IICA

DIRECCION GENERAL
Apartado 10281
San José, Costa Rica

PROGRAM EVALUATION IN AID

Lessons Learned

By: Herbert D. Turner
PPC/DPRE/PE
July 30, 1976

D10
6000
0000
60
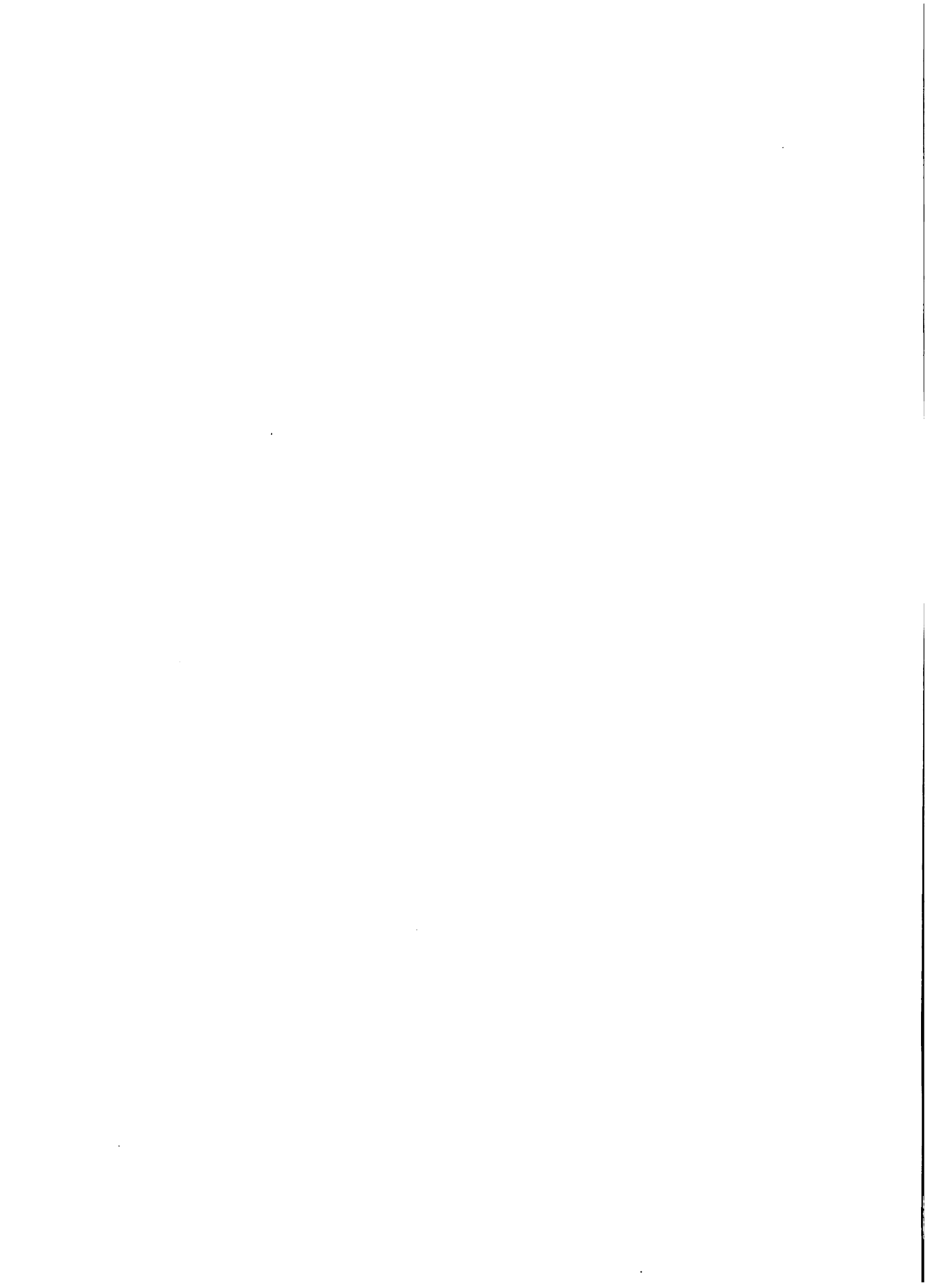
PROGRAM EVALUATION IN AID

Lessons Learned

Herbert D. Turner
PPC/DPRE/PE
July 30, 1976

00005261

001353

PROGRAM EVALUATION IN AID
Lessons Learned

Program evaluation, as an instrument for management and policy formulation, has grown rapidly in developed and developing country governments and multilateral organizations. While producing much useful information, this rapid growth inevitably has been accompanied by confusion about purposes, methods, organizational arrangements and, not least, terminology.

This paper attempts to summarize a few of the principles, methods and operational lessons learned in program evaluation in AID during the past few, eventful years.

DEFINITION

In this paper, we define <u>program evaluation</u> as the retrospective analysis of experience to see if we achieved our stated objectives and to determine how and why it happened. Thus, program evaluation is concerned with results.*

It is useful to differentiate program evaluation from routine implementation monitoring, and at the same time to note their interdependence.

---

*Some development agencies use the term "ex ante evaluation" to describe pre-approval decisions about project feasibility, and the studies upon which these decisions are based. Similarly, they use the term "ex post evaluation" to describe any measurement/verification process occurring after project approval. This paper will not use such terminology.

Program evaluation is a discontinuous function. The evaluator is disengaged from day to day operations and, hopefully, detached emotionally and intellectually from the program/project. The evaluator:

   ...examines the relevance of and need for the project;

   ...questions the design and the underlying assumptions;

   ...assesses induced change and progress toward planned
      targets;

   ...identifies unplanned change;

   ...attempts to identify causal factors and assess their effects;

   ...feeds his findings into redesign and improved execution.

Implementation monitoring is a continuous function. The monitor is intimately engaged in day to day operations and is usually emotionally and intellectually involved in the program/project. The monitor is concerned with:

   ...the procurement, delivery and installation of resource inputs;

   ...adherence to implementation plans;

   ...compliance with required standards and procedures.

In projects evaluated by the project management team, the evaluator and monitor may be one, requiring a periodic shift in attitude and behavior during the evaluation.
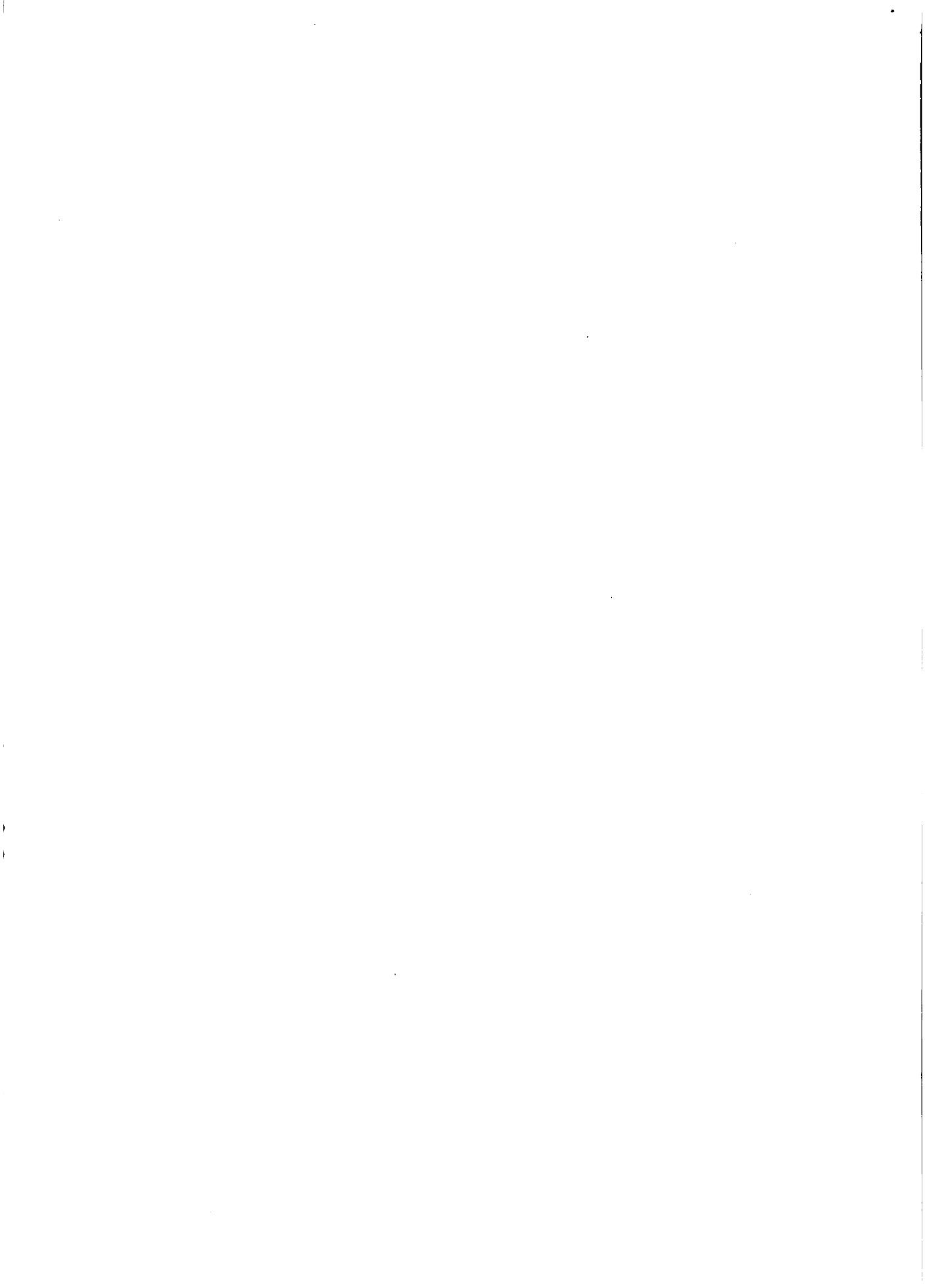
The two functions are interdependent and necessarily somewhat over-
lapping. The monitor generates and collects progress data needed
by the evaluator; the monitor may call for an evaluation when dif-
ficulties arise; the evaluator's findings are translated into
replanning actions and implemented by the monitor.

PURPOSES

The purposes of program evaluation in most development agencies
are several:

1. As an integral element in project management, evaluation is
concerned with project performance or effectiveness and would
include regular periodic evaluations of all ongoing projects on
a decentralized basis for purposes of redesign and improved exe-
cution. This performance evaluation is intended to deal with
three considerations.

    ...Reaffirming the continued relevance of the project in the
    light of changes over time in host country circumstances.

    ...Measuring progress toward planned targets, and where progress
    is discernable during implementation, at the impact level;
   examining unplanned results.

