that is to be made on the basis of the measurement the more rigorous the criterion of acceptability for the reliability of the measuring instrument, because lack of reliability means impreciseness of measurement and a greater liklihood of misclassification based on the measurement. Therefore, for example, a higher reliability would be required for operational use of a measuring instrument than

for research use. Similarly, a more reliable instrument would be required for placing individual units in classifications that would receive differential treatment on the basis of classification than would be required for detecting average differences between large groups of units.

Internal Consistency Reliability.

Split-half reliability coefficients for the 23 item Investment Index are available for 13 separate data collections.* The median coefficient is .82 within a range from .74 to .90.** Thus the internal consistency of the 23 item Investment Index is within the marginal and acceptable categories which are generally used in the behavioral sciences for research purposes. Therefore, evidence exists to support an Investment Index, composed of 23 items reflecting something in common, that has a marginal to acceptable degree of internal consistency when scored by summing across the 23 items. The RTG operating agencies (CD and ARD), however, have been provided a scoring system which does not sum across the 23 component items, but rather is produced by summing across three sub-index scores (called Political, Economic, and Social), each of which is given equal weight although each is composed of different numbers of items.*** No information has been published

---

* "Split-half reliability" is estimated by calculating one score based on half the items in the index and another score based on the other half of the items, computing the correlation between these two scores, and then correcting this for the reduction in number of items in the full scale.

** Final Report, Table 3.13, p. 38.

*** See Handbook, Appendix B.

regarding the internal consistency reliability of an Investment Index calculated by this currently employed scoring method or about the potential effects of this scoring method on the technical characteristics of the Investment Index. Thus one cannot really judge the reliability of the instrument as it is being employed currently in the field.

Test-Retest Reliability.

No information is available concerning the test-retest reliability of the Investment Index -- that is, its consistency over short intervals of time. There has been little indication, however, that the results of this measurement process might vary over short intervals. Indeed, data are presented that indicate that Investment Index scores are remarkably consistent over much longer periods of time; the correlations between scores obtained one year apart are approximately .70, and those obtained two years apart have an average of approximately .80.* These figures are comparable to the internal consistency reliability of the Investment Index reported above; however, such consistency over time is quite unusual. Possible interpretations will be discussed below.

Alternate-Form Reliability

Information about alternative-form reliability is rather crucial. During the research program, AIR tried out several alternative sources of information upon which to base Investment Index scores.

---

* Final Report, Table 3.18, p. 48.

Most of the data were collected using one of two preferred information sources -- the Community Development Worker (CDW) assigned to a village or the Village Headman. These two sources apparently were preferred primarily because of the economy and efficiency of collecting data from them. Published accounts, however, do not differentiate between Investment Indexes based on these two different sources.* Further, norms based on CDW information were used to transform raw scores based on Village Headman information into standard scores.** Thus, it is implied that the Investment Indexes derived from the two sources of information are comparable.

Unfortunately, the data do not support such an implication. In at least two studies, data on the 23 indicators comprising the Investment Index were collected in a limited number of villages from both the CDW and the Headman. The data from these two independent sources for a sample of villages thus may be compared to provide an estimate of the alternate-form reliability coefficient.

One of these studies, using a sample of only 15 villages, obtained a correlation of .54 between the two 23 item Investment Indexes, one based on information from CDW's and the other based on information from the Village Headmen, collected at roughly the same point in time.*** The figure of .54 is, of course, far below the marginally acceptable criterion of .70.

---

* See, for example, Final Report, Tables 3.13 and 3.14, pp. 38-39, or Table 3.17, p. 46.

** This practice is not explicitly described in the Final Report. For further discussion of scoring norms see pp. 40-44 below.

*** Ibid., Table 4.2, p. 71.

The second study collected data on the 23 item Investment Index from both CDW's and Headmen in approximately 130 villages; however, the results of this data collection have not been reported.* Given the exceedingly low alternate-form reliability data reviewed in the immediately preceding paragraph, it is regrettable that such low priority was given to analyzing and reporting these data.

Two other sets of data provide, rather indirectly, information relevant to the issue of alternate-form reliability. One study conducted relatively early in the research program, using a sample of 50 villages, obtained a correlation of .37 between a 29 item version of the Investment Index based on information from CDW's and a 33 item version of the Investment Index based on information from Village Headmen collected at roughly the same time.** It cannot be determined how many overlapping items were in these two versions, but the results certainly do not support the idea of high alternate-form reliability.

The other study collected data on the 23 item Investment Index on 93 villages at two points in time, using information furnished by CDW's in the initial data collection and information provided by Village Headmen in the second data collection 15 months later. The correlation between these two sets of Investment Index scores was .68.*** This figure is considerably higher than the other estimates

---

* AIR, Report of Progress, July 15, 1972-January 14, 1973, p. 18.

** Final Report, Table 3.10, p. 35.

*** Ibid., Table 3.18, p. 48.

of alternate-form reliability and approaches the criterion of marginal acceptability, namely, .70. Interpretation of the figure is complicated, of course, by the fact that the two data collections were not concurrent, but were separated by an interval of 15 months.

All these data, taken together, do not support a conclusion that the Investment Index based on information provided by a CDW is equivalent to an Investment Index based on information provided by a Village Headman. Rather, quite different, non-comparable results are derived from the two sources of information. Nevertheless, the Final Technical Report does not differentiate between these two "Investment Indexes."*

Summary and Interpretation of Reliability Data

In summary, available data indicate that the 23 item Investment Index which is now being used in ARD has marginal to acceptable internal consistency reliability if summed across the 23 component items (but unknown reliability when scored with the scoring system in use in ARD), unknown but probably acceptable test-retest reliability, and an unacceptable level of alternate-form reliability. That is, both the Investment Index based on information from Community Development Workers and the Investment Index based on information from Village Headmen are internally consistent, are probably consistent over time, but are not comparable to each other.

---

* See, for example, Table 3.13, page 38.

A probable interpretation of these facts is that the 23 indicators that make up the Investment Index, rather than having a single thing -- the construct Investment -- in common, have two things in common: the construct Investment and a single source of information. Thus, the summed score across the indicators, the Investment Index would reflect not only the construct Investment, but also the single information source's global judgment of the village. This would produce internal consistency and consistency over time within each of the two versions of the Investment Index -- one based on information from CDW's and the other based on information from Village Headmen -- but lack of comparability between the two versions. Indeed, the relative sizes of the three kinds of reliability coefficients suggests that the Investment Index is reflecting the global judgments of the source of information more than the construct Investment, i.e., it tells us more about CD workers or about Headmen that it does about villager investment.

This interpretation of the data on reliability in turn raises questions which are relevant to the issue of "validity" -- the degree to which an index actually measures what it purports to measure. We now turn to the question of the validity of the Investment Index.

Validity of the Investment Index

The fields of psychometrics and sociometrics usually consider evidence regarding validity under three rubrics: (1) face validity, (2) criterion validity, and (3) construct validity. The test of face

validity is whether or not the indicators, on their face, reflect the construct that the index purports to measure. The test of criterion validity is whether or not the index correlates highly with an independently derived criterion that is accepted as a valid measure of the construct, i.e., the index may be developed as a more economical substitute for a measure which is accepted as valid but is more expensive and time consuming to use. The test of construct validity is whether or not the measure relates to valid measures of other constructs as predicted by an independently substantiated theory or model. The same data cannot be used both to test the validity of an index and to test a model against reality because each of these tests assumes that the other has been demonstrated.

Face Validity.

The construct Investment was defined as expenditure of time, energy or other resources for economic, social and political improvements, and was further operationally defined by assembling a list of 250 behaviors as exemplars. Inquiry into face validity of the Investment Index involves determining the extent to which the 23 indicators used in the index are representative of the 250 exemplars that operationally define the construct, and the extent to which the 23 indicators reflect the verbal definition.

The 250 exemplars of the construct are not listed in any generally available document so it is not possible to examine the degree of representativeness directly; however, some indirect evidence on

representativeness may be adduced inferentially. As noted earlier, an important consideration in the selection of the 23 items used in the present version of the Investment Index appears to have been the ability of preferred informants, i.e., Community Development Workers and Village Headmen, to report on them.* This selection procedure probably produced some degree of bias in the sample of indicators but neither the amount nor nature of the lack of representativeness can be specified.

The items used in the Investment Index are presented in Table A. Several people who are knowledgeable about Thai villages and the development process have raised questions about whether a number of the indicators really reflect levels of Investment. Some of this criticism of individual indicators may have been the result of a lack of understanding of the measurement model, but some of it seemed to be legitimate questioning of the face validity of the items. The criticism involved two main points: first the possibility that a given indicator was in reality an indicator of a lack of investment rather than of investment, e.g., a large number of small rice mills in a village may reflect a failure to support and sustain investment in a single more efficient, high capacity mill; and second, that the indicator so overwhelmingly reflected something other than Investment that it could not possibly be a "good" indicator of Investment, e.g., the number of monks and novices in a temple fluctuates with the seasons. Thus, the evidence regarding

---

* Ibid., pp. 23 and 26.

Table A

ITEMS IN THE 23 ITEM INVESTMENT INDEX

Economic

- Number of daily bus trips
- Number of village-owned bus s
- Per cent of farmers using fertilizer
- Per cent of farmers using insecticide
- Per cent of farmers using improved seed
- Shops per 100 households
- Number of rice mills in village
- Per cent of farmers in economic groups
- Ratio: seasonal employment/subsistence families

Social

- Number of monks & novices
- Temple condition
- School condition
- No. of students beyond P4 per 100 households
- House condition
- Per cent of households in social groups

Political

- Ratio: merchant leaders to total leaders
- Number of requests to amphoe
- Number of petitions to amphoe
- Frequency of visits to amphoe
- Frequency of development committee meetings
- Number of village groups
- Total number of village meetings last year

Source: Final Report, Table 3.12, p. 37.

face validity is somewhat equivocal. There really were not sufficient data available to make a final judgment about the face validity of the Investment Index, although it may be possible to demonstrate face validity with additional information.

Criterion Validity

Assuming the face validity of the kinds of behaviors that are included in the Investment Index, the issue of criterion validity becomes crucial because the information on which the operational 23 item Investment Index is based is not obtained from direct observation of these behaviors by objective observers, but, rather, on the reports of generally knowledgeable but potentially biased informants -- Community Development Workers and Village Headmen.

Two studies have collected data on the currently used 23 item Investment Index for a sample of villages, drawing information from the Village Headmen and also from intensive direct observation by an objective research team. If the Investment Index based on direct observation is accepted as a criterion, the data may be used to investigate the criterion validity of the Investment Index based on information obtained from Village Headmen.

The first study used a sample of only 15 villages. The correlation between the Investment Index based on information from the Village Headmen and that based on direct observation was .65. This same study also produced a correlation of .65 between the Investment Index based on information from CDW's and that based on direct observation. If these obtained sample values -- .65 in both

instances -- were accepted as estimates of the actual validity coefficient within the total population of villages, they, of course, would not meet the marginally acceptable criterion of .70.

The second study collected data for a sample of 129 villages, but the results have not been published: "A version [of the Investment Index] based on site visits is possible also, but we have given this development lower priority and are not ready to report on it."*

Two other studies using an earlier version of the Investment Index -- not the currently used 23 item version -- compare the scores of Investment Indexes using CDW's as the source of information with Investment Indexes based on 6-8 man hours of direct observation in a village by objective research teams. These are reviewed here because they provide indirect evidence regarding the criterion validity of the currently operational 23 item Investment Index.

One of the studies used a sample of only six villages and presents scores for the villages on the CDW-based Investment Index and on the direct observation-based Investment Index, but does not report a correlation coefficient for the relationship between the two sets of scores.** The correlation is estimated to be .63.

The second study used a sample of 27 villages and provided Investment Index scores based on three sources: CDW's, Village

---

* Report of Progress, July 15, 1972-January 14, 1973, p. 18.

** Final Report, Table 3.8, p. 31.

Headmen, and direct observation by a research team. Using the results of direct observation as a criterion, the study obtained a correlation of .64 between the criterion and the Investment Index based on CDW information, and a correlation of .62 between the criterion and the Investment Index based on Village Headmen information.

The results of all these studies indicate a considerable lack of correspondence between Investment Index scores based on information from preferred informants -- Community Development Workers and Village Headmen -- and those based on direct and intensive observation by trained objective observers.

The researchers have also used another criterion to test the validity of Investment Indexes, namely, the "judgments" of knowledgeable officials about "village development levels." Correlation coefficients of .65, .80, and .73 are reported between such judgments and early versions of Investment Indexes based on CDW information.* These data have two weaknesses as tests of the validity of the currently operative 23 item Investment Index. First, the Investment Index used in these studies is not the current 23 item version. The scores used in the first two studies reported immediately above come from an index employing 20 items of which only 8 overlap with the current version of the Investment Index.** The third study used a composite score for "Investment" based on four different sources of information using from 29 to 33 indicators each.*** These data,

---

* Ibid., p. 40.