    ...Determining causality, i.e., what internal elements of
    project design and/or what external factors affected pro-
    ject performance and how did they operate.
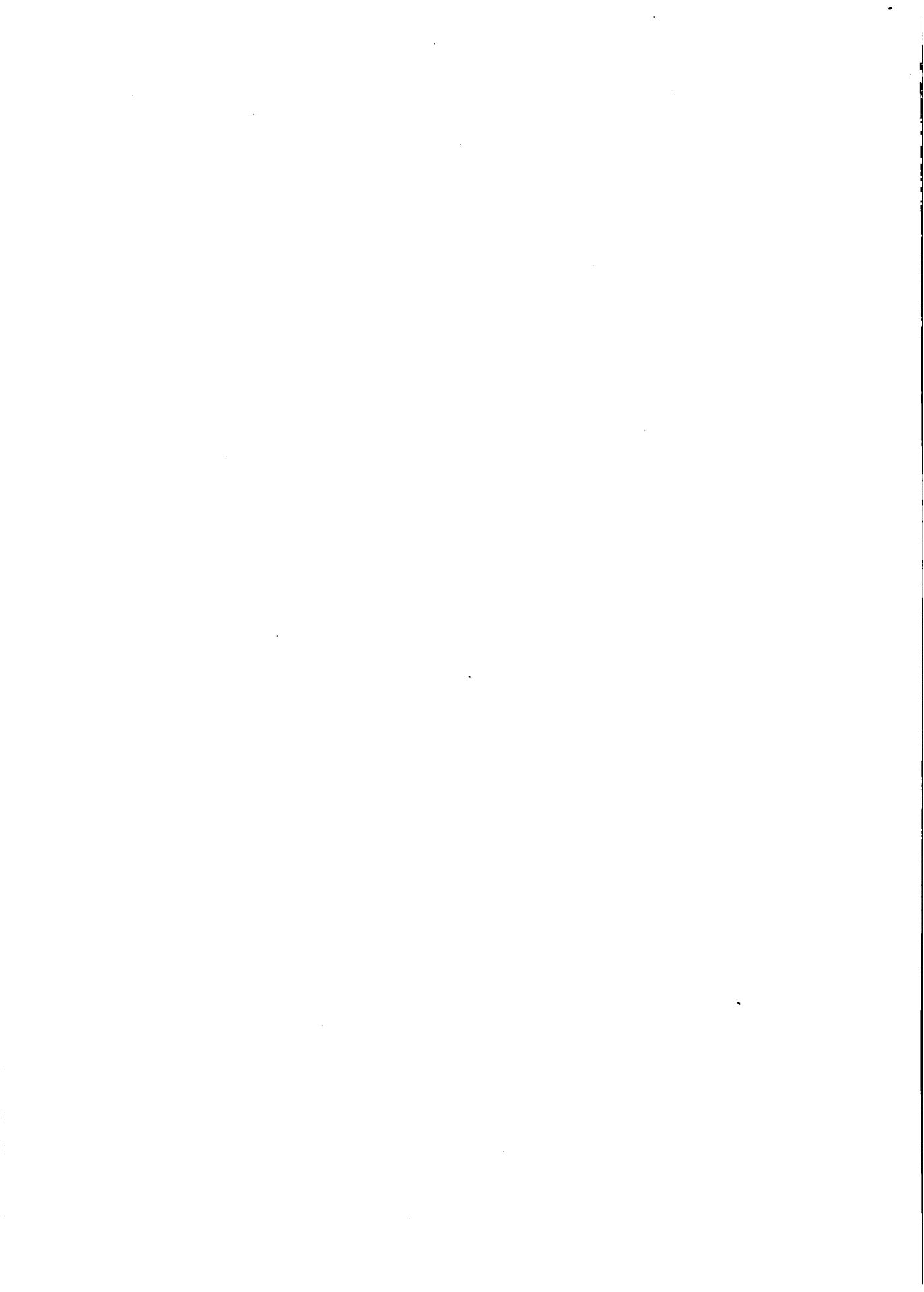
2. As a device for improving resource allocation and program management, evaluation would be concerned with impact or significance and would consist largely of centrally-managed, highly selective ex post studies in depth of the impact of individual projects and sets of projects on development goals. These evaluations would complement the interim impact measurements noted above and would focus on those situations where project impact on sector/program goals was readily apparent only after project termination. The objective is the distillation of operationally useful lessons for application elsewhere.

These indepth evaluations differ from the performance evaluations noted above in that they require a greater investment of skills and other resources as well as more extensive and intensive data collection and analysis.

3. As a means of policy formulation, evaluation is concerned with
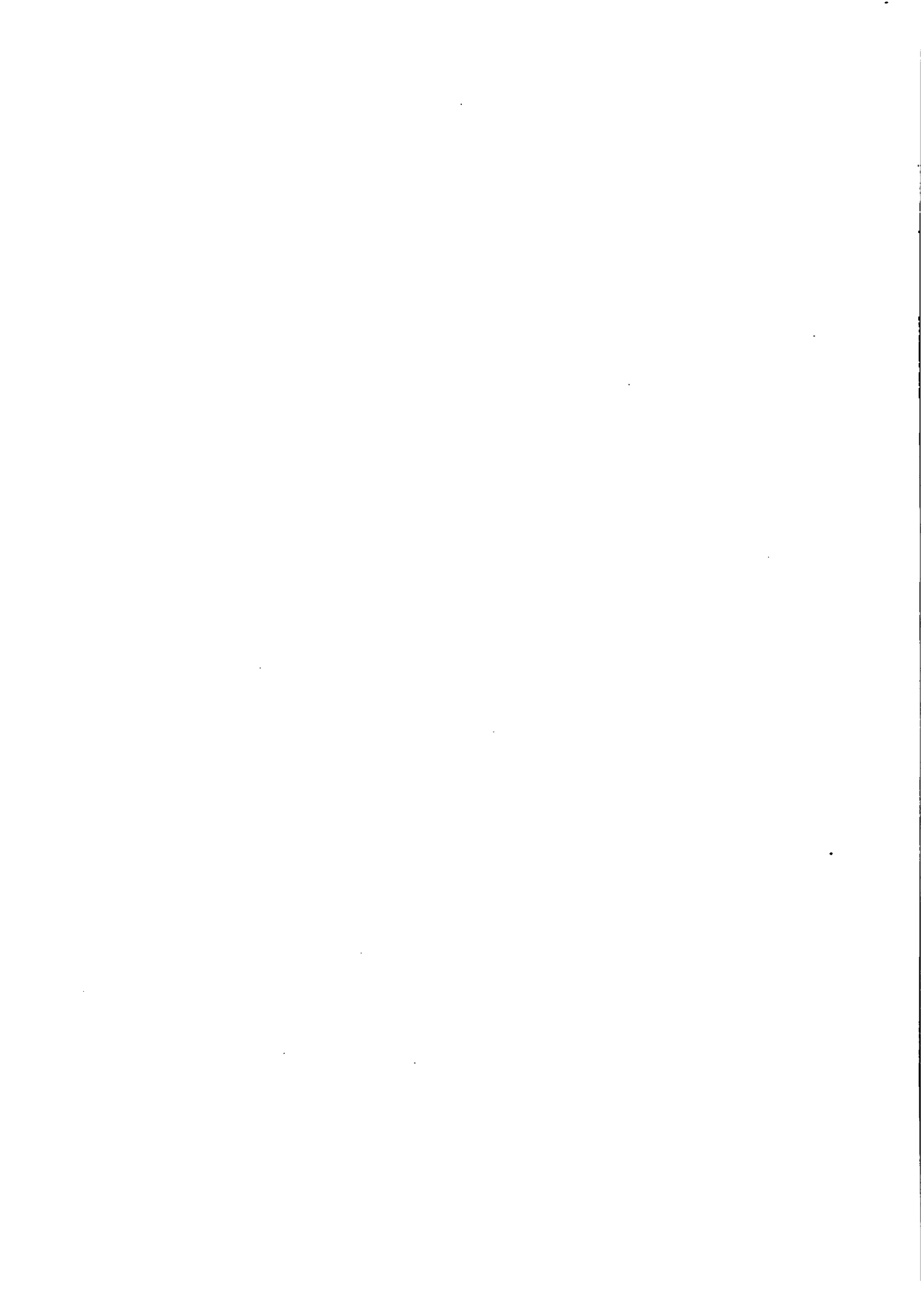
the retrospective examination of program issues which are not country-specific. These would include such subjects as the role of agricultural credit in rural development, the effects of price policies on agricultural production, the impact of land reform/ land tenancy arrangements on rural development, etc.

PRINCIPLES

Several operational principles have emerged from recent experience which govern the design and evaluation practices of a development organization.

1. The effectiveness of the evaluation process is largely dependent upon the quality, explicitness and rigor of program/project design; the quality of design is the major limiting constraint in evaluation.

2. Evaluation must comprehend the total program/project. It is neither feasible nor productive to limit the evaluation process to the fractional resource input of a single source or donor.

3. Evaluation should not be conducted for its own sake or to generate potentially useful information. Evaluation should be decision-driven i.e., evaluation should be undertaken only in response to a need for a decision or for formulation of plans, programs or policies based upon the evaluation findings. Consequently, evaluations findings
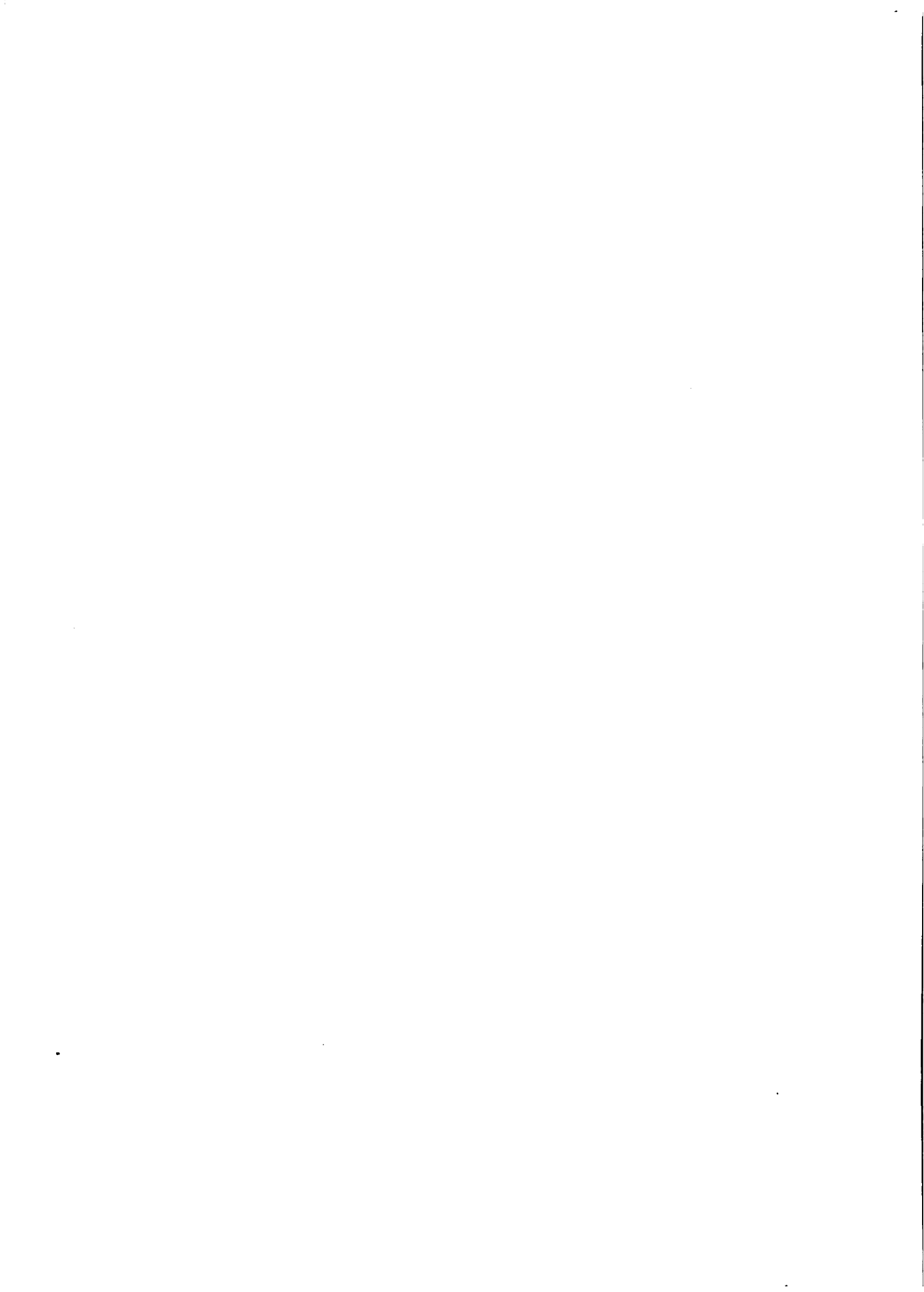
should be packaged and timed for the needs of the decision maker.

4. The responsibility for evaluation should be placed, ·func÷
tionally and organizationally, as close as possible to the user
who will base his decision on the evaluation findings, i.e., the
feed-back loop should be as small as possible.

5. The host country should take the leading role in evaluating
donor assisted projects with the donors playing a supportive role.
Where the host country does not have adequate capacity for evalu-
ation, the donors should offer training in evaluation methodology
both on a general basis and on a project-specific basis.

6. In designing and implementing evaluation studies, maximum use
should be made of host country skills and resources, e.g., local
universities, consulting firms, etc.

7. Achieving maximum transfer and utilization of evaluation findings
to similar projects/programs in other countries, requires an infor-
mation system capable of collecting, storing, matching, retrieving,
and disseminating experiential data.

THE PRECONDITIONS FOR EVALUATION

The key element both in project design and evaluation in AID is the
establishment of a logical framework for the project. The logical

framework assists the designer to structure the project design in the following manner:

1. To define a causal hierarchy of project inputs, outputs, purpose, and higher goal in measurable or objectively verifiable terms:

Project Outputs are defined as the specifically intended kind of results (as opposed to their magnitude) that can be expected from good management of the inputs provided. Outputs are also defined as the preconditions necessary to the achievement of the project purpose.

> Example: Manpower, training, machinery and building materials (inputs) can be managed to produce an irrigation network, trained operational staff, a water utilization schedule and a user rate scale (outputs).

Project Purpose is defined as the primary reason for the project, i.e. the result which is expected to be achieved or the problem which is to be solved if the project is completed successfully and on time.

> Example: An irrigation network and associated facilities and services (outputs) are intended to produce increased per hectare yield (project purpose).

Sector/Program Goal is defined as the programming level beyond the project purpose, i.e., the next higher objective to which
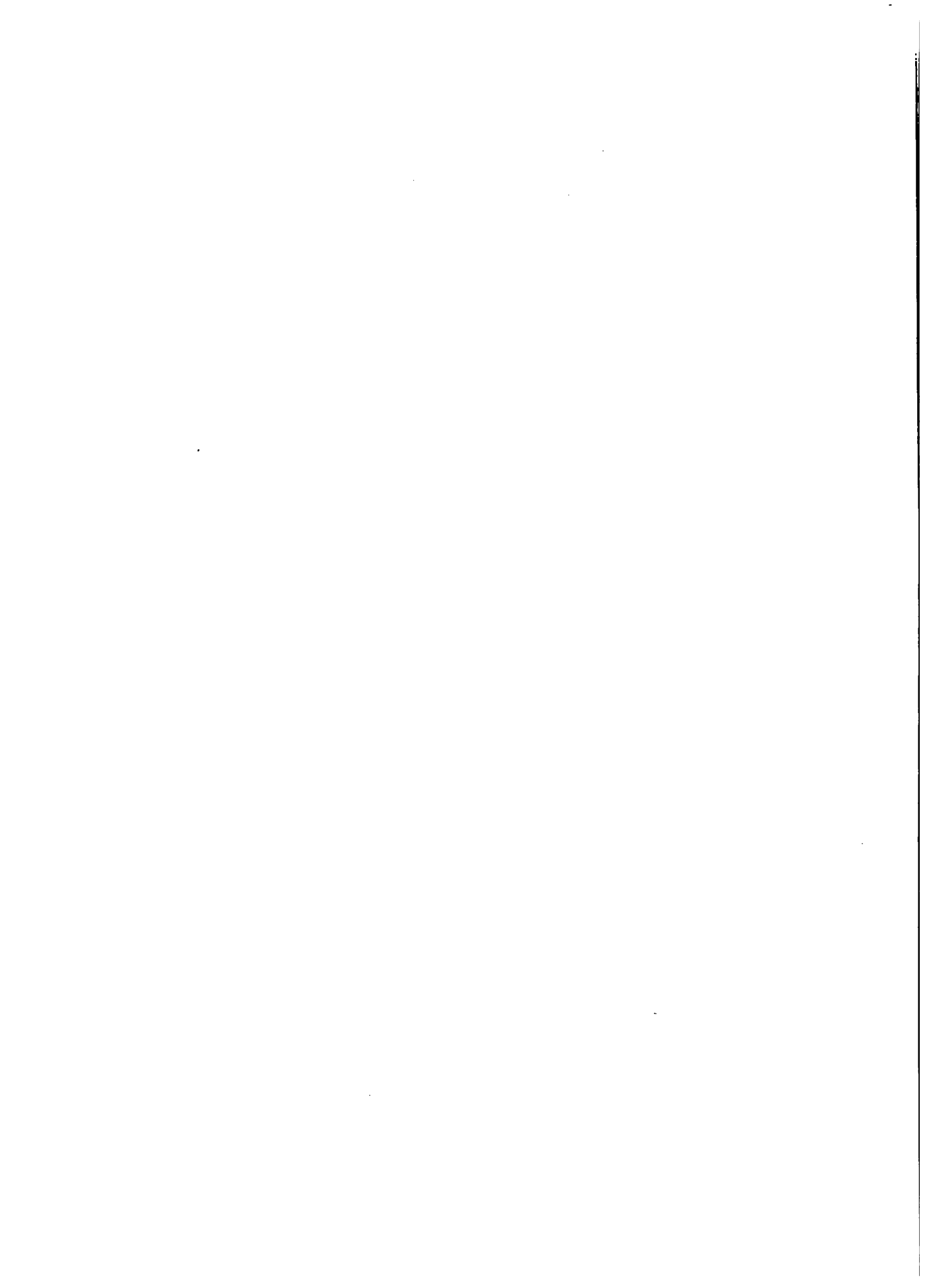
the project is intended to contribute.

> Example: Increased per hectare yield (project purpose)
> is intended to result in expanded exports of agricultural
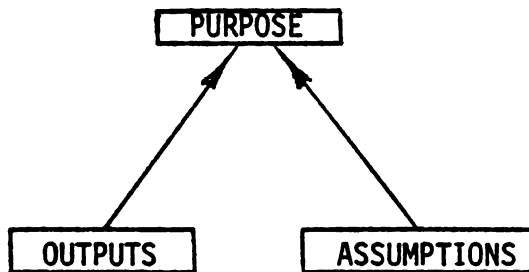> crops (sector goal).

2. To hypothesize the causal (means-end) linkages between inputs,
outputs, purpose and goal. Each of these target levels must
exist before the next can be achieved. The existence of one of
the levels does not however guarantee that the next level will
be achieved. Factors outside the project design (farmer atti-
tudes, access to credit) may prevent achievement of the next level.
Thus the causative relationship between any two levels in the
hierarchy must be stated as a hypothesis, the external factors must
be explicitly identified, and assumptions made about their behavior
and probability of happening.   Evaluation must then verify whether
or not the hypothesis was realized.