** Cf. AIR, Village Investment in Development: An Index of Counter-insurgency Impact, October, 1970, p. 16 and Final Report, p. 37.

*** Village Investment in Development, pp. 40-42, and 48.

therefore, provide only indirect evidence regarding the validity of the currently used 23 item Investment Index.

Second, although a criterion for testing the validity of an index should be independent of the index, the criteria used in these studies were not independent of the Investment Index. In the first two studies, the CDW provided both the "judgment about village development level" (i.e., the CDW's were the "knowledgeable officials") and the information on which the Investment Index was based.* Thus the correlations should not be interpreted as indicating validity -- or the lack of it -- of the Investment Index, but, rather, as indicating the consistency of the CDW in evaluating a village. The CDW's global judgment is perhaps based on much the same information that is tapped by the indicators that comprise the Investment Index. Similarly, in the third study cited, the source that provided the global judgment also contributed to the Investment score; hence the correlation is not between independent sets of scores.

Construct Validity

The data cited above regarding criterion validity are certainly not a sufficient basis for a claim of validity of the current version of the Investment Index. Indeed, the Final Report, at the end of the discussion of the Investment Index, states, with reference to criterion validity, that "... we would not claim validity on this basis...."** Then the Report continues:

---

* Ibid., pp. 24-25 and 27.

** Final Report, p. 40.

> "Our claim for validity, and we believe it is a very strong claim, ultimately rests on the relationships that have been demonstrated to exist between Investment and other elements of the impact model. We will therefore turn to these elements, returning in a later section to the question of the validity of the concept and the measure of village investment."*

This is referring of course to the idea of "construct validity" and it prompts two comments.

First, the question of validity of the Investment Index is not discussed again in the Final Technical Report. Thus the question is left hanging with an assertion of a strong claim for validity but with no data to support it. Second, a claim for validity based on the demonstration of relationships between Investment and other elements of the impact model can only be made if the model has been independently verified. The same data cannot be used to verify the model and to test the validity of a measure; an adequate test of one of these issues assumes that the other issue has already been resolved in the affirmative. Since virtually no data regarding the model have been produced, the model remains an unverified assumption and may not be used to demonstrate the validity of measures of the concepts used in it.

Summary of Validity

Thus the status of the validity of the Investment Index at this writing is that the issue is still open. Some questions have been raised about its face validity by persons knowledgeable about rural Thailand; data on alternate-form reliability suggests that the

---

* Ibid.

I. [illegible]tment Index probably is reflecting primarily a given information source's global judgment of a village; and high agreement with an independent criterion has not been demonstrated.

Sub-Indexes of Investment

Two sets of sub-constructs of the Investment construct were identified during the R&D program. One set involved the differentiation of Investment into Economic Investment, Social Investment, and Political Investment. The other set involved a differentiation between Community Investment and Individual Investment. Attempts were made to develop sub-indexes of the Investment Index to correspond to these two sets of sub-constructs.

The idea of separate Economic, Political, and Social Investment constructs is not supported by the empirical data. First, the corresponding three sub-indexes are not empirically differentiated from each other, i.e., the items in a given sub-index often correlate more highly with another sub-index than with the sub-index it is supposed to reflect.* Second, the internal consistency reliabilities of the sub-indexes are not high enough to be acceptable. The median reliability coefficients, over 13 separate data collections, were: Economic Investment, .74; Social Investment, .62; and Political Investment, .49. Since these sub-indexes are not empirically differentiated and the sub-indexes are not reliable, one must conclude that they should not be used. Unfortunately, this

---

* Ibid., Table 3.12, p. 37.

conclusion is not clearly stated in the Final Technical Report. Indeed, one statement may be read to mean just the opposite.* This finding has not been communicated to CD or ARD, but should be because CD has used these sub-indexes specifically to evaluate performance and guide action.

The idea of separate Community Investment and Individual Investment constructs was also not supported by the empirical data. The items in one sub-index often correlated more highly with the other sub-index than with the sub-index in which it was placed;** and the internal consistency reliabilities of the sub-indexes did not meet the criteria of marginal acceptability.*** The conclusions to be drawn from these facts are not adequately stated in the Final Technical Report; however these indexes have not yet been used in operational contexts.

Scoring Norms for the Investment Index

In order to make it easier to communicate with people who are not very technically oriented, the raw scores from the Investment Index are converted into "stanines" and reported to operating officials in terms of these stanine scores.

Stanine scores are relatively common in educational testing and are achieved by a simple transformation of raw scores into a

---

* Ibid., p. 35: "The summary conclusion was that there was an existing capability (instruments and procedures) for measuring the extent to which a village is investing its resources in political, economic and social development." Emphasis as shown.

** Ibid., Table 3.20, p. 53.

*** Ibid., Table 3.21, p. 55.

specific distribution that is technically known as the "normal" distribution. The use of this transformation assumes that the underlying construct is also "normally" distributed, even though the empirical raw scores are not. Thus the transformation into stanines forces the scores into a symmetrical "normal" distribution.

The procedure used in the assessment techniques research program to transform most of the reported scores on the Investment Index into stanines was a simple conversion table that presents a stanine value that corresponds to each raw score value. One enters the conversion table with a raw score and the table indicates the corresponding stanine score. This conversion table was produced by a transformation into stanines of the raw Investment Index scores, based on Community Development Worker information, from a sample of 380 villages. Since a conversion table based on data from one sample is used to convert raw scores from other independent samples into stanines, it involves the use of "scoring norms."

The use of scoring norms based on a sample of scores is a common practice and in itself is not objectionable. However, the practice assumes that the norm sample is representative of the population that the norms are to be used with. And the use of a set of scoring norms for determining stanine scores also assumes that the central tendency and the dispersion of the raw scores for the norm sample are similar to the central tendency and dispersion of raw scores from the population with which the norms are to be used. If the assumptions are not met, the use of the norms will

cause distortions of the transformed scores for the later independent samples.

The procedures actually used in the impact assessment research, and subsequently recommended for use operationally, raise several potential problems. The first is that the legitimacy of the conversion of raw scores into stanines is based on the assumption that the distribution of the underlying construct is "normal" in the technical sense. A very plausible argument can be made that the distribution of Investment in an LDC is not "normal" but, rather, is skewed in the direction of high Investment; that is, many villages can be expected to have low levels of development and hence score low on investment, whereas a very few villages will show relatively high investment ratings. If the distribution of "Investment" is thus skewed in the direction of high Investment, then the result of transforming raw scores into stanines would be to lessen the differentiation of villages at the higher levels of Investment and to exaggerate the differentiation at the lower levels. In other words, the assumption of "normality" implicit in the use of a stanine scoring system distorts the raw scores one would expect from the Investment Index such that the few relatively high Investment villages would be pulled back toward the center and the low Investment villages would be pushed further away from the center.

A second, and perhaps more important issue, is whether or not the sample of villages on which the scoring norms are based is representative of the total population of villages to be considered.

There is no evidence available to show that the norm sample of 380 villages is representative. If the norm sample of villages is not representative, the use of the norms based on this sample may distort the distributions of scores from later samples of villages. The result would be a spuriously "high" or "low" average and a restricted range for the distribution. For example, most of the scores could wind up being high scores, e.g., in stanines 7, 8 and 9. Then, assuming a reasonable level of continuing development activity, all of the villages would shortly be jammed into the same high stanines and there would be no way to differentiate among them. Such a result may already have occurred in one of the research studies carried out by the contractor,* but this may also be a product of the inadequacy discussed in the next paragraph.

A third and potentially quite serious problem with the use of the scoring norms that have been provided to ARD in the Handbook is that they are based on information provided by Community Development Workers, whereas all of the data collected by ARD in 1973 and 1974 have used the Village Headman as the source of information. This fact is not reported in either the Final Technical Report or the Impact Assessment Handbook and appears not to be recognized by Thai personnel using the techniques. Such a procedure may produce distortions in the Investment Index scores in two ways.

---

* Final Report, Table 3.14, p. 39.

First, as was pointed out in the discussion of the reliability of the Investment Index, the correlations between scores based on CDW information and those based on Village Headman information were too low for the two to be considered as equivalent forms. Therefore, norms for one are not appropriate for the other. However, even if there were a relatively high correlation between the scores on the two forms of the Investment Index -- indicating that a high score on one form would have a high score on the other, and a low score on one would be accompanied by a low score on the other -- it would still be possible for the central tendency and dispersion of the two distributions to differ. If the mean of the CDW-based raw scores is lower than the mean of the Village Headman-based raw scores, then the stanine scores of the Village Headman-based Index will tend to be in the upper categories and there may be a restricted range of the stanine scores. Unfortunately, no information has been published regarding the comparability -- in terms of central tendency and dispersion -- of the Investment Index based on the two different sources; but the kind of potential distortion just described is reported for one sample in the Final Report.*

Summary of Technical Adequacy of the Investment Index

The current 23 item version of the Investment Index that is based on information from either Community Development Workers or Village Headmen has a marginal to acceptable level of internal consistency reliability (if scored by summing across the 23 items in

---

* See Table 3.14, p. 39.

the Index) and an unknown but probably acceptable level of test-retest reliability. However, the correlation between the two forms -- one based on information from CDW's and the other on information from Headmen -- is far below acceptable standards for alternate-form reliability. Thus the two forms should not be considered equivalent and scores from the two forms should not be considered comparable.

The validity of the Investment Index as a measure of the degree to which villagers are expending time, energy or other resources for economic, social and political improvements is still an open question. People who are knowledgeable about village life in Thailand have questioned the extent to which some of the items in the Investment Index reflect the construct Investment and the mode of selecting the items raises the possibility that those finally selected are not representative of the population of behaviors that define the construct.

Neither of the current forms of the Investment Index correlate highly with a criterion Investment Index based on intensive observation of villages by trained objective researchers. The combination of acceptable internal consistency and test-retest reliability but low agreement between alternate forms based on different sources raises the possibility that the Investment Index is reflecting primarily the global judgment of a village by the informant. Indeed, scores on early versions of the Investment Index do correlate relatively highly with overall judgments of the village by the same informant.

There are also two problems with respect to scoring methods. First, the current procedure of providing scores by use of a particular set of scoring norms that may not be completely appropriate may create distortions in the Investment Index scores. Second, the three proposed sub-indexes have not proved to be empirically differentiated from each other and they have not shown adequate internal consistency reliability. Thus they should not be used.

Although the validity of the Investment Index as a measure of the concept "Investment" is open to question, the fact of its marginal to acceptable level of internal consistency reliability and probable test-retest reliability suggests that it may be measuring something. This "something" may be of interest and the Investment Index may yet be found to be related to antecedants and consequents that are of interest to people concerned with rural development and useful to policy makers, administrators and planners; however, a substantial research program would be required to develop these relationships.

The Opportunity Index

A second intervening construct in the conceptual model is "Opportunity," defined verbally as "anything in the environment which permits an increase in the range or magnitude of investments in political, economic, or social improvement."* This construct is measured by an "Opportunity Index," the current version of which is composed of 12 items. The Opportunity Index, like the Investment

* Ibid., p. 40.

Index has two forms, one based on information from Community Development Workers and the other based on information from Village Headmen. Indeed, information for both the Opportunity Index and the Investment Index are collected at the same time using the same interview schedule or questionnaire. The Opportunity Index is scored in a fashion similar to that used with the Investment Index and uses scoring norms based on the same sample of villages and the same source of information.

Reliability of the Opportunity Index

Several of the twelve items that comprise the Opportunity Index have quite low correlations with the sum of the other 11 items.* This indicates that the items are reflecting to a rather small degree the commonality among the items -- presumably the construct Opportunity -- and reflecting to a large degree something idiosyncratic to the item. That is, the items are not very good indicators of whatever the Index is measuring. It also means that there will be a lot of "static" in the Index scores and that low internal consistency reliability will probably be obtained unless large numbers of indicators are used in the index.

Split-half reliability coefficients have been reported for the 12 item Opportunity Index for 13 different data collections. These have a median of .70 and a range from .53 to .80.** Thus, the internal consistency reliability is very marginal, if acceptable at

---

* Ibid., Table 3.15, p. 43.
** Ibid., Table 3.16, p. 44.

all. No data are presented regarding the test-retest reliability over short intervals of time such as several weeks. Opportunity Index scores are not highly stable over longer periods of time, e.g., one or two years.*

Few data are presented regarding the alternate-form reliability of the Opportunity Index. The data that are available indicate a lack of comparability between scores based on information from CDW and those based on information from Village Headmen. The study in which data were collected from the two sources concurrently for a sample of 15 villages found a negative correlation (-.23) between the two forms of the Opportunity Index.** Two other studies, in which the interpretation is complicated by the fact that the data from the two sources were collected one year apart rather than concurrently, found correlations between Opportunity Indexes based on CDW information and on Headman information of .50 and .43.*** Other data that might be related to the issue have not been published. Equivalency of the two forms and comparability of the scores from the two forms has certainly not been demonstrated, and indeed, the data raise considerable doubts about such equivalency.