3. To articulate the assumptions about external influences and factors
which will affect the causal linkages. Assumptions are defined as
external situations or conditions which must be assumed to exist or
to be brought about if the project is to succeed, but over which
the project management team has little or no control. Example:
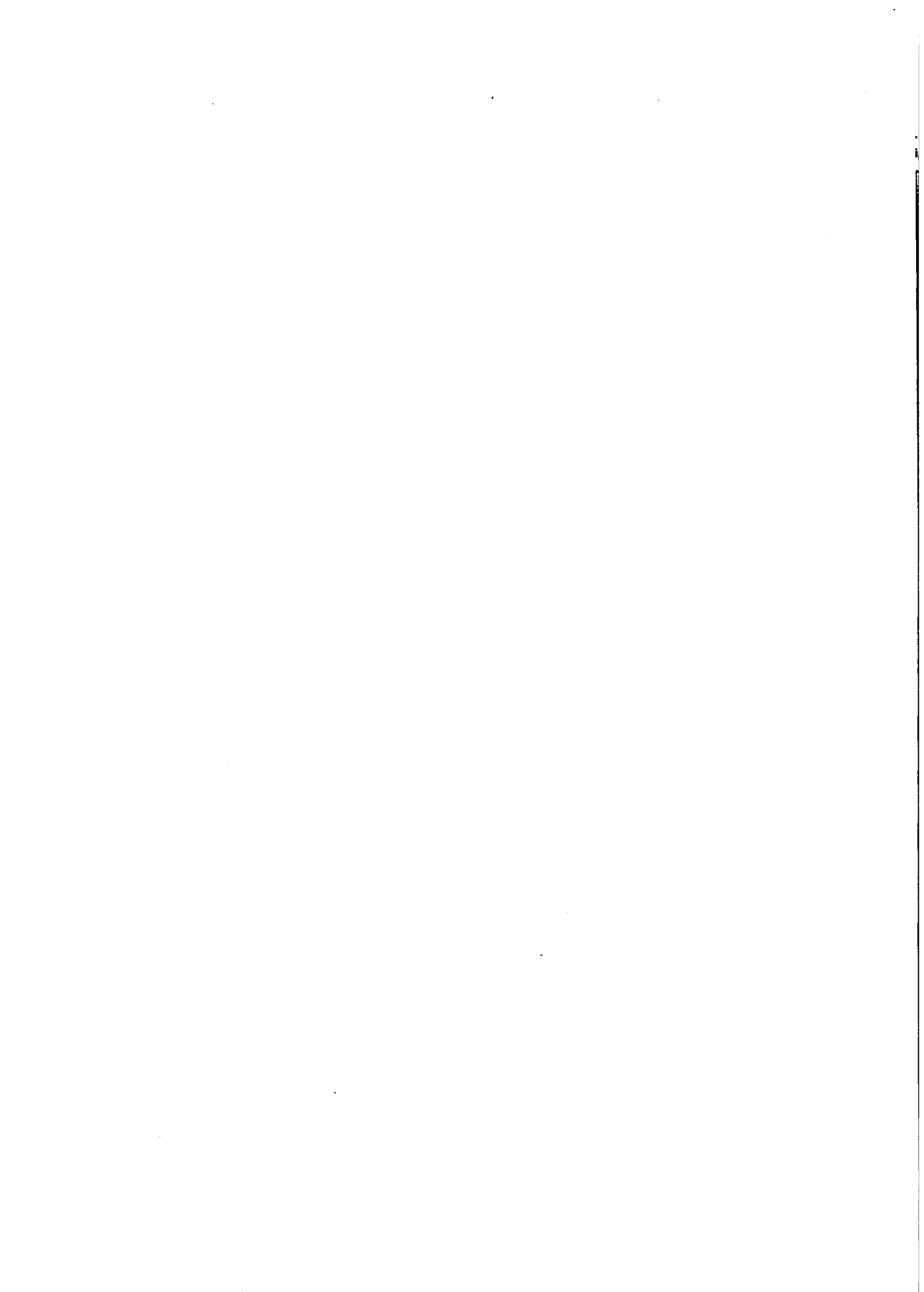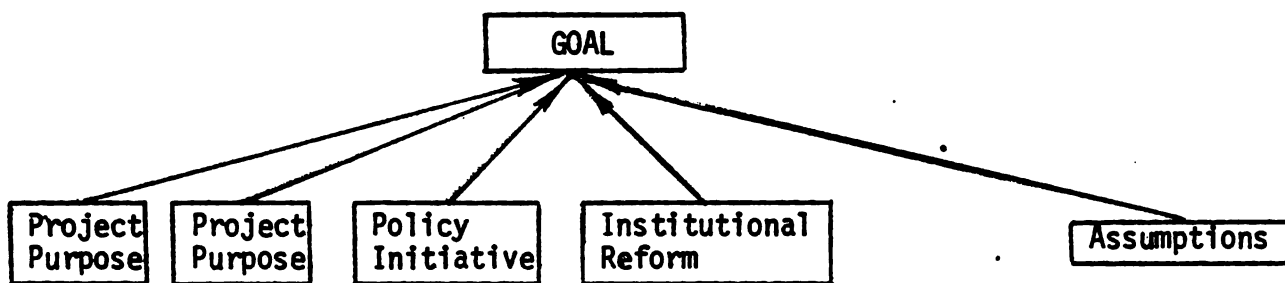Increased crop yield (project purpose) will contribute to expanded

export of agricultural crops (sector goal) only if price and market conditions are favorable (assumption).

The concept of causality rests on the basic premise that each level in the causal hierarchy can be shown to be not only necessary but also sufficient to cause the next higher level to be achieved. Each causal linkage must be tested to assure that a given target level (e.g., outputs), in concert with the assumptions at that level, are necessary and sufficient to achieve the next level (purpose).

```
            ┌──────────────┐
            │   PURPOSE    │
            └──────────────┘
               ↗        ↖
             /            \
            /              \
  ┌──────────────┐  ┌──────────────┐
  │   OUTPUTS    │  │ ASSUMPTIONS  │
  └──────────────┘  └──────────────┘
```

At the purpose-to-goal linkage, the causal model is basically the same although somewhat more complex in scope. This is because the necessary and sufficient conditions must include all projects as well as nonproject initiatives (e.g., institutional, policy and other reforms) - contributing to the goal.

```
                            ┌──────────┐
                            │   GOAL   │
                            └──────────┘
                          ↗   ↑  ↑  ↖         ↖
              ↗      ↗        ↑   ↖              ↖
┌────────┐ ┌────────┐ ┌──────────┐ ┌──────────────┐      ┌──────────────┐
│Project │ │Project │ │Policy    │ │Institutional │   .  │ Assumptions  │
│Purpose │ │Purpose │ │Initiative│ │Reform        │      └──────────────┘
└────────┘ └────────┘ └──────────┘ └──────────────┘
```

4. To establish the progress indicators which will permit subse-
quent measurement or verification of achievement of the defined
outputs, purpose, and goal. Progress indicators are defined as
pre-established criteria or measures of an explicit and specific
nature designed to provide objective assessment of project progress.
Progress indicators should be objectively stated so that both
a proponent of a project and an informed skeptic would agree that
progress has or has not been as planned. Pre-establishing objec-
tively verifiable progress indicators and targets helps focus
discussion on evidence rather than opinions.

The logical framework is primarily a project planning device. It
also is used for re-examination of the original design of ongoing
projects as a necessary prelude to evaluation, i.e., it sets the
stage for determining and validating whether or not the project
outputs are being produced, whether these outputs in fact are
serving to achieve the project purpose; and finally whether this
achievement is making a significant contribution, as planned, to
the higher order goal.
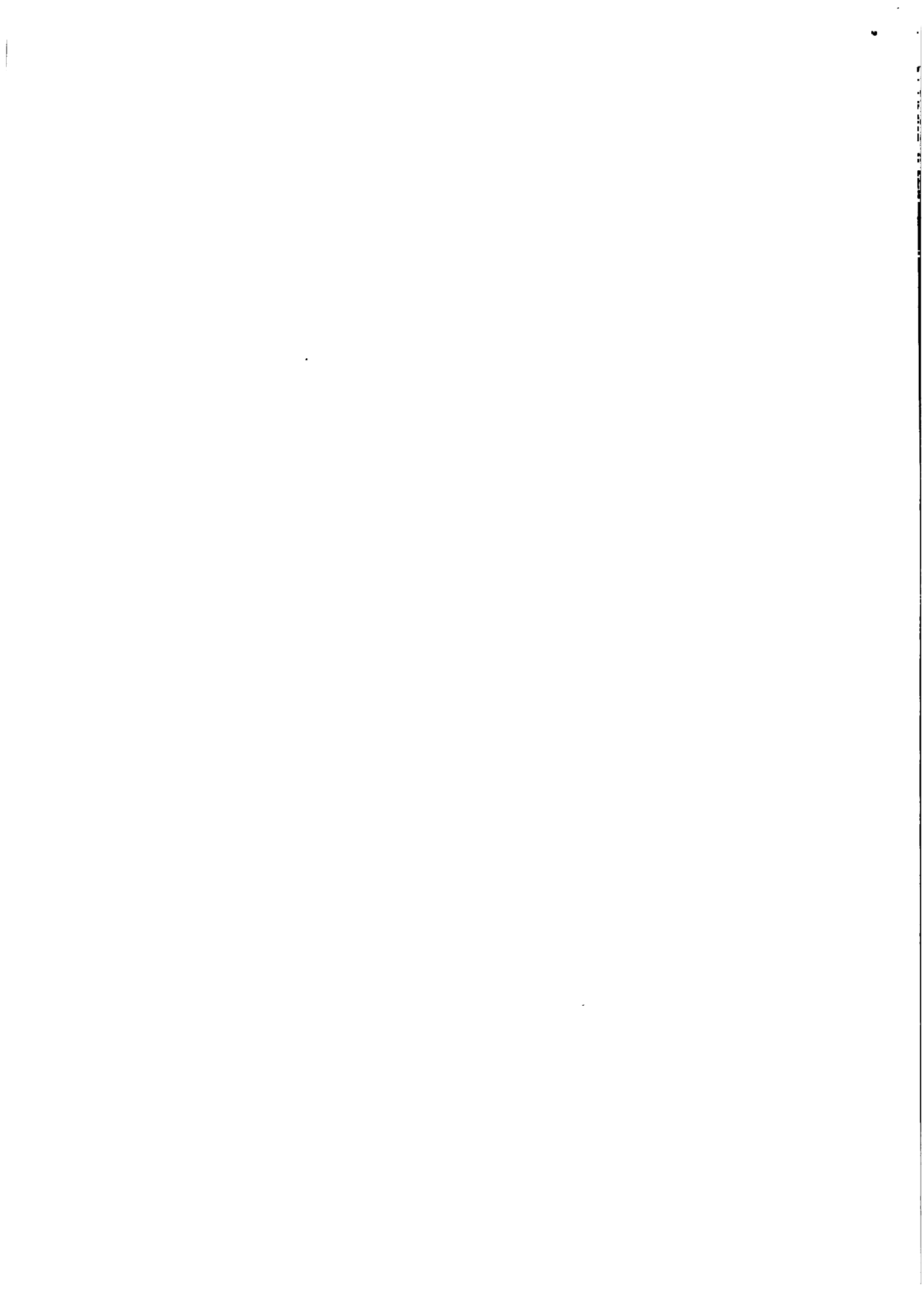
The logical framework also is used to set the practical limits of
responsibility of project management. Articulating the project
planning assumptions in explicit and operational terms permits
a clearer separation between manageable interests and those
factors which appear to be beyond the control of the project
management team. The input-to-output level is largely suscept-
ible to managerial control with relatively few uncontrollable
external factors. At the output-to-purpose level, the possibility
of managerial control decreases while external factors become more
important. At the purpose-to-goal level, the ability of project
management to predict and control events usually is further dimin-
ished.

In addition to the logical framework elements noted above, the
project design process must also contain such evaluation information
and actions as:

  ...collection of baseline data;

  ...review of prior experience with similar projects elsewhere,
     to the extent possible;

  ...provision for experimental, quasi-experimental or other
     evaluation approaches.