Thus the currently available Opportunity Index does not appear to have sufficient reliability to be a widely useful measure. The researchers would appear to agree with this evaluation: "The reliability is less impressive than is the case for the Investment

---

* Ibid., Table 3.18, p. 48.
** Ibid., Table 4.2, p. 71.
*** Ibid., Table 3.18, p. 48.

measure.... All that is desired of an opportunity measure is that it serve as a crude control for observed differences in investment. The present measure seems adequate in this sense."* Unfortunately, the Opportunity Index has been recommended for use in ways other than as a "crude control" -- for instance, in the application called PAM (discussed in Section V). Such use in an operational setting is obviously much more than a "crude control," and it certainly requires a higher degree of reliability than the Opportunity Index has been demonstrated to have.

Validity of the Opportunity Index

Generally, if a measure is not reliable, it cannot be valid. Considering the extremely marginal internal consistency reliability of the Opportunity Index, one should not invest very much time inquiring into its validity. However, some data are available comparing the Opportunity Index scores derived from information provided by CDW's and by Village Headmen with criterion scores derived from direct observation of trained objective researchers. One study, using the current 12 item Opportunity Index with a sample of 15 villages, found criterion validity coefficients of .48 of the Village Headmen version and .31 for the CDW version. Another study, using a different version of the Opportunity Index in a sample of 96 villages, found a correlation of .58 between scores based on information from CDW's and scores based on direct observation by trained objective researchers.** Clearly, acceptable criterion validity has not

---

* Ibid., pp. 42 and 45.

** Report of Progress, July 15, 1972-January 14, 1973.

been demonstrated for the 12 item Opportunity Index based on information from CDW or from Village Headmen.

Scoring Norms

Stanine scores are derived for the Opportunity Index in the same manner that they are for the Investment Index and use similar norms. The criticisms of this procedure and these norms made above for the Investment Index also hold for the Opportunity Index.

Summary of the Technical Adequacy of the Opportunity Index

The internal consistency of the Opportunity Index is very marginal and there is no agreement between scores based on the two preferred sources of information, and very little correlation between these two sources and a criterion based on direct intensive observation of villages. Consequently the current version of the Opportunity Index should not be used generally, and certainly not by operating agencies to make decisions regarding individual villages.

Empirical Tests of the Theoretical Model

The theoretical model guiding the assessment techniques research postulated relationships between a class of antecedants, Inputs from Development Programs, two intervening constructs, Opportunity and Investment, and a class of consequents summarized by the term Polity Strength. Given development of measures of the intervening constructs -- an Opportunity Index and an Investment Index -- it would be possible to test this theoretical model. The Final Technical Report of the research program presents several correlations among measures of the antecedent, intervening, and consequent variables.

These empirical relationships among indexes of the several constructs in the theoretical model can be used either as tests of the construct validity of some of the measuring instruments, if it is felt that the model has been verified, or to test empirically the theoretical model, if the measures are assumed to be valid. Since the model has not been empirically verified, it is more appropriate to consider the available data as a test of the model.

The Relationship between Opportunity and Investment

Data are presented on the correlation between the Opportunity Index and the Investment Index, measured concurrently, for 13 separate data collections. The median correlation coefficient is .59 with a range from .52 to .73.* Correlations between independently measured constructs that are conceptually differentiated in a theory are seldom so high. Indeed, considering the internal consistency reliability of the two measures -- median values of .82 of the Investment Index and .70 for the Opportunity Index -- the observed correlations between the Opportunity Index and the Investment Index are very nearly at the maximum possible. This suggests that the two Indexes are measuring essentially the same thing -- not independent constructs.

It may be recalled that both the Investment and Opportunity Indexes had much higher internal consistency reliability than alternate-form reliability, suggesting that these Indexes were

---

* Final Report, Table 3.17, p. 46.

primarily reflecting the global judgments of the source rather than objective characteristics of the villages: If one now notes that in these data the Opportunity Index and the Investment Index are not independently derived but rather are based on information from the same source -- either a Community Development Worker or a Village Headman -- one has a clue to a probable interpretation of these high correlations. Both the Opportunity Index and the Investment Index are measuring much the same thing -- the information source's global judgment of the village. It is thus very doubtful if these data say anything about the theoretical model.

The theoretical model, of course, states that Opportunity leads to Investment. The proper test of this relationship is a correlation between Opportunity measured at one point in time and Investment measured at some later point in time. Several such correlations are presented.* One has to be very careful interpreting these figures, however, because the data sometimes come from the same versions of the Indexes and sometimes from different versions, and sometimes from the same source and sometimes from different sources.** Generally, indexes derived from the same source, even though in different years, have the higher correlations, suggesting that the source's relative global judgments are reasonably constant over time. However, the correlation between the Opportunity Index based on CDW information and the Investment Index based on Village Headmen

---

* Ibid., Table 3.18, p. 48.
** Ibid., cf. p. 47.

information a year later is .59. This value is very nearly the maximum possible given the internal consistency reliabilities of the two indexes -- .69 for the Opportunity Index and .90 for the Investment Index. This is too high because it would indicate that Opportunity is the sole determinant of Investment a year later. Since this is not consistent with other findings to be reported (that size of village is a principal determinant of Investment) nor with the intent of the theoretical model (which has built into it Disposing Conditions that influence Investment independently) perhaps one should consider that this high correlation is a chance occurrence resulting from sampling or that it is an artifact of procedure, as some other high correlations seem to have been.

The Relationship with Antecedants and Consequents

The theoretical model postulates a relationship between the antecedant Development Program Inputs, and Opportunity directly and Investment indirectly. During the research program, a few attempts were made to examine this relationship. The two definitions or indicators of Development Program Inputs used on these occasions were, respectively: (1) the number of ARD projects completed in a village, or (2) the total baht funding for all known RTG projects in a given village, regardless of sponsor.

A significant correlation is reported between number of ARD projects in a village and Opportunity Index scores and Investment Index scores based on Village Headmen information for a sample of 127 villages, although no correlation coefficient is given.* It

* Ibid., Table 3.19, p. 49 and pp. 47 & 50.

would be very surprising if a significant correlation were not obtained between the number of ARD projects in a village and Opportunity Index scores, because the consequences of ARD projects are rather directly reflected by several of the Opportunity Index items.*

Three studies, each using large samples, are reported which provide correlations between total RTG funding over the preceding three years for projects in a given village and the Opportunity and Investment Indexes based on Village Headmen information. Correlation coefficients between funding and the Opportunity Index were .54, .20, and .01 in the three studies; correlations between funding and the Investment Index were .35, .21, and .17.** Thus, there appears to be a real but rather small relationship between funding for village projects and both the Opportunity and Investment Indexes, but the relationship with "Opportunity" is not larger than the one with "Investment." Indeed the researchers concluded that "... the hypothesized role of O[pportunity] as a mediator [between inputs and Investment] is not established by the data."***

Correlations are also reported between a truncated 15 item "Individual Investment" Index, with 14 items overlapping the 23 item Investment Index,**** and "Development Inputs" which are not otherwise defined. These appear to be based on the same three studies using large samples that provided the data discussed in the immediately

---

* Ibid., cf. Table 3.15, p. 43.
** Ibid., p. 50.
*** Ibid.
**** Ibid., cf. Table 3.20, p. 53.

preceding paragraph. The correlations are .52, .53, and .45. These are difficult to interpret because "Development Inputs" are not defined, and also because it appears that the source of information on "Development Inputs" probably was the same source as that for the information on which the Investment Index was based.

It should be noted that all of the above data are in terms of correlations which reflect similarities in the pattern of relative relationships among a set of entities on two separate attributes, but are rather insensitive to a general movement of the distribution of the entities up or down an attribute dimension. On the other hand one would expect an impact assessment technique to be sensitive to movement <u>along</u> a dimension. However, no data are reported concerning <u>changes</u> in Investment Behavior -- for example, in terms of mean Investment Index scores for a group of villages -- over time, although it is reported that there were repeat measures collected for certain villages over one or two year periods.

Another finding raises questions regarding the model's postulation that Development Inputs lead to Investment Behavior. Although village "size" was considered to be primarily a variable which might moderate the impact of Opportunity -- and hence Development Inputs -- on Investment, in reality "Size" seems to exert some considerable, independent influence in determining Investment. A correlation is reported between the Investment Index and Size of Village for a sample of 1117 villages, although no coefficient is given.* However, the coefficient may be estimated

---

* <u>Ibid.</u>, p. 56.

from the data in Table 3.22 of the Final Report to be approximately .50. This observed correlation between Size of Village and Investment Index appears not to be an artifact of the indicators used in the Investment Index because these were either corrected for village size by being converted into a ratio based on size or they are logically independent of size.

This independent and real influence of Size on Investment was recognized by the researchers, who said that "The observed relationship between Size and the Investment Index is not a statistical artifact and it is not produced by Opportunity."* Indeed, one may estimate (based on the assumption that the correlation between Size and Opportunity, which is not provided in the Final Report, is between .30 and .40) that between two-thirds and three-fourths of the association between the Opportunity Index and the Investment Index is a result of Size.**

In considering the adequacy of the theoretical model one should note that Size accounts for more than five times as much of the variance in Investment Index scores as does total Royal Thai Government funding over the preceding three years in a given village.***

Finally, as noted previously, no empirical data have been presented to test the assumed relationship between Investment and "consequents" such as Polity Strength.

---

* Ibid., p. 58.
** Ibid., cf. p. 56.
*** Calculated from data presented on pp. 50 and 56 of the Final Report.

Summary of Test of the Theoretical Model

In summary, there has not been an adequate test of the relationship between Opportunity and Investment (that is, with the two variables independently measured).

The role of Opportunity as a mediator between Development Inputs and Investment has not been established. There is a small correlation between Development Inputs, as measured by number of ARD projects or total government funding for village projects, and Investment. The magnitude of the relationship is such that between 5% and 10% of the variance in Investment is attributable to Development Inputs. Size of village accounts for considerably more of the variance in Investment than does Development Inputs as measured in this research. No data are presented to relate Investment to "consequent" variables, e.g., Polity Strength, that might be of interest to development planners, policy makers, or administrators. The theoretical model remains a plausible and intuitively appealing, but empirically unverified assumption.

Conceptual and Methodological Conclusions

Several conclusions that have been presented in the Final Report as "Implications" of the research data are stated in language which is open to misinterpretation, especially by LDC's. These conclusions of the contractor presented below in the form of quotations leading off each paragraph, and are discussed in the light of the data which have been reviewed above.

"First, it has been demonstrated that objective and reliable measures of the principal constructs exist."* In our review, the Investment Index was found to have a "marginal to acceptable" degree of internal consistency reliability, if scored by summing across the 23 items in it, and the Opportunity Index was found to have a very marginal to unacceptable degree of internal consistency reliability. The small amount of data presented on the issue indicate that both the Investment and the Opportunity Indexes reflect great inconsistencies between the two alternative sources of information -- Community Development Workers and Village Headmen -- on which the Index scores are based; and these two principal versions of the Indexes provide scores that are not highly correlated with a criterion using the same indicators but information derived from direct, intensive observation by trained objective researchers. The data do not appear to substantiate the asserted conclusion.

"The measures are useful in assessing program impact."** There does seem to be a rather small correlation between Development Inputs, as measured by number of ARD projects completed or total RTG funding in a given village, and both the Opportunity and Investment Indexes. The magnitude of the correlations is such that roughly 5% to 10% of the variance in both Indexes can be accounted for by such Development Inputs. Size of village accounts for more of the variance in the Index scores than do the Development Inputs. These

---

* Final Report, p. 58.
** Ibid.

data would suggest that the Investment and Opportunity Indexes have not yet been demonstrated to be a highly useful measure for assessing program impact.

"The measures ... can be used by RTG agencies ... in routine monitoring..."* The Opportunity Index does not have sufficient internal consistency reliability to allow it to be used to identify individual villages for decision making purposes. Both the Opportunity Index and the Investment Index have such low correlations between alternative sources of information and between each of these sources and the best criterion -- direct observation by trained researchers -- that the validity of scores for monitoring what is happening in villages is questionable. In the one situation in which these instruments have been used for extensive monitoring -- in the Community Development Department -- the Political, Economic, and Social sub-scales of the Investment Index were employed; these sub-scales have been demonstrated to have low internal consistency and should not be used to identify specific individual villages for any purpose.

"The measures ... can be used by RTG agencies ... in program planning."** Again we must note that the Opportunity Index does not have sufficient internal consistency reliability to allow it to be used to identify individual villages for decision making purposes and that the degree to which both the Investment and

---

* Ibid., p. 59.
** Ibid.