  ...establishment of schedules for recurring evaluations with
   timing keyed to decision-making.
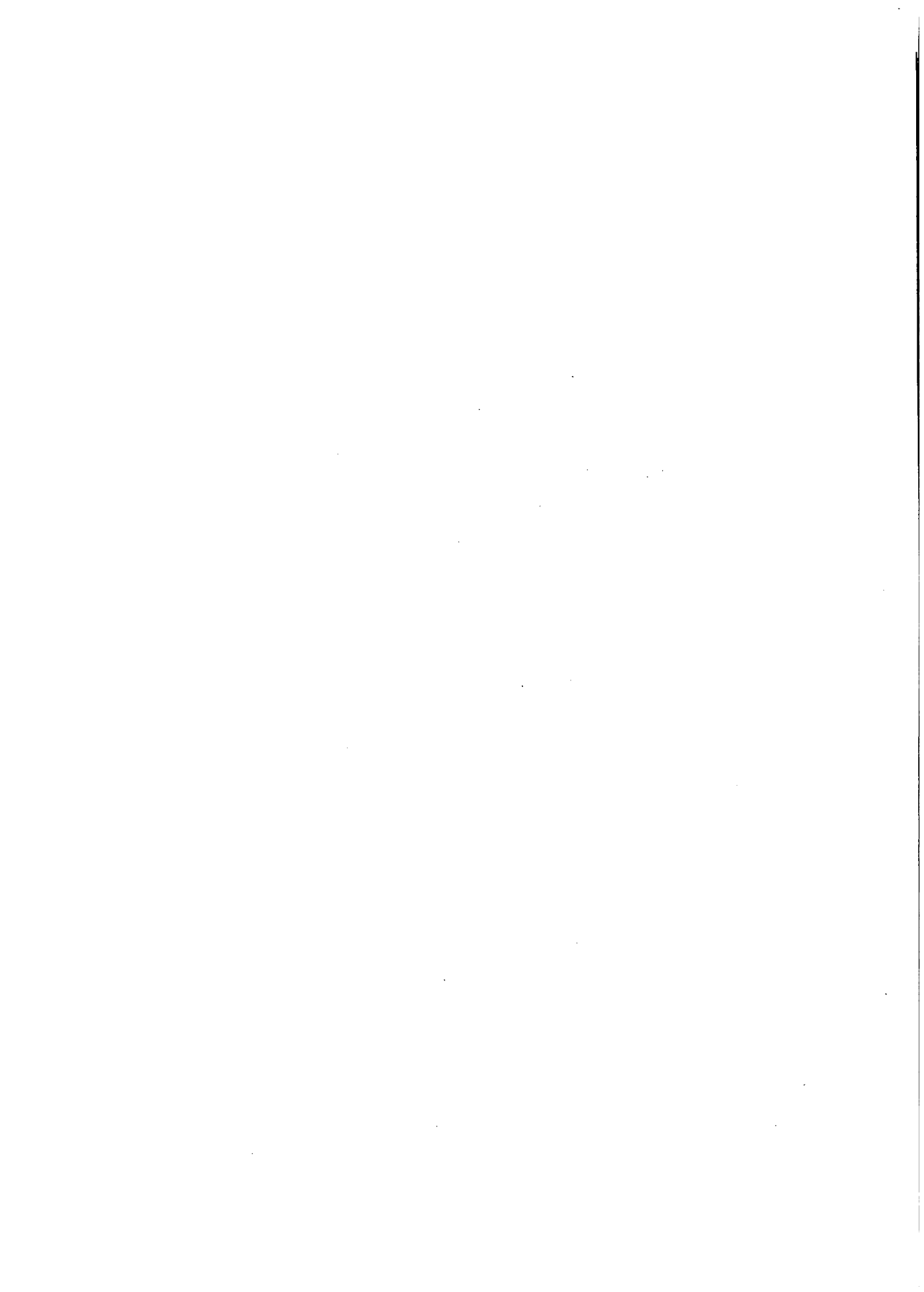
For ongoing projects, if these evaluation elements and actions were not built into the design, evaluation will be difficult.

Experience in evaluating ongoing development programs/projects has shown:

For most projects it is not possible to evaluate progress toward established targets in a meaningful way until (1) the evaluator has considered any changes in the host country socioeconomic setting which may have significantly affected the project, and (2) the existing project design has been reexamined and clarified.✓

The first of these considerations, the assessment of changes in the host country socioeconomic setting, may well be demanding and difficult, requiring data collection and analysis as well as extensive discussion with all collaborating agencies. It includes such questions as:

a. Changes in the nature and magnitude of the problem to which the project is addressed.

b. The continuing validity of original feasibility data and estimates.

c. Changes in physical and environmental conditions.

d. Changes in demand and other economic variables.
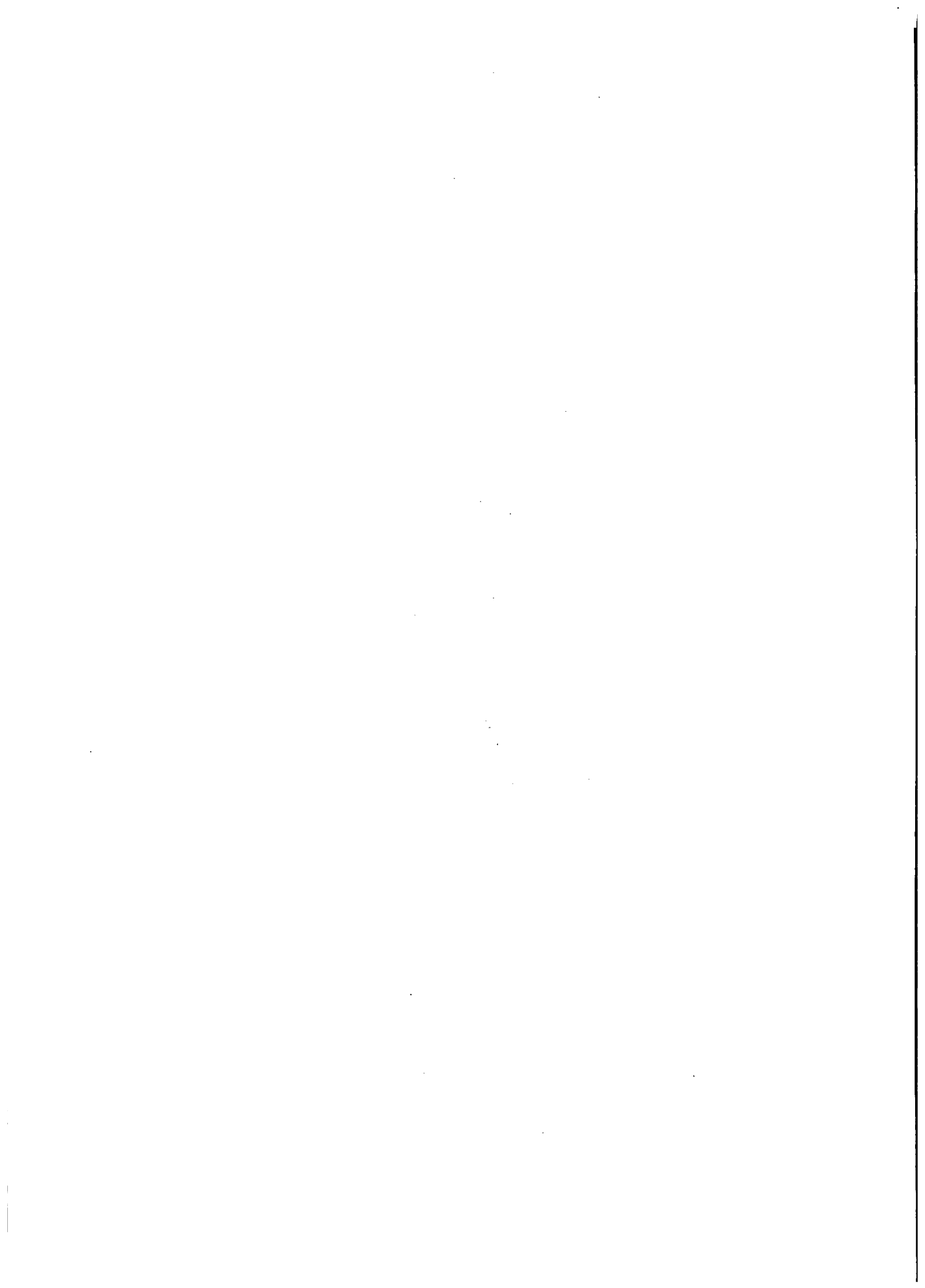
e. Changes in attitudes and other social variables.

f. Changes in host country development policies and
   priorities.

\# When this assessment of changes in the project setting
   results in substantial redesign, it may not be possible
   to evaluate progress toward the new targets immediately.
   Depending upon the extent of the redesign and the revised
   implementation plan, it may be appropriate to postpone
   the evaluation for a year or more.

KINDS OF EVALUATION

There are several forms and types of evaluations used in develop-
ment organizations.

Formative evaluation is used when the project/purpose is not
readily definable in precise and explicit terms and when the
strategy for achieving that objective is even less clearly under-
stood. In such circumstances formative evaluation is used period-
ically to explore experience to date in order to permit managers
progressively to sharpen the definition of the purpose and to
formulate a viable strategy for achieving it.
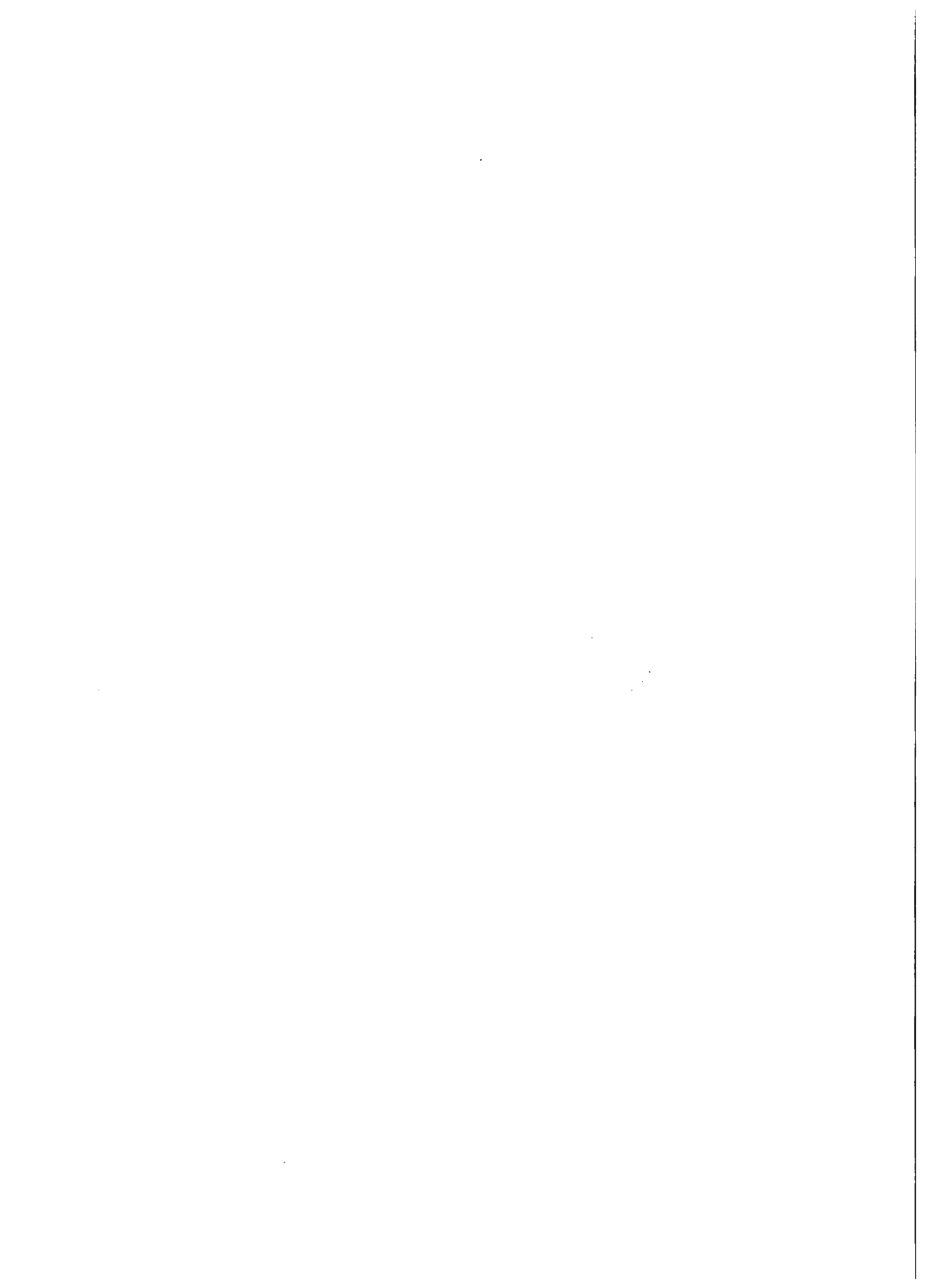
Accumulation

√ Summative evaluation is used when the objective is clearly defined and there exists a high level of confidence in the strategy for achieving that objective. In this instance summative evaluation merely attempts to measure progress towards the objective,

The goal attainment evaluation model is used when the project/ programme has a single predominant objective. This model is widely known and used; it is relatively low-cost and imposes modest skill requirements.