Opportunity Indexes are reflecting anything in the village other than size is probably very small. The one instance in which these Indexes are being used as part of a program planning process -- the Project Allocation Method in security support planning activities -- seems to place greatest emphasis on the Opportunity Index, which is the weaker of the two instruments. Indeed, as noted above, the researchers themselves have stated that the technical adequacy of the Opportunity Index is such that it should be used only "as a crude control for observed differences in investment,"* and thus would appear to restrict its application. In the light of these considerations, it would seem unwise to use the Opportunity Index in PAM. The role of the Investment Index in PAM is clouded by administrative problems in applying PAM (discussed in Section V), but even on methodological grounds the Investment Index's value would appear to be uncertain at best, taking as a whole the various comments discussed above regarding its reliability, validity, the scoring norms issue, etc.

"Finally, the model from which the constructs derive must be accepted as a plausible representation of rural development in Thailand."** In contradiction, however, the report also states, "but the hypothesized role of O[pportunity] as a mediator [between development Inputs and Investment] is not established by the data."*** Further, as noted, the data indicate that Investment may be more

---

* Ibid., p. 45.
** Ibid., p. 59.
*** Ibid., p. 50.

reflective of village size than of Development Inputs. No data are presented regarding the relationships between the intervening variables of Investment and Opportunity and consequents such as Polity Strength. To repeat, the theoretical model may still be rationally plausible, but the empirical data presented thus far do not add an empirical basis to the original rational plausibility.

Thus the above "conclusions" or "implications," as asserted on pages 58 and 59 of the Final Report, do not appear to be well grounded in the data presented elsewhere in that report.

The thrust of our comments on the data available regarding the reliability and validity of the Opportunity and Investment Indexes certainly raises serious doubts about their utility for operational purposes such as development program planning, assessing program impact, etc. The internal consistency of these measures, for example, is such that considerable error would be involved in measuring a given individual village on either of these dimensions. This error could result in gross misclassification of individual villages, a significant drawback for the uses of these techniques proposed for operating development agencies. There are, however, some potential uses which the reported deficiencies would not rule out.

For example, if an agency were to decide that the indicators that comprise these scales on their face reflected some concept of importance, and if the agency were interested in the impressions of CD Workers or Village Headmen about villager behaviors related to the concept, then the Investment Index could be used for selected

purposes. These selected purposes would involve situations where error in measurement would be outweighed by other factors. Two kinds of situation come to mind. One is where rather large effects or differences are expected and the size of the effects or differences would overshadow the amount of error inherent in the measurement. For example, one might use the Investment Index to detect the impact of doubling the wealth of a village by means of cash grants to each household. A second situation is where one wishes to compare the central tendency of two or more relatively large groups of villages. In this instance the measurement errors due to lack of reliability would tend to be random and would tend to "cancel out" in large groups to provide a more reliable measure of the central tendency of a large group. Thus, for example, the Opportunity and Investment Indexes might be used to compare two relatively large groups of villages categorized by another independent criterion -- such as the absence or presence of a road, or the size of the village (e.g., villages categorized by number of households) -- or to detect change over time in the mean value of a relatively large group of villages. These uses, however, are much less ambitious than the applications of the assessment techniques attempted to date in Thailand.

V

RTG USE OF THE IMPACT ASSESSMENT TECHNIQUES

The assessment techniques, or variations of them, have been or are being used in two RTG agencies: the Office of Accelerated Rural Development (ARD) and the Community Development (CD) Department, both located in the Ministry of Interior. CD's use, a one-time research and evaluation project, was completed in the period January-June 1974. A similar project may be undertaken in the future. ARD maintains two uses, one occasional and one continuing.

ARD is basically a rural infrastructure construction agency, with emphasis on rural roads, but now desires to become a more broadly-based development agency. The CD Department trains and sends individual workers into villages to promote the idea of villager self-help. Following the October 1973 revolution the RTG decided to merge the two agencies into a single Rural Development Department but did not implement the decision at that time. During the RJBA Team's visit in August 1974 it was announced that the merger would now proceed and each agency was directed to submit reorganization proposals to the Minister of Interior by the end of August.

Although the two agencies use the assessment techniques, neither can be said to have assigned them a central role in planning activity. The techniques are present and visible, but peripheral to the basic planning, programming, and evaluation activities of the two agencies. They do not influence the allocation of a

significant portion of the resources in either agency. A cynic might say that the techniques have been accommodated or co-opted in characteristic Thai fashion -- bits and pieces added on to (but not substituted for any part of) the existing management mechanism, on the theory that they can do no harm and might do some good, and perhaps for display purposes as sophisticated management techniques which are nonetheless used to support whatever the agency was likely to do anyhow. An optimist would consider the fact that two Thai agencies are actually trying out the techniques or variations of them, to bode well for the future and trust that with continued use, and some additional technical assistance and/or training, they would prove their worth. The RJBA Team has concluded that ARD and CD are neither particularly cynical nor particularly enthusiastic about the techniques, but are interested in them, and for the present are viewing them in a rather tentative, experimental light. In any event, their use has been restricted to certain specialized instances.

As noted above, the CD leadership has rejected the Polity Strength outcome of the model and ARD ignores it. It is probably true that most of the professional staff below the ARD and CD leadership levels do not understand the concept, though they may be more confused by the jargon used to describe the measurement aspects associated with the assessment techniques and their use than they are by the concept itself. Both agencies tend to focus on the Opportunity and Investment Index portions of the model, with

CD trying out its own version of these indexes predominantly in an evaluation context, essentially for monitoring or scorekeeping purposes. ARD has been using the assessment techniques as an aid to planning the location of infra-structure projects in certain security sensitive areas; it also uses a very limited project evaluation procedure that assesses "impact" but has nothing to do with the basic model and the O/I Index approach. ARD does not use the techniques for program evaluations. A more detailed discussion of these uses and their relationship to the principal planning systems in the two agencies is presented in the Appendix.

Community Development Department

Starting with CD, the Director-General of the Department made a policy decision about two years ago to undertake "concentrated" or "intensified development" in the one or two villages considered most suitable for development in each tambol (township) where the CD program operates. Accordingly, it was decided that "community development work will start in selected villages which have higher potentiality in development and are capable of extending the result of growth to other neighbouring villages in future."* A list of 23 criteria was given to each CD worker to use in making selections. As a result 1088 villages out of the over 20,000 villages covered by the CD program were designated for the new program.

---

* Community Development Department, Research and Evaluation Division, Mission and Accomplishments of Community Development Program in Thailand, 1962-1972, pp. 5-6.

Early in 1974 the CD Department's Research and Evaluation Division decided to modify the AIR Opportunity/Investment (O/I) survey questionnaire for use in attempting to structure the relative standing of each of the "intensified development" villages, i.e., to look at their ranking vis-a-vis one another. So far as could be determined, no attempt was made to catalogue actual CD inputs, or other development inputs in those villages, or to relate them directly to the scores obtained. Not surprisingly, many of the villages showed up in Stanines 7, 8 and 9 -- indeed it is a wonder that virtually all of them did not appear as "9's," given that each was originally selected on the basis of its potential for development. All, of course, will be "9's" before long if nothing is done to re-establish the norms, as noted in Section IV. The principal motive for using this "evaluation" technique seemed to be to confirm the wisdom of the original village choices for inclusion in the "intensified development" program.

The CD Department also reported that it is sending the results of the O/I survey questionnaires back to provincial CD officers to serve as "guidance," in the form of a supplement to the village benchmark surveys, five year plans, and other documentation normally used for CD planning, if they choose to use it, in accordance with the Department's "village-up" planning philosophy.* The CD Workers apparently are free to accept or ignore these scores. CD headquarters

---

* Discussed in the Appendix, see pp. 1-4.

calculated Investment scores on the basis of the sub-constructs Economic, Social and Political, and in reporting these back to CD workers in the field has suggested, for example, that if a village is scoring well on Economic and Political, but poorer on Social, then attention might be devoted to upgrading the Social aspects of the village.* Unfortunately, as the discussion of methodology in Section IV above makes clear, the internal consistency reliability of these measures is so low they should not be used for making decisions about individual villages; at best they may be useful for making some overall judgments between fairly large groups of villages. We do not know how seriously CD officials in the changwat are taking either the scores or the guidance from headquarters. In any event, this particular input is essentially of a "one-shot" nature because the evaluation survey is not planned on an annual basis.

In sum, CD has used a particular variation of the assessment techniques to confirm its original village selections for the "intensified development" village program and to identify, in a rather loose way, villages among the favored group that have done less well than others. Secondarily, it has sent the scores back to the field with the intent that CD workers use them as an input to their planning activity in "intensified development" villages. The calculation of scores by summing across the Political, Economic and Social

---

* The interest in P, S and E may stem from the Director-General's conviction that both the major problems and opportunities found in rural Thailand have their roots in social and political factors.

sub-indexes of the Investment Index should cease because this procedure is unreliable. The evaluation exercise may be repeated in about three years, but no immediate additional use of the assessment techniques is contemplated by the Department.

In the context of AID's interest in a system that would feed back data on program or project impact to the planning and resource allocation process, followed by repetitions of the cycle, CD's use of the Indexes does not seem to qualify. At most, CD expects the Indexes to provide information to help guide planning and to assess the impact of development inputs in a very general way. CD-modified O and I Indexes have simply been used, on a test basis, to monitor a particular class of villages (and they may be used again for that purpose), but they are not considered to be a major planning or resource allocation device by the Department.

Office of Accelerated Rural Development

The ARD case is more complex. Four ARD offices ostensibly have a role in the use of assessment techniques. As could be expected with an experimental procedure, the attitudes among them differ. Only one -- probably the least essential of the four in terms of planning and programming responsibilities -- is enthusiastic about the assessment techniques; a second is only superficially involved. Another uses the techniques when directed to do so by the Secretary-General, but clearly lacks conviction as to their value, and a fourth rejects them intellectually and resists their use.

Project Allocation Method (PAM)

One special use seems to have achieved the status of standard operating procedure. It appears that whenever the Internal Security Operations Command, or ISOC, (formerly CSOC) requests support from ARD for operations in selected security-sensitive areas, ARD will apply O and I Indexes within a procedure called the Project Allocation Method (PAM), as part of its planning process. This has been done in two security support operations and appears likely to be repeated for a third.

Given a geographic area selected by ISOC on the criterion of security, plus policy decisions by the Secretary-General to undertake only certain types of infrastructure projects within those areas and to give priority to villages therein which are felt to lack opportunities, O and I scores are collected by ARD's Planning Division and used to select villages which are to receive first consideration for project inputs. In the first use of PAM, however, the final "sort" of villages from among those ranked by O and I scores seemed to be made on other, unspecified criteria.

Answers from some of the questions on the O/I questionnaire are also used by ARD headquarters staff as proxies for village "needs" when checking specific changwat project proposals for the villages in question. In other words, for example, if the answers to the two questions concerning water indicate an inadequate water supply, the village is said to have a 'water need." Only three types of infrastructure project are permitted -- water, road and a

limited range of "village improvement" activities. A more detailed description of PAM is presented in the Appendix, pp. 11-18.

The 0 and I scoring provides a convenient means of identifying villages that fit the categories selected by the PAM rationale, and numerical scores are also easy to communicate and display at policy level bodies such as the Committee on Accelerated Rural Development (CARD). ARD staff do not profess to know whether the 0 and I scores are accurate, but as one senior Thai official put it, "something is better than nothing," i.e., having some criteria is better than having no criteria. As a practical matter, the emphasis on "low Opportunity" villages in PAM is basically an endorsement of historic ARD practice, namely, to determine if a village has an infrastructure "need," and if so, to then try to meet it. The implicit point, given that ARD now wishes to minimize its involvement in security operations and to control tightly its resource commitment within the security areas it must enter, is that if this rationale and method of ranking satisfies CARD and ISOC requirements, then it serves a useful purpose.

In short, ARD presently limits use of the assessment techniques for planning and resource allocation purposes to the special case of conducting a few infrastructure-oriented projects in support of ISOC operations in security-sensitive areas.* These security support operations are increasingly viewed by ARD as exceptions to its normal

---

* The PAM design also was intended to cover a choice of some non-infrastructure projects as well, but thus far has hardly been used for that purpose. See Appendix, p. 15.

activities. When such occasions do arise, ARD undertakes a planning process tailored for the purpose. Within that framework, PAM plays a role in the selection of villages and for crude "indicator" or "verification" purposes with respect to the choice among the three types of infrastructure projects that may go into the village.

ARD does not use the assessment techniques for planning the two types of activities which consume the vast majority of its resources, namely, opening up new provinces to ARD programs (or maintaining a presence in old ones) and concentrating activity in "economic development Growth Areas" in very secure areas.* In theory, at least, there is no reason why PAM could not be used for these purposes. It was difficult to obtain good answers as to the reasons for this choice from personnel below the level of the Secretary-General himself who, when the question was put, said that he considered the assessment techniques to be highly experimental and as such wished to restrict their use until their value had been tested.**

AID personnel familiar with ARD's creation will recall the great emphasis placed on the development of province-level development and planning capabilities. The potential for achieving meaningful decentralization of government was in fact a major factor in the U.S. decision to proceed with the ARD program. The AIR

---

* Discussion of the ARD planning system(s) appears in the Appendix, pp. 4-18.

** Staff members usually gave two explanations: "because the Secretary-General said so" and/or that it required too much effort and resources to apply the techniques more broadly.

materials describing PAM lay considerable stress on the role of the changwat in using it; indeed, one gets the impression that PAM was being designed primarily for changwat (province) planners. The Handbook, for example, referring to PAM and use of the decision-matrix, states that it is "the task of the changwat-level officials to (1) identify actual villages of this type; (2) identify what type of ARD project will most effectively meet the needs of the people in these villages; and (3) formulate official plans for these projects..."* In practice, however, PAM as presently conducted seems to foster, if not require centralization of planning responsibility in Bangkok.

Individual Project Impact

The second ARD application of impact assessment techniques deals with project impact in villages. The Evaluation and Reports Division, which basically audits the physical completion of individual projects via on-site inspection visits, is using a spin-off developed specifically for it by AIR. Called the Routine Impact Assessment Form it encourages the auditor or "spotchecker," by means of a few questions, to determine whether the villagers know that a specific project (limited to wells and roads/streets within or between nearby villages; ARD's "standard" rural roads, for example, are excluded) exists and when it was finished; know (without prompting) which RTG agency built it; and can recall whether they

---

* Handbook, p. III-2.

contributed to it in some fashion (e.g., donations of land, labor or food to ARD equipment crews). The answers to these questions, it is felt, provide a measure of specific project impact that adds spice to otherwise mechanical reports that well X in village Y was spotchecked and was found to exist and to be in Z condition. This sort of project impact information would appear to be most valuable in a "hearts and minds" setting, i.e., the sort of environment within which ARD was originally created, where it is deemed important to know if the "government" or the "system" is really registering in the minds of the villager. It does not appear to be particularly helpful in assessing general developmental impacts or providing data for planning purposes. This sort of "impact assessment" information, though perhaps useful for ascertaining a project's visibility, is unrelated to the basic model and to the O/I measurement activity; hence it sheds no light on the impact assessment techniques of primary interest to AID.

In addition the RIAF contains a second section which calls on the spotcheckers to administer the standard O/I Index questionnaire in each village they visit;* however, it is not clear just why this is being done and no use is being made of the data collected. There seems to be a general feeling that if the questionnaire were to be readministered 2-3 years later (perhaps in the security support

---

* The Division apparently will administer the O/I section when it is evaluating projects, such as wells and village improvement projects, for which the Village Headman is the respondent; in a recent attempt to evaluate the COMPAC program, the Division chose not to administer the O/I section because COMPAC members, not the Headman, were the respondents.

areas where PAM was used), some comparative judgments could be reached about "progress," but this is all very vague and the Division is uncertain about which element of ARD has or should have responsibility for long-term impact assessment work. One explanation is that collection of these data on a routine basis may have been conceived originally for R&D purposes, i.e., to assure a continuing flow of data for research use. ARD, however, does not have an in-house R&D capability to further develop the impact techniques and it has been given no instructions or game plan as to what to do with such a data base for operational purposes. The data, in our view, are more suitable for research purposes than for operational use.

As we understand it, the RIAF was designed primarily to give the Evaluation and Reports Division a boost in morale and a capability to collect some elementary impact information. It has succeeded in both instances, although the Division Chief laments that nobody at higher management levels reads his reports.

Constraints and Problems in Using the Assessment Techniques

For the sake of brevity, the twin issues of constraints influencing use of the assessment techniques for planning and evaluation within Thai agencies and the identification of problems encountered by those agencies in using them, are discussed jointly.

An important and pervasive constraint is the aforementioned attitude of Thai officials toward the conceptual content of the model. Either because they explicitly reject the model, or feel

uncertain about it, or do not understand it, enough doubt exists to serve as a significant constraint on use of the impact assessment techniques. As we have seen, they have been tried for program monitoring purposes (the one-time CD study), and been given a role in selecting villages for certain project activities (occasional security support planning exercises), but they are not viewed as significant planning and evaluation "systems" at this time. To achieve such status, it would probably be necessary to undertake more work deliberately designed to show that there really is some connection (as the model hypothesizes) between Programmed Development Inputs on the one hand and Opportunity and Investment on the other. Neither agency has a capability to do such research (discussed below). It is worth noting that the experimental use of the techniques in CD and ARD operations that has occurred has been almost exclusively due to the interest and desire of the two agencies' respective leaders. No significant source of enthusiasm exists at the professional staff level at this time.

One plausible source of the lack of enthusiasm by many Thai staff is the gradual realization that the assessment techniques work has not gone very far in the direction of particularized program assessment. Virtually all the research, save one example, assumed or defined the "Development Inputs" portion of the model to cover program inputs generally, without regard to sponsoring agency.*

* The exception was work published by AIR under the title "Some Evaluations of ARD Program Impact in Four Amphoe," November 1972. This work, however, did not involve the Investment and Opportunity Indexes or the basic formulation of the model.

The average program administrator is really concerned with what his specific program is doing and ARD managers -- whether their interests are Amphur Farmer Groups, roads, COMPAC, village wells, agri-business, etc. -- are no exception. The applications of the impact assessment techniques developed for ARD are not designed to provide that information, a fact that ultimately will serve to dampen the generation of significant interest in them on the part of many program directors.

The technical aspects of the assessment techniques, and the manner in which these have been presented, have confused many of the Thai participants. The methods are simple for persons with the requisite training (which itself is not unduly demanding), but the jargon can be difficult to comprehend and has been over-used in making explanations. This has served as a constraint on the ready acceptance and expanded use of the assessment techniques and will definitely continue to do so in Thailand.

The serious methodological problems discussed in Section IV should operate as a constraint on the use of the impact assessment techniques, but may not because Thai personnel are unaware of some of them (thus raising the issue of informing Thai agencies about these problems). Neither CD nor ARD has an in-house capacity to undertake serious developmental work to improve the assessment techniques significantly or even to keep them current. Speaking of ARD's shortage of personnel with training and experience in social science research, AIR has said that the problem "could simply not be resolved in the time frame of the project; a five-to-seven-year

effort, which includes graduate level training of ARD staff, would be the minimum required."*

The shortage of research personnel is due in part because the Thai civil service does not offer careers in social science research and in part because operating agencies such as CD and ARD are quite properly more interested in results than in research and feel they cannot afford the luxury of using staff positions for large research-oriented groups. All the Divisions must hire entry-level recruits to the Civil Service, few of whom are likely to have any research training or experience and virtually none of whom will be interested in a career in research. (ARD's Rural Survey and Research Division is largely staffed with temporary hires). This fact of life limits not only the prospect of doing R&D work itself, but also the agencies' abilities to use operationally techniques which are the experimental product of R&D. Thus, lack of a civil service research tradition is a handicap, though not necessarily an insurmountable one.

Apparently none of the staff associated with the assessment techniques in ARD has had training either in statistics or data processing (computers). They freely acknowledge that they do not understand the processing and analysis phases of the work very well, if at all. They seem to be capable of collecting the data in field surveys and coding it for input to a computer, but since the contractor's departure, they have not been able to undertake the key

---

* Final Report, p. 82.

punching and computer steps which must precede analysis of the data. The contractor developed suitable computer output formats, but had all processing done in a U. S. government facility, which is not directly accessible to ARD; however, the computer programs have not been transmitted to an RTG computer center and adapted for use at such a facility. CD also has not yet obtained suitable computer support. Since the assessment techniques, and particularly calculation of the O and I Indexes, depend heavily on computer support, it is likely that the indexes will not be used at all until the data processing problems are solved. The need for reprogramming and lack of computer access will serve as an obstacle to continuing use of PAM unless overcome.

Even given resolution of the data processing matter, the question of data analysis remains. The Division Director (Evaluation and Reports) most interested in using the assessment techniques asked for U. S. advisory help to train staff to do the actual analyses and preparation of reports from the computer output. There appear to be very few staff members who are able to take computer printouts (all in the English language) and develop them into written analyses and reports that are of value to planners and managers in the organization.

PAM, the most direct application of assessment techniques, clearly has its complexities. AIR has described them as follows:

> Since PAM allocations are based on impact assessment surveys in target areas using the investment and opportunity indices and internal village security ratings, its maintenance probably requires more

> technical expertise than any use of the impact assessment techniques: concerned ARD staff must be able to perform I/O/S surveys properly, guide ARD Central and changwat-level officials through its formal decision-making and data-based allocation mechanisms, monitor changwat applications, and be prepared to modify the PAM mechanisms in response to continually changing conditions.*

This capability does not presently exist in ARD. As AIR has correctly reported, "the 'critical mass' [of trained personnel] necessary for self-sustaining refinements and improvements" in the impact assessment techniques is not available.** ARD staff, for example, indicated that they do not feel adequately trained to change from the Investment Index scoring procedure in the Handbook to scoring that Index on the basis of summing across its 23 items, or to confirm the internal reliability of the Handbook scoring system.***

The assessment techniques are complex and costly to administer. Data collection, processing, and analysis all consume a significant amount of staff resources and time, and require the most highly skilled people available. In the ARD case, sheer availability of staff resources has become a serious problem. The Planning, Evaluation and Reports, and Research and Rural Survey Divisions in ARD headquarters, for example, are approximately the same size now -- with 42 changwat to cover -- as they were in 1966 when ARD was concerned with perhaps fifteen changwat. Freeing up people for work considered by most of the supervisors to be marginal will not

---

* Handbook, p. VI-8.
** Final Report, p. 81.
*** Supra, pp. 26-27.

occur, unless directed by the Secretary-General. If done in that fashion, there is likely to be a residue of prejudice against the assessment techniques themselves sufficient to reduce or neutralize their value, unless they succeeded in scoring rapid and unusually high successes. Unfortunately, prejudicial feelings toward the techniques already exist to some extent in two key Divisions as the result of alleged disputes between the division chiefs and contractor research personnel and they persist even though the contract has terminated. The highly personalized nature of relationships between senior staff (and their divisions) with one another, and with advisors and others with whom they must deal, is an aspect of Thai administrative behavior which remains dominant and difficult to contend with directly.

ARD has other organization challenges to meet. It has already been pointed out that four divisions are or have been involved in collecting data and working with the contractor. With the contractor serving as the linchpin, the practice of assigning staff (by order of the Secretary-General) from several divisions to work on the impact assessment techniques research succeeded reasonably well. However, the last U. S. and Thai contractor personnel have departed, leaving a vacuum. As a result, use of the techniques and further development of them has either lapsed or proceeds more or less proforma. None of the divisions feels responsible for the impact assessment work as a whole and none looks any further than the piece formally assigned to it. Hopes for more active participation by the

Rural Survey and Research Division and the exercise of overall organizational responsibility for assessment techniques work by the Office of the Assistant Secretary-General have not materialized.*

The major alternatives are to create a new Division for impact assessment or to establish a high level coordinator in the Secretary-General's office, or perhaps to set up a small group to try to carry on research-type work designed to improve the value of the techniques to the offices using them. All of these alternatives generate a claim on limited resources, especially talent, and would require some professional training for the people involved. The final decision (if any decision is made) is likely to be based on judgments about the importance of the impact assessment work vis-a-vis alternative uses for the people involved. The "alternative uses" are about to grow in number because ARD's current policy thrust is to transform itself from a counter-insurgency agency devoted to basic infrastructure activity to a more broadly-based development agency with a wide-range of "production-oriented" economic development and service programs. This is a formidable task and it apparently will have to take shape simultaneously with the stresses and strains of merger with the CD Department.

To add more uncertainty to the mix, the new Policy and Planning staff (OPP) in the Office of the Undersecretary of Interior is at least showing signs of becoming more assertive. Thus the merger

---

* See ibid., pp. 80-81.

of ARD and CD headquarters staffs and the evolution of their relationship with OPP will be primary points of concern -- careers and programs will be at stake -- and worry about the most appropriate institutionalization of responsibility for the impact assessment techniques is not likely to rank high on the priority scale.

In brief, given the operating agencies' commitment to operational programs, the lack of a research tradition or a mandate for research programs, the experimental nature of the impact assessment techniques, the need for technical sophistication both in using them operationally and in further developing them, and the burden on staff resources which results when attempts are made to use the techniques in these circumstances, the cautious application which has occurred to date is both understandable and appropriate.

VI

FUTURE USE OF IMPACT ASSESSMENT TECHNIQUES IN THAILAND

The future use of impact assessment techniques in Thailand can be considered in two contexts: (1) immediate operational applications in RTG agencies, and (2) additional research and development work.