The systems evaluation model ·is utilized for institution building and for complex projects/programmes with multiple (and possibly competing or conflicting) objectives. It is not yet widely used because it is not well understood and its cost is high in terms of the skills required.

Many institutions engage in an either/or debate about the comparative advantages of ongoing evaluation vs expost evaluation of completed programs/projects. Both are needed.

Evaluation of ongoing projects is a program management function intended to:

    # assure the efficient use of scarce resources;
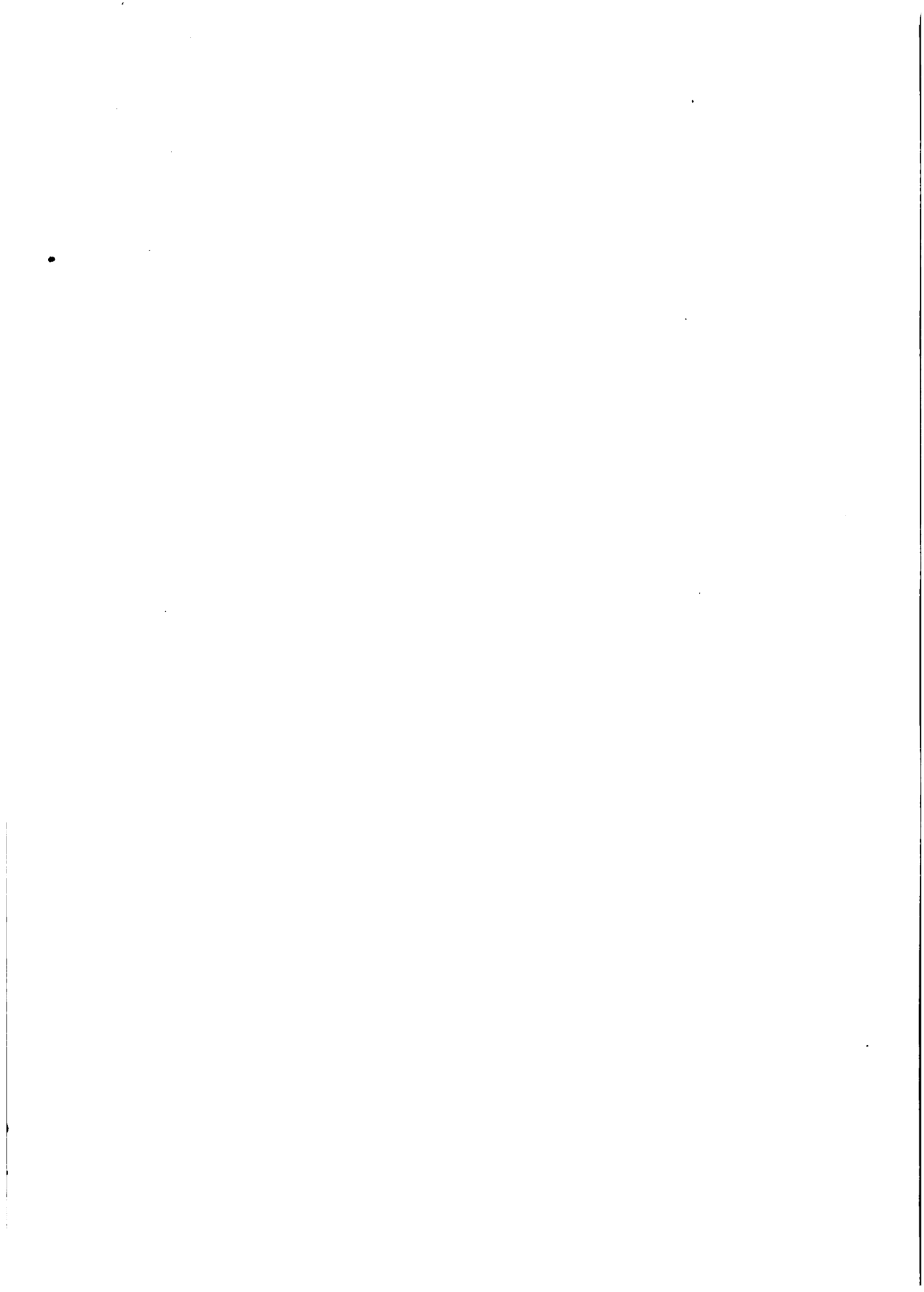
> # measure effectiveness in achieving planned project objectives;

> # confirm the continuing relevance of the project in the light of changing circumstances.

<u>Ex post evaluation of completed projects is a program policy and management function</u> and is intended to:

> # give a final accounting to the sponsors and supporters of the project;

> # identify possible follow-on opportunities for the host country;

> # derive lessons from past experience which can be applied to the design of similar projects and to the formulation of program criteria and policy.

MEASUREMENT AND VERIFICATION

The design of the program/project rests on a structured hierarchy of objectives (inputs-outputs-purpose-sector/program goal) connected by <u>hypothesized</u> causal linkages (e.g., if outputs, then purpose). The task of the evaluator is to verify these hypotheses.

If the project design provides for the evaluative preconditions described above, then measurement and verification can be relatively simple and routine, i.e., measuring change using the previously formulated progress indicators and drawing on the progress data collected during implementation.
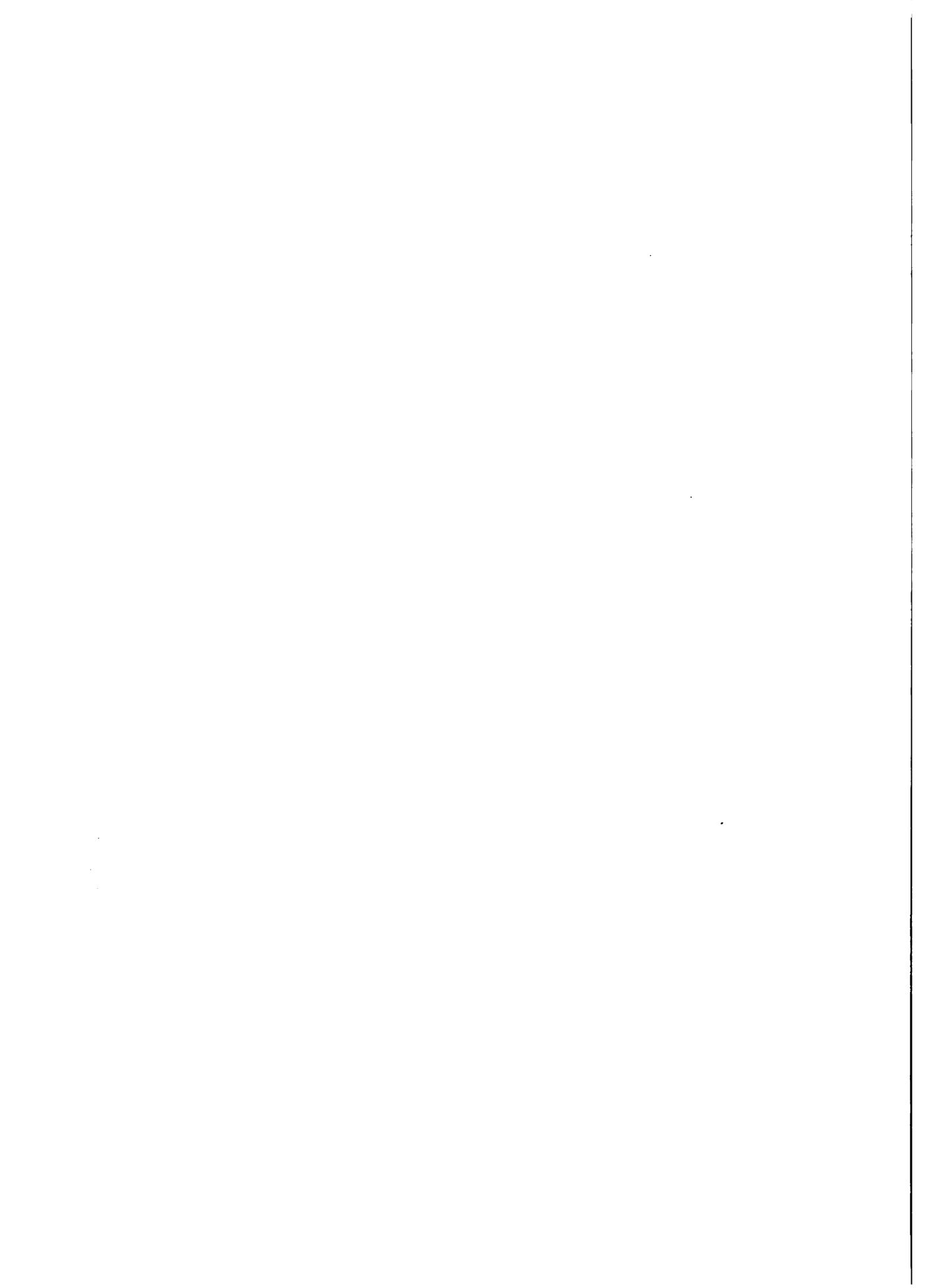
Program and policy evaluation, involving aggregations and inter-
country comparisons of projects and programs generally follows the
same basic evaluative process as is followed in project evaluation.

If the desired changes did not occur - or if they occurred and an
explanation is needed of how and why they occurred - then the
evaluator must delve deeper.

The resources available to AID and the operational conditions in
the less developed countries have inhibited extensive use of
experimental or quasi experimental design, except in special cases.
This does not reduce the need for analytical rigor in design and
evaluation. It is possible for the evaluator to simulate the rigor
of a controlled experiment by the proper use of baseline data,
explicit definition of targets, and the retrospective verification
of project hypotheses.

Such retrospective verification includes careful examination of the
nature of the processes through which the hypothesized causal
linkages occurred or failed to occur. Verification/validation
of the hypotheses might be approached through direct observation;
the use of proxy or surrogate measures; examination of differences
in results/impact among target subgroups within the treated areas,
and of the independent variables associated with those differences;
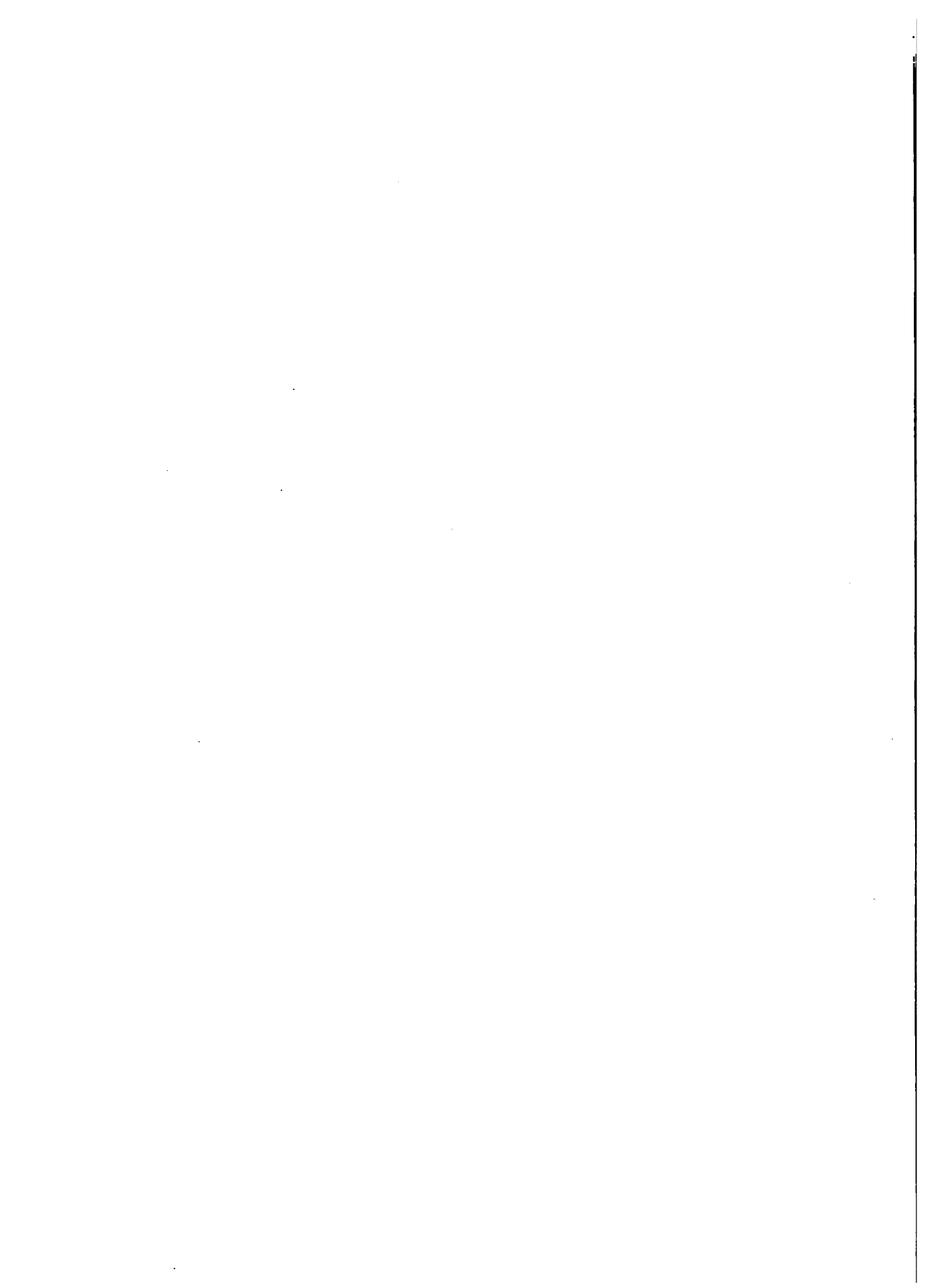
statistical techniques such a multivariate analysis; etc.

Verification can also be approached by formulating alternative hypotheses to explain changes. If none can be developed or if plausible alternative explanations can be disproven, this suggests at least partial validation for the original hypotheses. Conversely, if the alternative hypotheses can be supported to the point that they become persuasive, then the original hypotheses lose credibility.

While such retrospective methods would appear less satisfactory than the classical experimental design, the operational realities of a development assistance program usually favor the former.

The evaluation findings should be reviewed to establish implications and conclusions for further action. The review should include all interested parties: the host country, the implementing agents, all donors. The climate should be collaborative and constructive, not threatening; replanning, not recrimination is the intent.

ROLES AND RELATIONSHIPS

Two closely related pairs of issues require thoughtful consideration.
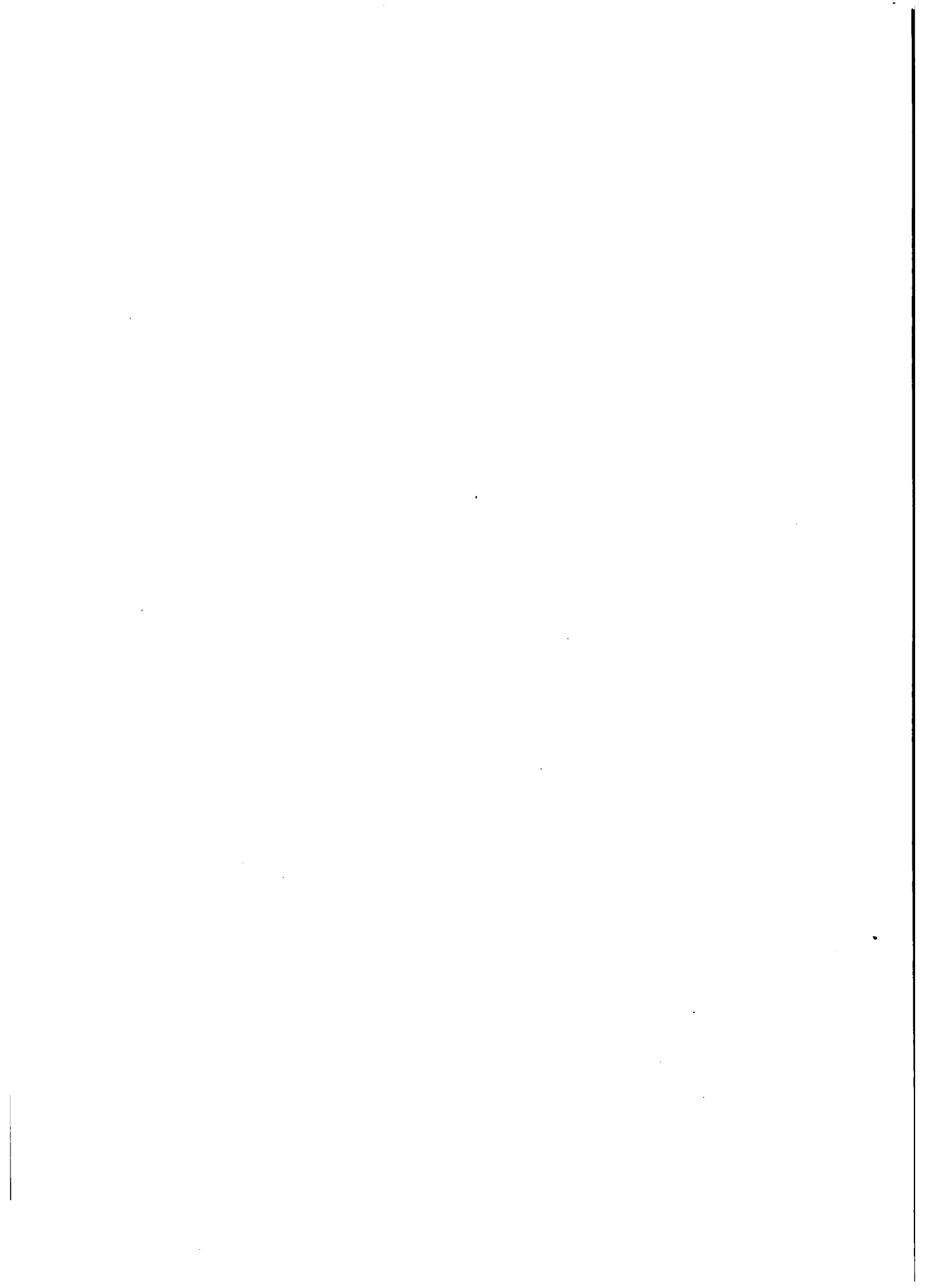
1. Should the evaluation responsibility be centralized (i.e.,
evaluations performed by a central evaluation unit) or decentra-
lized?

The critical criterion is that responsibility for conducting
evaluations should be located functionally and organizationally
as close to the decision-making point as possible, i.e., the
feedback of findings into decision-making should be as direct
and immediate as possible. Application of this criterion suggests
that:

- \# evaluation of progress and of continuing relevance of
  individual projects be decentralized to the project
  management team.

- \# evaluation of experience to formulate policy and program
  criteria be done by a centralized unit concerned with
  policy and program coordination.

- \# evaluation of programs and groups of projects for program
  management purposes be done at an intermediate level, e.g.,
  regional or technical offices.

2. Should evaluations be conducted by in-house staff (e.g., those
with program/project implementation responsibilities) or by outside
experts? Operational experience shows that project managers can
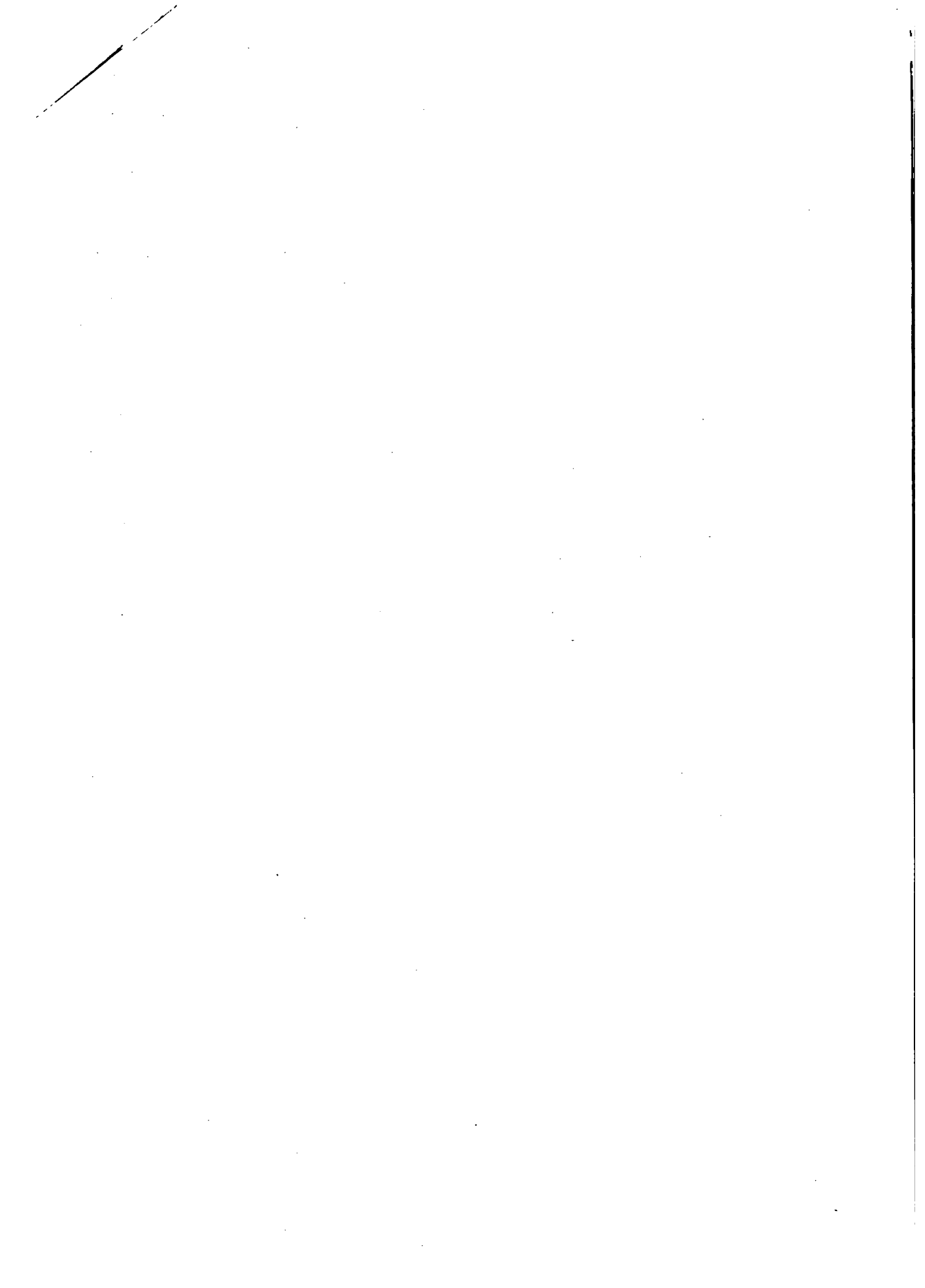
evaluate progress and the continuing relevancy of their own projects
candidly and objectively if there is a climate of constructive
inquiry rather than a climate of recrimination. There are simple
criteria to guide the choice.