Restrictions On Use of Opportunity and Investment Indexes

Given the technical deficiencies discussed in Section IV above, as supplemented by the information about usage in development agencies presented in Section V, there are definite hazards in routine, mechanical use of the O and I Indexes. At their present stage of development, a reasonable degree of technical sophistication is needed to make judgments about appropriate uses and to distinguish between what is acceptable from an R&D point of view and what is acceptable from an operating point of view. The requisite technical sophistication for such flexible use of these instruments is not present in the two agencies. Three examples of current inappropriate usage warrant attention.

First, ARD and CD are computing scores on both the Investment and Opportunity Indexes using scoring norms based on information furnished by Community Development Workers for a sample of 380 villages. These norms may or may not be appropriate for newly obtained Indexes based on information from CD Workers (no data on this matter are available). But such norms clearly are inappropriate for Indexes based on Village Headmen interviews. No data have ever been presented that indicate that the distribution of raw scores based on

the two sources for the Indexes are comparable in terms of central tendency and dispersion, and this would be a very minimum requirement for the use of scoring norms based on one source to provide scores from data based on the other source. Stanine scores for the Investment Index for 1117 villages based on Village Headmen information deviate significantly from normal in the direction of skewness toward low scores.* That is, most of the villages show relatively high levels of development and relatively few are found with low investment scores. In brief, the 1117 villages show higher levels of development than do the original norm group of 380 villages; hence the original norms are not appropriate for scoring the villages to which they are currently being applied. Thus a new set of scoring norms should be calculated. If this is not done, one cannot feel confident that the stanine scores derived actually reflect the distribution of Investment or Opportunity indicated by the raw scores from the Index. It would not require a particularly high degree of technical sophistication to develop a new and more appropriate set of scoring norms. Since ARD personnel assert, however, that they lack the knowledge, it would be necessary to arrange technical assistance, either in the form of training for ARD staff or use of a Thai university consultant or foreign advisor.

Second, the Final Technical Report presents data indicating that the Political, Economic, and Social sub-indexes of the Investment

---

* Final Report, Table 3.22, p. 56.

Index do not have a sufficient degree of internal consistency to be used in singling out individual villages in terms of these characteristics. Unfortunately, the scoring norms for these sub-indexes are presented in the Impact Assessment Handbook (Appendix B), and the Community Development Department has used the sub-index scores to identify specific villages on the basis of alleged economic, social or political inadequacies. Recognition of the inappropriateness of this use, in the face of the presence of the scoring norms for the sub-indexes in the Impact Assessment Handbook, requires not only a relatively high degree of technical sophistication but also a thorough familiarity with data that are published in the Final Technical Report, but not in the Impact Assessment Handbook.

Third, the fact that the Opportunity Index does not have a sufficient degree of internal consistency to be appropriately used to select individual villages to receive or not receive funds or projects raises serious questions about PAM. As it stands, the Impact Assessment Handbook presents both a set of scoring norms for the Opportunity Index and a method -- the Project Allocation Method -- for using these scores to assist in making determinations as to which villages are eligible and which are not eligible for receiving various types of development projects. Although language in the Final Technical Report (pp. 66-68) warns that this system should not be used rigidly and mechanically, again it would require a relatively high degree of technical sophistication and a thorough

familiarity with the data which are in the Final Technical Report but not in the Impact Assessment Handbook -- not to mention a rather venturesome operating official -- to question the uses that are implied in the Impact Assessment Handbook. The Opportunity Index is sufficiently unreliable that it should not be used to differentiate among individual villages, i.e., continued use of the Opportunity Index will mean that many individual villages will be incorrectly classified in priority rankings. If it really is important to the agency to target specific villages for specific programming purposes, this process should not be used. There are also sufficient questions still open about the reliability and validity of the Investment Index to warrant considerable caution in using it in PAM or any other operational context; on balance, it appears unwise to use it without specialist personnel, and this would amount to reinstitution of a research effort.

These three examples suggest three courses of action. First, CD (and ARD) should be advised against using the P, S and E subindexes of the Investment Index in the future. Since the assessment techniques are not in regular use in CD, this would not be disruptive to it.

Second, continuing use of the Indexes in PAM is unwise, especially given the central role of the Opportunity score in the present application of this methodology. Third, renorming of the O and I Indexes should be undertaken. Any decision to renorm, however,

should be taken in the light of the actual uses foreseen for O/I Index data. At this time, the only use of these Indexes occurs in PAM applications. Given the above comments regarding the role of the Opportunity and Investment Indexes in PAM, which renorming would not resolve, one can question the actual value to ARD of doing the renorming.

In brief, it is our judgment that the limitations of the current versions of the Opportunity and Investment Indexes are sufficiently serious that these two instruments should not be used routinely for operational purposes.

Maintenance of Existing Capability

Notwithstanding the above, the Team recognizes that decisions may be made to try to proceed with continuing collection and use of O/I data. In that event, several steps are warranted. First, the aforementioned renorming should be undertaken. Second, ARD would need to arrange for keypunching services and the writing of a computer program(s) compatible with whatever computer facility it will use. The keypunching could probably best be resolved by obtaining the equipment and training ARD staff to use it or working out an agreement with another government agency. Reprogramming, like renorming, would require outside technical assistance. Gaining computer access is something that senior officials in ARD would have to negotiate.

Successful maintenance and regular use of the O/I Indexes really warrants the provision of some additional training in

measurement and statistics for several ARD staff. We definitely believe that there are people on the staff fully capable of absorbing the training and would suggest the following type: a basic course in statistics (through bivariate descriptive statistics); a basic course in measurement; and guided independent study relating this basic material to the particular problems and issues arising in the agency's work. The measurement techniques are basically simple. On the other hand, they are not so simple that they can be turned over to untrained people for routine mechanical use. Without demanding great sophistication in research techniques, there nonetheless must still be some personnel with a deeper understanding of the statistical concepts on which the mechanism is dependent, if it is to function productively.

With respect to the issues of training, renorming and reprogramming, and making judgments as to the appropriate use of research instruments of limited technical adequacy for operational purposes, either foreign or Thai professional personnel could be used. We find it hard to believe that adequate technical expertise could not be found in Thai universities. The Secretary-General of ARD volunteered the opinion that he would like to have a cooperative relationship with a Thai university or individual professionals. Regrettably, the RJBA Team was unable to pursue this matter very far, in part due to the short duration of its stay in Bangkok and in part because no academicians could be found who already had some familiarity with the impact assessment techniques activity.

Interaction, then, had to be preceded by a rather long explanation of the entire subject. AID/W and USOM might wish to provide help in exploring the prospect of a liaison between technically knowledgeable Thais -- some of which may be in places such as the Educational Psychology Department at Chulalongkorn -- and ARD/CD; conceivably, it could be efficient to offer a U.S. advisor for several months to undertake that task and simultaneously provide some immediage advice to ARD on its renorming, reprogramming, and related issues. At the moment, the Thai agencies appear effectively isolated from the technical assistance they need for any but the most mechanical applications of the instruments left with them. We repeat, however, that the suggestions in this paragraph assume a decision to proceed with O/I usage for operational purposes, despite our aforementioned recommendations to the contrary.

An obvious question of interest is: what might be done, for operational purposes, with the large village data base which has been collected as part of the six-year research program and continuing RTG application of some of the techniques? We are inclined to think, as a practical matter, that very little can be done with it. While data on perhaps 2000 villages exist, much of it apparently on tape, there are serious problems that probably limit its use for specific given purposes: some of the data are based on different, not necessarily equivalent, versions of the Indexes; some of the data are based on different -- and definitely not equivalent or comparable -- informants; and some are repeat measures on the same

villages but not necessarily using the same index and the same informant.* Thus for any specific use many fewer than 2000 villages may comprise the data base. This may be one reason why the AIR team left no instructions for using these data with RTG agencies. In any event, it does not seem possible for an RTG agency to tap these data currently for assistance in planning and programming its activities. There may, however, be some research uses.

Possible Research and Development

Additional research can be viewed on both short and long term bases. The U.S. and Thai Governments made substantial investments of money, time and effort in the research program which culminated in the applications of impact assessment techniques discussed above. In our opinion (leaving asid the question of the desirability of continuing to test the original basic model) the instruments of the research are inadequate for routine use for operational purposes and the RTG operating agencies lack the base of technical knowledge needed to adapt uses to fit the technical limitations.

On the other hand, the central product of the research is a measuring instrument -- the Investment Index -- which has caught the imagination of many people engaged in development planning. Within a research framework, this Investment Index has been demonstrated to have marginal to acceptable level of internal consistency, and it has been shown to correlate at a rather low level with very gross indicators of development input; however, its connection to

---

* Cf. Final Report, p. 47.

more precise indicators of development input, its sensitivity to changes in development activity over time, and its relationship with desirable outcomes of the development process have not been adequately tested. As noted, data on Investment Index indicators for a rather large sample of villages in North and Northeastern Thailand have now been collected and are stored either as computer tapes or as computer printouts.

Building on this base, can something more of a research nature be done, at reasonable cost, to enhance the value of the impact assessment techniques for practical purposes?

Probably the single most useful piece of information for AID or the RTG to have would be some more direct confirmation (or rejection) of the proposed connection between Development Inputs and Investment. A research effort involving the following three steps would be required to develop more information on this point: (1) rescore the pool of Investment Index indicator data, using new appropriate scoring norms; (2) use this pool of data to explore relationships between new Investment Index scores and whatever readily available indicators of Development Input may exist; and (3) test the Investment Index's sensitivity to changes in development activity over time. As simple as this sounds, its chance of success would be very dependent on the ability to obtain information "readily" about past development program inputs -- their existence and when they appeared -- at the village level. As anyone who has quizzed Village Headmen and villagers, or has tried to use provincial

and district records, knows, this information can be very difficult to get, especially with regard to timing. Perhaps USOM would be in the best position to estimate both the availability and the cost of obtaining such information. It should be understood that the purpose of such an exercise would be to assess the potential value of the Investment Index, and lay the groundwork for decisions concerning additional R&D work with it. This would not necessarily affect operational use of the Index in RTG agencies. The resource commitment would involve: a Thai (or U.S.) professional to rescore the norms; some continued collection of Investment scores by ARD; collection of data concerning development inputs, perhaps in possession of USOM, NESDB and/or MOI; and, finally, use of the above to examine the Development Input-Investment connection, statistically and over time, under the supervision of a U.S. consultant. Ideally the latter would serve primarily as initiator and catalyst in a three-way cooperative effort involving Thai operating agencies and Thai technical people outside government.

If one is concerned about deeper questions such as the unreliability of the Opportunity Index or the basic construction of the Investment Index, the answer would appear to be that "much more research and development" would have to be undertaken. With respect to the latter, for example, serious questioning of the original list of 250 behaviors, and of the rationale for the eventual sorting down to the 23 indicators in the Investment Index, could lead to a process of starting the research program all over again. Given the central

role of the Investment Index in the evaluation of the impact assessment techniques, such a fundamental look probably is warranted before such an Index is used extensively to guide operational decisions. Such work, of course, would be far removed from the day-to-day concerns of an operating agency.

Ultimately, as was noted in Section III, the question of undertaking longer term R&D work should commence with a review by AID policy-makers of the model or models of the development process that they would most like to see tested, now. Debate about measurement approaches and devices would follow. A decision whether to proceed on additional research could only be made in the context of AID's overall R&D program structure and priorities.

Conclusion

On balance, it is our judgment that the impact assessment techniques now in use in RTG agencies should not be applied in the future for operational decision-making purposes. Some observations regarding additional research and development work to attempt to improve their suitability for operational use have been offered, but only the U.S. and Thai agencies concerned are in a position to determine the relative value of committing R&D resources to doing so.

VII

TRANSFERABILITY OF THE ASSESSMENT TECHNIQUES TO OTHER LDC'S

The PIO/T called for a statement "covering cost and complexities in an operational environment (as opposed to further research) of installing the system, as is in other LDC rural development agencies." The RJBA Team does not recommend installation of the "system," "as is," for operational purposes in other LDC's because at present there is no system, usage of the techniques to date in Thailand is both limited and experimental, and a number of conceptual and methodological problems remain. Use to date in Thai agencies does not support a case for export on operational grounds.

The research work underlying the design of the impact assessment techniques used in Thailand has involved three aspects: (1) a model of development; (2) a method of developing instruments (indexes) to measure concepts in that model; and (3) specific instruments (indexes) developed and used in the research.

The model is transferable. It has not been demonstrated yet to be empirically valid (or not valid) or practically useful, but it is a rational general model and if AID chooses to endorse it, it definitely can be transferred to other LDC's.

The idea of measuring intermediate variables that intervene between Development Inputs and Ultimate Outcomes in order to provide clues as to one's progress, and hopefully facilitate improved planning, is transferable.

The methods of developing measuring instruments (indexes) are transferable. The basic measurement model -- the psychometric or summated ratings model -- is a general measurement model that has been applied in a wide variety of educational, psychological, sociological, and anthropological research settings. If one wants to try to measure the behavioral aspects of the development process, then this is probably the measurement model to use. Use of this measurement model in a new LDC, however, implies another relatively long period of research and development to produce specific instruments appropriate for the new area.

Specific instruments -- such as the Opportunity Index and the Investment Index -- are not transferable. Indeed, there is no "off the shelf instrumentation" that can be used in a variety of places. The content of the indicators is extremely culture bound and a new set of indicators would have to be developed wherever the idea was used. They often change within countries and ARD, for example, was cautioned about this matter in Thailand:

> The current investment and opportunity indices and norms and internal security rating forms have all been designed to reflect the behavior of villagers in northern and northeastern Thailand. They are likely to be inadequate to serve as a guide for project allocations in southern Thailand, where opportunities and investment behaviors may be quite different due to the entirely different geographical and cultural situation.*

---

* Handbook, p. VI-9.

In brief, transfer of the assessment techniques as we have observed them in Thailand, for operational purposes, is inappropriate. Transfer of the model and the methods of developing measuring instruments for research and development purposes, is feasible.

PLANNING AND THE USE OF ASSESSMENT TECHNIQUES BY RTG AGENCIES

This appendix is intended to supplement Section V above and provide some additional details regarding the planning systems in effect in CD and ARD and the actual uses of the impact assessment techniques to date by the two agencies for operational purposes.

Community Development Department

Planning and project allocation decisions in the CD Department are rather decentralized, being essentially the responsibility of the village level Community Development Worker. In conjunction with the village Development Committee, the CD worker prepares a five year plan for each village in the tambol to which he is assigned, based on his and the villagers' perceptions of needs and likely resource availability. To assist in identifying needs and resources and to help the CD worker become familiar with the problems of villages under his purview, the Department requires that the worker complete a comprehensive and detailed "benchmark survey" of each village in his area. Once completed, a five year development program is drawn up using information regarding village need and potential gleaned from the survey and from interviews with villagers. Specific development proposals and budgets are submitted each year and approved by the Department if funds are available and if the project was in the original five year plan. Projects deviating from the 5 year plan require special justification before approval.

In line with its "village-up" planning philosophy, the CD headquarters staff is very small and its budget is a fraction of ARD's. Almost by definition the scale of its self-help projects is small.

Approximately two years ago, the Director-General of the Department directed that CD efforts be concentrated in one or two villages in each tambol. A list of 23 criteria was given to each CD worker and all villages under CD coverage were rated on this basis, resulting in the designation of 1088 villages (out of over 20,000 villages in CD program areas) as participants in the "concentrated development" or "intensified development" program.

Early in 1974, the Division of Planning and Research within the Department developed a modified version of the AIR O/I questionnaire and undertook to use it to determine the relative standing of each of the 1088 "intensified development" villages, i.e., the intensified development villages were being compared with each other, but not with villages which had not participated in that program. The CD workers answered this questionnaire on the basis of their knowledge of the villages where they were assigned. Answers given by CD workers -- 832 responses were received -- were verified in five provinces by sending staff from the Evaluation and Research Division to resurvey selected villages. The data were coded by staff in the CD Department and an unsuccessful attempt was made to run the AIR programs on NSO computers; thereafter AIR provided reprogramming assistance to the Department.

In addition to ranking the "intensified development" villages, the resulting O/I scores are also being used by CD in a particular way, intended to suggest the type of development emphasis each of the intensified villages should receive in the future. As discussed in Section IV, CD has divided the composite Investment score into component sub-scores for Social, Political and Economic Investment. Both the total O/I scores and the three Investment sub-scores are being sent back to the changwats, with the suggestion that CD workers use the scores as an input in considering the type of future assistance that the village might receive. For example, a CD worker might concentrate on increasing a village's political investment if the village had stanine scores of 9 for economic and social investment but a stanine score of only 6 on political. The origin of this use of investment scores as potential inputs to CD program decisions is uncertain. Unfortunately, it is an inappropriate application because of flaws in the methodology (see Section IV).

The Research and Evaluation Division also did an evaluation of the data collected from the 832 villages for the Director-General which concludes, among other things, that taken collectively the intensified development villages scored lowest on the Social dimension and hence "are still under-developed in the social aspects, namely, Education, Religion, Sanitation, Public Health, Community Order and discipline, as well as social grouping."* The Division

---

* Translation of CD Research and Evaluation Division paper entitled "Report on Country-Wide Intensified Community Development Classification," undated.

went on to conclude that "Emphasis of development should be placed on the [social] aspect more than any other so as to achieve a proportionate balance in all three aspects of economics, society and politics."

For the present, no immediate use of assessment techniques is foreseen in CD. The Research and Evaluation Division is talking about a possible resurvey of "intensified development" villages in two or three years.

Office of Accelerated Rural Development

ARD is using an application of the impact assessment techniques as part of a process for selecting villages to receive ARD project assistance in security sensitive areas and, to a lesser extent, to serve as a checklist against which to review choices as to the type of project that should be placed in those villages. A second procedure designed to provide some rudimentary information about the impact of wells or roads/streets within villages, on an individual village basis, is also in use, with "impact" meaning essentially villager awareness or recognition of the projects' existence and its sponsor. While dealing with project impact, it does not use the O/I Indexes and is not really germane to the main purposes of this report. The first activity is conducted by ARD's Planning Division and the second by its Evaluation and Reports Division. Both take place outside the mainstream of ARD planning and program activity.

The Planning Division is formally charged with overall responsibility for ARD policy planning and the coordination and integration of project activities in the ARD changwat (provinces). As a practical matter, the Engineering Division, whose programs account for about 90% of the ARD budget, exercises a great deal of autonomy, as do some of ARD's technical divisions at the headquarter level -- e.g., Youth and Agri-business.

Three distinct planning and project allocation "systems" were observed within the Planning Division: the ARD Blue Book, economic development Growth Area planning, and "security support" planning for specified ISOC counterinsurgency operations. The Project Allocation Method (PAM), which stems directly from an "assessments technique" background, is part of the latter.

The ARD Blue Book

Until quite recently, separate plans and budgets were prepared annually by each of ARD's operating divisions in Bangkok and by each of the ARD changwat. Because the plans were drawn up separately and independently, often at different times within the budget cycle, and usually without reference to other ARD activities in the same changwat, ARD administrators and RTG budget officials found it difficult if not impossible to ascertain either the extent of ARD-proposed activities on a comprehensive basis or their aggregate cost. This customarily resulted in extensive revision of ARD plans and delays in budget approvals.

In February 1973, ARD adopted a "Blue Book" planning system patterned after the Malaysian Red Book. The Blue Book system was intended to rectify the piecemeal planning of the past. Four "books" designed to increase the coordination of ARD planning and project execution make up the system. Book I establishes an overall planning cycle and schedule for the preparation, submission and review of changwat plans for all ARD program activities. It also includes relevant policy guidance and provides detailed instructions for the submission of annual plans and budgets for each type of ARD activity. Book II is the resultant compilation of all ARD project activities proposed by the changwat for the coming fiscal year, as determined in accordance with the Book I directions. Book III establishes an ARD Management Information System (MIS) and Book IV is comprised of the MIS submissions which each changwat is required to submit quarterly.*

The Blue Book system has been used by ARD for one budget cycle. Minor revisions are currently underway to show more clearly how RTG inputs are related to program outputs at the amphur level. There was also some reference during the Team's visit to complete revision of the system during the course of the next year, but no definitive information was provided.

The Blue Book system is seen by ARD officials as a significant improvement in ARD planning. For the first time ARD activities are

---

* For a more complete description see Michael Dwyre, The ARD Blue Book (Bangkok: USOM, May 1974).

submitted and processed according to a uniform cycle, applicable ARD policy and procedures are written down in a single document, a standardized format for project submissions has been established, changwat plans and budgets for all ARD project activities are filed in a single document, and a feedback MIS system exists to monitor project implementation (largely in terms of categories such as baht obligations, baht expenditures, start-up and completion dates, etc.) against approved plans and budgets. The Blue Book system still has a way to go before it becomes more than a catalogue of ARD activities and takes on the attributes of a more dynamic planning system which analyzes needs, identifies opportunities, develops a strategy and plan for solution of problems identified, and allocates resources in conformity with an overall development strategy.

ARD is presently using the Blue Book system for the standard, on-going ARD activities, i.e., for planning work in new changwat or sustaining conventional ARD activities previously initiated. Planning of ARD activities in security sensitive areas and in the newly-designated, high priority Growth Areas in the changwat takes place outside the Blue Book system. Opportunity/Investment impact assessment techniques are not involved.

ARD Economic Growth Area Planning

A second planning technique is just now in the process of evolving within ARD for use in conjunction with ARD's latest priority development concept, namely, concentrated economic development "Growth Areas." The idea is to focus and integrate a multiplicity of program

activities in a small area judged to have economic potential rather than scattering them over a changwat. ("Small" area originally meant an area with a radius of 5-10 kilometers, with perhaps 10-15 villages covering one or two tambol; due to pressure from the engineers, who prefer to operate over larger areas to achieve economies in equipment utilization, the size shows signs of increasing). The Growth Area selection process has been taking place during FY 2517 (October 1, 1973-September 30, 1974) and as of May 1974 thirty-one sites in thirty-one changwat had been approved.*

The criteria for selection as a Growth Area are high population density, "good" economic growth potential, "social viability," a high level of security, a history of previous ARD activity, and evidence of interest on the part of local officials and villagers. The objectives to be achieved in these areas are stated predominantly in terms of increasing farm output and efficiency, expanding occupational opportunities, and raising farmer incomes. Indeed, the Chief of the Planning Division described the ultimate objective as "increasing the farmers' income at the margin." Four types of ARD projects will be emphasized: community physical improvements; occupational and income development, specifically agri-business, COMPAC and land improvement programs; health services in the form of assignment of permanent tambol paramedics and regular visits by mobile medical teams; and youth activities, with an apparent emphasis

---

* Some ARD documentation lists thirty-one sites, plus several "probables;" Planning Division personnel usually refer to twenty-nine Growth Areas in discussion.

on teaching new skills to young adults and young marrieds and on forming youth groups to try communal farming and communal handicraft production.

Clearly, "the focus is changing from infrastructure to area development,"* and in doing so ARD has decided that a new approach to data acquisition and planning is also required. Planning has been divided into short-term and long-term categories, both under the tight direction of ARD headquarters. In the short-run (1-2 years), physical improvements will be emphasized, i.e., more infrastructure; in addition, basic social and economic data will be collected about the Growth Areas for two purposes: (1) to determine resource availability and development potential, and (2) to serve as baseline information against which to measure progress at a later date.

Accordingly, Planning Division staff are being sent from Bangkok to the field to collect data on agricultural production, trade, banking, industry, communications, transport, public services, the farmers' marketing habits, etc.** The procedures for these "Area Studies" were apparently worked out with officials in the Regional Planning Division of the NESDB. The data will be analyzed to determine each Growth Area's development potential; then projects will be proposed and an "operating plan" prepared at ARD headquarters' initiative.

---

* ARD, "ARD Growth Areas," mimeo, June 1974, p. 1.

** These studies are quite separate from the "baseline studies" that are conducted by ARD's Rural Survey and Research Division in new ARD changwats and periodically up-dated at 3-5 year intervals.

The Area Studies seem to encompass geographic areas that are larger than the initially conceived boundaries of the Growth Areas and hence could support expansion of the latter.*

The long-range plan is expected to emerge gradually and cover 5-7 years. The main elements of it are intended to flow in large measure from the communities themselves on the theory that each community "must identify for itself the nature of growth and development it wants."** ARD envisages a strong role for tambol councils, e.g., in selecting sites for water resource projects; arranging for community contributions to and community maintenance of facilities; supervising and supporting youth activities; and generally articulating their own view of the nature and type of development desired.

It is ARD's intention to set growth objectives for the Growth Areas, to establish indicators to measure progress, and to schedule regular evaluations (the first is tentatively set for January-March 1975) on a triennial basis. The indicators will be defined in terms of income increases, new farming and occupational practices, access to markets, health conditions, etc. There apparently is no intent to use impact assessment techniques in this evaluation.

---

* ARD hopes eventually to relate the Growth Areas effort to NESDB thinking about regional growth and to a possibly forthcoming Ministry of Interior (Office of Policy & Planning) plan covering the relationship of secondary cities to rural development. All three agencies -- ARD, NESDB and MOI (OPP) -- appear to be strongly influenced at the present time by various interpretations of the "growth pole" theorizing found in the development literature.

** ARD, "ARD Growth Areas," p. 11.

Conceivably, once Growth Area planning becomes routine, it may become part of the Blue Book system. But for the foreseeable future it seems likely to evolve as a separate, and perhaps paramount, form of planning in ARD. The Opportunity/Investment Index approach is not involved.