IN-HOUSE	OUTSIDE EXPERTS
-- Greater knowledge of the environment and of specific operations is required.	-- Disinterested objectivity is paramount and must be demonstrated.
-- Direct and immediate feed-back of findings into replanning.	-- Greater and more recent technical knowledge is required.
-- Lower unit cost - broader evaluation coverage.	-- Evaluator must be free from operational workload for extended period.

Basically, it should not be an either/or choice. Both kinds of
evaluation are needed and a combination of inhouse and outside ex-
perts often provides the best arrangement.

Irrespective of the extent to which program evaluation is decentral-
ized, there is a need for a central program evaluation unit to assure
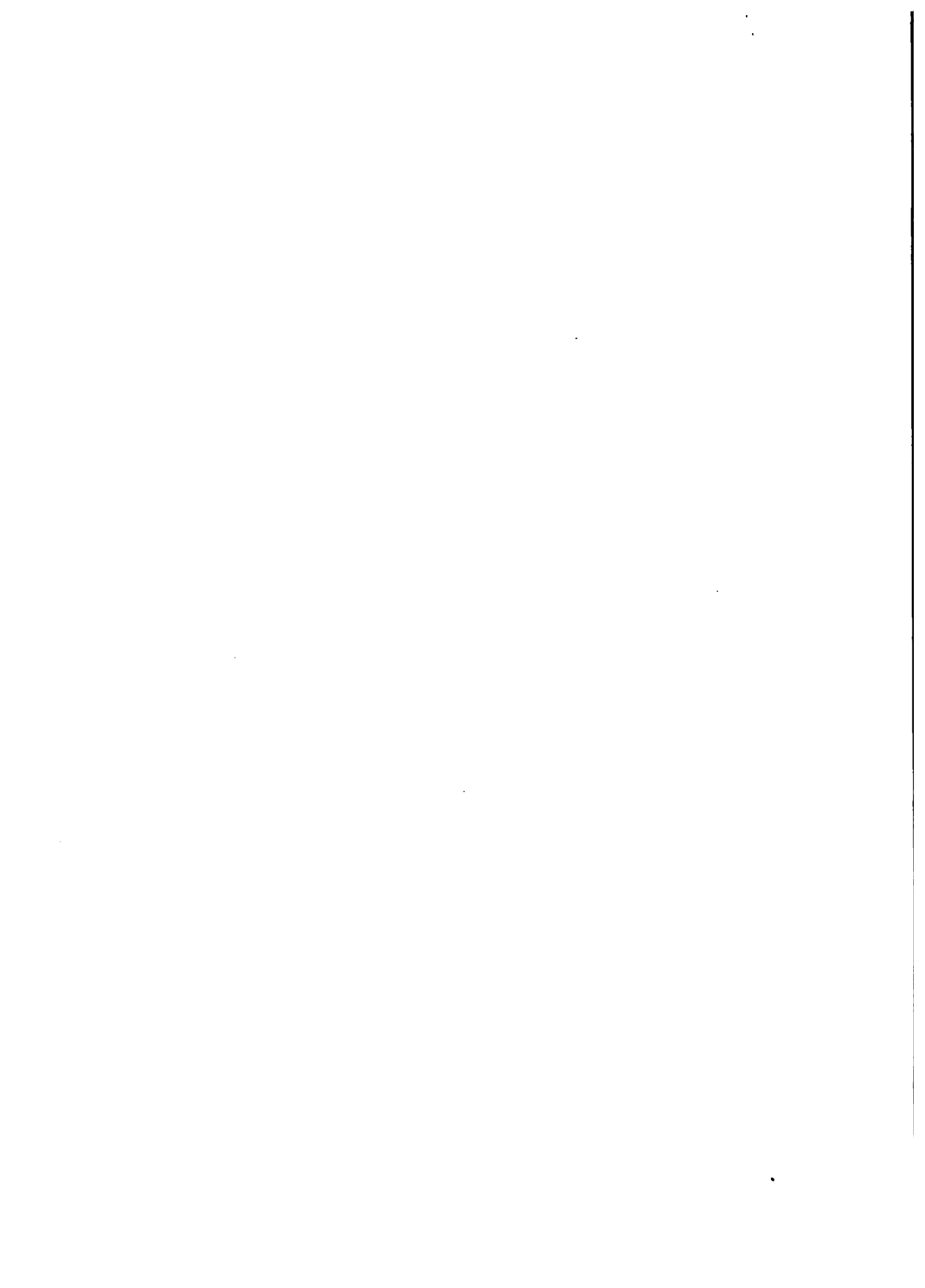that evaluation is occurring and:

   ...is professional/rigorous;

   ...addresses operational concerns;
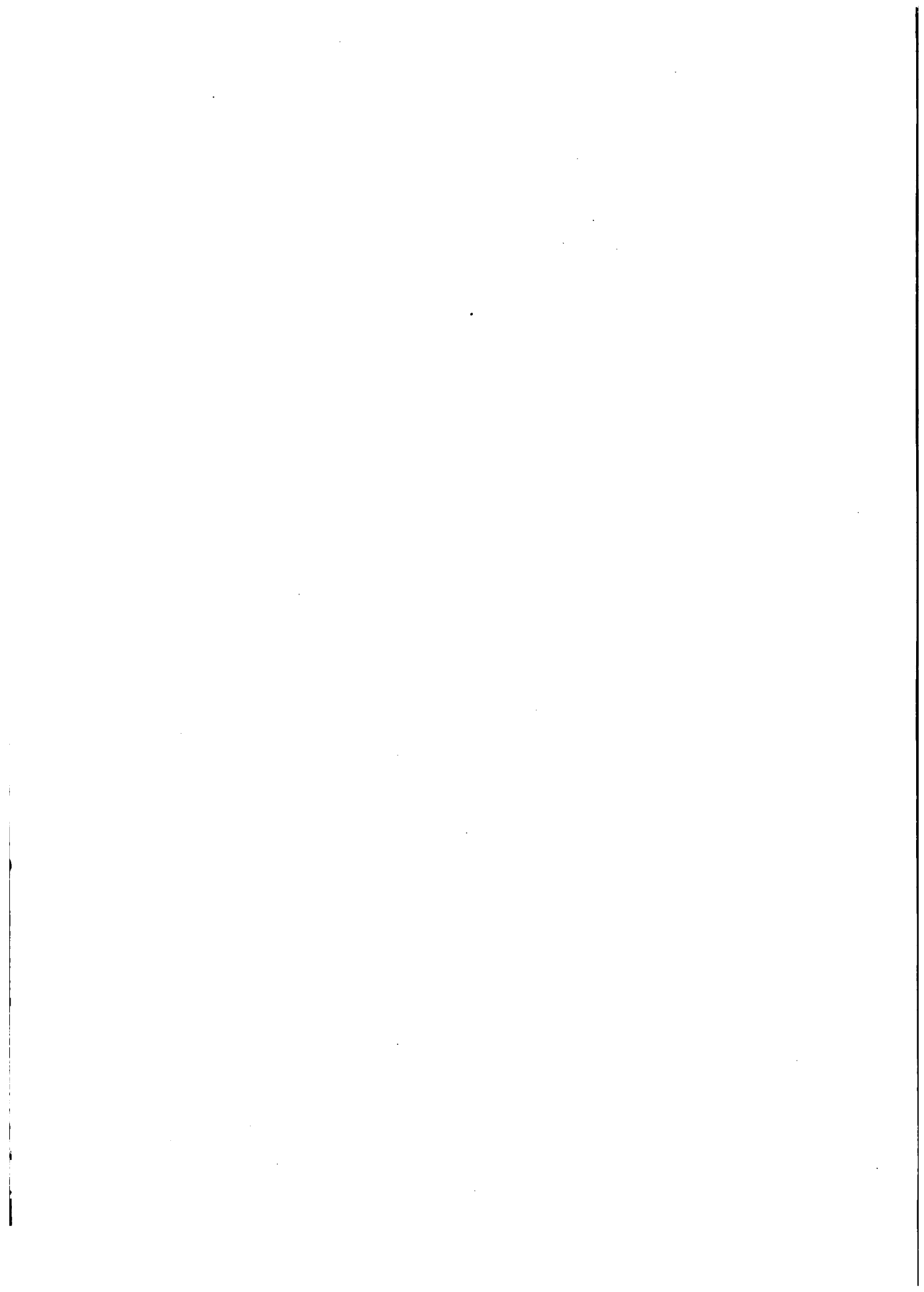
...evaluation components are built into programs and projects;

...effective methods/techniques are available, understood, used
   skillfully;

...findings are channelled into replanning and decision-making.

The services which such a central program evaluation unit might
give to evaluators in the organization would include:

   # coordinate evaluation activities and the preparation
      of annual evaluation plans;

   # advise on, and participate in evaluations;

   # administer evaluation contracts and consultant services;

   # develop improved methodology, guidance, instructional
      material;

   # provide training in evaluation methodology;

   # provide central evaluation document and reference
      service.

FECHA DE DEVOLUCION			
17 NOV 1988			
14 MAR 1990			