Security Support Planning and the Project Allocation Method

From time to time ARD is requested by the Internal Security Operations Command (formerly CSOC) to assist the latter's security operations in rural areas. These operations are normally carried out under an overall Civilian-Police-Military (CPM) plan. Using a low (Red)-intermediate (Yellow)-high (Blue) security coding system, geographic areas are assigned a security rating.* As a rule, ARD will not operate in Red areas, which generally are at or near insurgent base camp areas. It will, however, undertake "security support" operations on request in Yellow areas which are adjacent to Red areas. Some villages in Yellow areas will, of course, have poorer security levels than others.

Planning for security support operations takes place outside the normal Blue Book system. As part of the security support planning process, ARD has included use of the Project Allocation Method on two occasions and a third may commence shortly. The purpose of PAM is to assist in the selection of: (1) "types of villages in which ARD infrastructure projects can make the greatest potential

* ARD may try to adjust or refine ISOC intelligence data on the basis of interviews with amphur officials. AIR developed a questionnaire for that purpose. See Handbook, Section VI.

contribution in maintaining the loyalty of threatened villages,"* and (2) the types of development inputs those villages should receive.

In many respects, PAM is a rationale for making explicit a number of choices which the policy-maker or program administrator has to make, and would normally make, in planning and/or conducting a program. This appears to have been true in the ARD instance, where ARD had been involved previously in supporting security operations. The O/I Indexes were simply introduced as part of that decision-making process.**

The PAM commences by requiring or accepting several policy decisions. The first is selection of the target area by ISOC. The second includes decisions by the ARD Secretary-General regarding the most appropriate use of development inputs in Yellow areas, where security can be a problem. The impetus for, and the dominant influence on, ARD's choices in this regard were two memoranda written and circulated by Dr. R. E. Krug of AIR in late 1972 and early 1973.*** The discussion which followed resulted in formal articulation of ARD policy for threatened or "security support" areas: (1) no development

---

* Handbook, p. 3. Descriptions of PAM are found in the Handbook, Chapters III and VI and Appendices D and I; the Final Report, pp. 64-68; and in the AIR Report of Progress, Fifth and Sixth Quarters, January-July, 1973, Appendix A.

** Some observers limit the phrase Project Allocation Method (PAM) to the actual use of the O/I Indexes; in our judgment, the "Method" includes, but goes beyond, the mechanical scoring role assigned to the Indexes.

*** See "Some Thoughts on an Index of Village Security," December 8, 1972 and "Some Rationales for Development Programming," February 22, 1973.

resources would be placed in low security villages because the insurgents would be the most likely beneficiary, (2) non-infrastructure projects, i.e., income-generating activities, would be permitted only in villages ranked high on security; and (3) emphasis and priority in the Yellow areas would be given to villages judged to have the fewest "opportunities," a condition to which a limited range of infrastructure resources exclusively would be directed.

The Secretary-General's choices resulted in a priority ranking displayed in the form of a decision matrix with up to twenty-seven combinations of Opportunity, Investment and Security. The Secretary-General further decided to put projects only into villages which fell within six specified combinations each for infrastructure and non-infrastructure categories, as shown in Table A. As discussed below, the non-infrastructure category was barely used and hence can be ignored. The significant point to note with respect to the infrastructure category is the key role assigned to "low Opportunity." Certainly in the minds of officials using this matrix, the central notion is to seek villages that are low on this dimension and provide them with infrastructure projects.

Given the above decisions as to security, the geographic areas of interest, the respective roles of infrastructure and non-infrastructure inputs, and the preferred programming strategy (only priority villages 1-6 would be eligible for assistance), the O/I scoring technique comes into play. First, teams are sent from Bangkok to collect data for the calculation of O and I scores in

Table A

Infrastructure projects (Roads, water, village projects)

	Opportunity	Investment	Security
1.	Low	High	High
2.	Low	High	Mid
3.	Low	Mid	High
4.	Low	Mid	Mid
5.	Low	Low	High
6.	Low	Low	Mid

Non-Infrastructure (such as COMPAC, Economic-Business groups)

	Opportunity	Investment	Security
1.	High	High	High
2.	Mid	High	High
3.	Low	High	High
4.	High	Mid	High
5.	High	Low	High
6.	Low	Mid	High

Source: Handbook, Appendix E-1.

the target area by means of interviews with Village Headmen. The process used in the first test, covering eight districts in the tri-province area of Loei, Petchaboon and Phitsanuloke, involved two steps. The first was to obtain the average O/I scores from two or three "representative" villages in each tambol using AIR's standard O/I questionnaire. This resulted in the identification of 15 tambol as eligible for ARD programming, i.e., tambol that had villages which would warrant development activi r in accordance with the decision matrix discussed above. The second was to obtain O and I scores for all 136 villages in the 15 tambol and to match them against the matrix. This resulted in a determination that 55 villages were eligible for infrastructure projects and that 23 could qualify for non-infrastructure projects. Application of the decision matrix appears to have been helpful in making ARD's support strategy explicit and facilitating a "sort" among villages.

At this point in the process, the O/I assessment technique drops out as a factor, except for some residual use of answers to several questions on the O/I questionnaire as proxies for hard data on actual village "needs" (discussed below). Continuing to use the tri-province case as an example, the next step in the planning process was to ask changwat officials to propose specific projects for the 78 target villages. Although 78 villages were eligible, the changwat only proposed 25 infrastructure projects and two non-infrastructure projects to ARD. We could obtain no explanation for the elimination of 51 villages altogether, nor information about the

criteria used in doing so, although such a significant reduction raises questions about the value of going through all the preceding "decision steps." With respect to the small number of non-infrastructure project nominations, ARD officials stated that there was some confusion in two of the changwats regarding ARD development policy for non-infrastructure projects. There also appeared to be some confusion in ARD headquarters about the suitability of applying PAM to non-infrastructure projects (mirrored partially in Section III of the Handbook, which suggests that PAM was designed for infrastructure projects, but nonetheless makes some mention of the non-infrastructure category).

In any event, when ARD headquarters receives the changwat project proposals, it undertakes a review or verification procedure. In the first instance this involves a simple check to confirm that the village in question satisfies priority criteria 1-6. Next, the ARD headquarters' reviewers resort to answers given to several questions on the O/I questionnaire previously administered to the Village Headman and use them as measures for judging the validity of the changwat officials' choice of projects. Thus if the changwat requests a well or a road for village X, and the questionnaire response indicates that there is a need for water or for access in that village, the project will be approved. In other words, if the O/I form (questions 2 and 3) answers indicate an inadequate water supply, the village is deemed to have a "water need." If the questionnaire responses also suggest that the school is in poor condition,

headquarters might also direct the addition of a school improvement project in order to take advantage of the equipment that will be in the village. If the changwat proposes a project that runs counter to an answer on the questionnaire, headquarters may disapprove or question the choice.

Actually AIR proposed that the answers to questions pertaining to "needs" (which it called "indicators") were to be used by changwat officials in making their project recommendations, i.e., the questionnaire answers should only be taken as being indicative of need and be subject to modification by the judgment of those in the area. Given the nature of the questions, the responses are at best "indicative" and lack the precision necessary to serve as an accurate description of a "need." Indeed the Handbook states that "actual allocation of projects to villages is a local decision."* Notwithstanding that advice, they have become a more direct proxy for "need" in ARD headquarters.**

Ultimately, ARD approved 24 of the 25 infrastructure projects in the tri-province area for FY 2517 as well as the two non-infrastructure projects -- one each for youth and agri-business -- for FY 2518. It sought a supplemental budget for the former, but as of August 1974 approval had not been received from the Budget Bureau

---

* Handbook, p. III-4.

** There is much to be said for headquarters' willingness to match up the views of the Headmen with those of their officials in the field -- if only to suggest differences of opinion for further investigation. Too few agencies will do this. The danger comes in using data intended only to be an indicator as a somewhat more direct measure of actual project need.

(ARD staff said that "a few" of the projects had commenced using funds at hand). The total estimated cost of the infrastructure projects is 1,370,000 baht, a rather small fraction of ARD's 30,000,000 baht budget for other activities in the three changwat.

A second application of PAM commenced in April 1974 as part of a security-support operation in parts of Kalasin, Sakorn Nakorn, Chiengrai, and Nan. About 480 villages in 46 tambol of eight amphur were surveyed. Changwat proposals had not been submitted as of August and we do not know if non-infrastructure projects will be included this time. Furthermore, given ARD's data processing problems, there may be difficulty in providing the data in useable form, on a timely basis.

During the RJBA Team's visit a third ISOC request was received, asking for ARD support activities in parts of Loei, Udorn and Nonghan. ARD headquarters staff expect the Secretary-General to approve use of PAM in the above process again, i.e., it is becoming standard operating procedure in ARD to consider PAM for "security support" planning activities. As part of its planning for participation in ISOC-inspired CPM operations, ARD must seek the approval of CARD (on which ISOC also sits) and the use of Opportunity and Investment scoring in the selection of villages has become a feature of its presentations. It provides the briefing official with a scientific rationale for selecting villages and projects and no one has chosen to dispute the method or its findings.

In summing up, a variety of points can be made about the Project Allocation Method (PAM). Clearly, the majority of ARD's resources are planned and allocated without reference to PAM, i.e., it is peripheral to and has relatively little influence on the way in which ARD allocates its resources. This is not necessarily an indictment of PAM, but in the context of the PIO/T's interest in exporting operational planning systems it is important to stress that PAM has been applied on a very limited basis. Furthermore, as the Handbook points out with respect to PAM's potential relationship to the Blue Book System, "the actual form of this interaction is as yet unclear."* Thus PAM has been restricted to use in security-inspired settings and it has served principally in the procedural role of providing a decision-matrix for use in ranking villages as suitable for infrastructure projects, following the prior determination of a number of important policy and operational decisions. The answers to some of the questions in the basic questionnaire -- the "indicator" responses -- are also used as rough indicators of "needs" in each village and hence serve as a checklist for headquarters staff in reviewing the propriety of changwat project selections.

Given the serious doubts expressed in Section IV of this report about the reliability of the Opportunity and Investment Indexes, they should not be used as the bases for decisions about specific individual villages receiving or not receiving projects.

---

* Handbook, VI-7.

Project Evaluation

As noted in Section V, the AIR research group devised a Routine Impact Assessment Form (RIAF) for ARD's Evaluation and Reports Division. The "impact assessment" aspect of RIAF, which pertains to individual projects in individual villages, does not involve the O/I Index procedures and really is not related to the basic concept under discussion. Some confusion on this point results in part because the phrase "Impact Assessment" appears in the RIAF title and in part because the RIAF questionnaire contains a separate section which calls for administering the standard O/I questionnaire while the project auditors are in the village. The latter section, of course, has no connection to the particular projects in the village which the Evaluation and Reports Division has come to spotcheck, i.e., the questionnaire is administered on a "target of opportunity" basis.

In order to minimize the aforementioned confusion, a few additional words about the RIAF are in order. The Evaluation and Reports Division is responsible for "evaluating" individual ARD projects, e.g., a well or a road, in specific villages. Historically this has been done by making spotcheck visits to project sites to verify that projects reported as having been completed by the changwat are in fact finished, i.e., performing an audit.

The RIAF was designed to help the Division go a bit beyond the mere audit function to some evaluation of the impact of certain projects, namely, ARD well projects and intravillage infrastructure

projects. ARD's standard rural road projects, for example, are not included. The form is routinely administered by the Evaluation Division in conjunction with its regular spotcheck activities. It is divided into two parts: the first part asks a series of questions to ascertain the degree of villager awareness about the project, its completion, and ARD's paternity of it (as opposed to some other agency), and the level of villager participation in doing the project. Separate sets of questions have been developed for use with water and village improvement projects. The second half of the questionnaire is devoted to collecting O/I data on the general level of village opportunity and investment, using the AIR questionnaire.

In a recent change, the Evaluation and Reports Division developed a modification of the first part of the questionnaire in order to develop some impact information about the COMPAC program. In this instance the spotcheckers did not apply the Part II O/I questionnaire because their informants were COMPAC members, not the Village Headman (whom the ARD staff felt would be inadequately informed about COMPAC). Interviews were held with members of seven COMPAC projects in five changwat (Roi-et, Nonghan, Udorn, Petchaboon, and Uttradit). The projects involved sericulture, livestock, fish, and cash crops. The data are being processed by hand.

Spotchecks and RIAF's have been undertaken in all ARD changwat, except those newly opened, at one time or another. However, the latest computer output is dated February 1974. Since then, spotchecks of some 50 additional projects in 30 changwats have been

collected and coded, but the data have not been processed because of the lack of a continuing arrangement between ARD and the U.S. government processing facility previously used by AIR. Likewise, the computer programs have not been rewritten for use on RTG computers. When the Evaluation and Reports Division does administer Part II of the RIAF and collects the O/I data, it is simply stored. At present no use is being made of it.

Reports on the individual project impact assessments done by the Evaluation and Reports Division are routinely distributed to all Division Chiefs for information, but it is fair to say that they have little or no impact on ARD planning and resource allocation decision-making.