DOCUMENT RESUME

ED 214 983

TM 820 ,235

AUTHOR   Stanfield, Jonathan
TITLE   Management Review of Evaluation Practice.
INSTITUTION   Northwest Regional Educational Lab., Portland,
                Oreg.
SPONS AGENCY   National Inst. of Education (ED), Washington, D.C.
REPORT NO   NWREL-REP-58
PUB DATE   Jun 81
CONTRACT   400-80-0105
NOTE   70p.

EDRS PRICE   MF01/PC03 Plus Postage.
DESCRIPTORS   *Administration; *Consultants; Elementary Secondary
                Education; Evaluators; *Management Systems; *Program
                Evaluation; Staff Role; *State Departments of
                Education
IDENTIFIERS   California; *Evaluation Research; Montana;
                Washington

ABSTRACT
        Evaluation practice within state education agencies
(SEAs) is reviewed from a management consultant's perspective. The
study is based upon a review of literature, discussions with the
Northwest Regional Educational Laboratory Research on Evaluation
Program, and visits to SEAs in California, Montana and Washington.
The main findings of the study deal with the current practice of
evaluation in SEAs, and with management metaphors which throw light
upon the evaluation unit considered as a supplier of services within
the SEA, and upon the unit's clients who are its customers for such
services. The central theme of the study is that the basic paradigm
underlying SEA evaluation is in the process of change, moving from
the big study, advanced methodology glamor of the late 1960s to an
organization-based function supporting SEA management in a variety of
ways, alongside other staff specialists. The study suggests that in
the long term the SEA evaluation function will condense into two sets
of roles, namely, that of evaluation systems analysis, and a
consulting role. The study then goes on to identify three possible
intervention strategies: (1) clarification and confirmation of the
new paradigm; (2) assistance to the supply side of evaluation
services; and (3) the building up of informed and understanding
demand among evaluation clients. (Author/GK)

No. 58   MANAGEMENT REVIEW OF

EVALUATION PRACTICE

JONATHAN STANFIELD

Private Consultant
Kirkland, Washington

June 1981

Nick L. Smith, Director
Research on Evaluation Program
Northwest Regional Educational Laboratory
300 S. W. Sixth Avenue, Portland, Oregon 97204

PREFACE

The Research on Evaluation Program is a Northwest Regional
Educational Laboratory project of research, development, testing,
and training designed to create new evaluation methodologies for
use in education. This document is one of a series of papers and
reports produced by program staff, visiting scholars, adjunct
scholars, and project collaborators—all members of a cooperative
network of colleagues working on the development of new
methodologies.

What is the nature of evaluation practice in state departments of
education when viewed by a management consultant rather than a
traditional educational researcher? What management perspectives
might be used to better understand the operation of state
department evaluation units? These and related questions are
answered in this report by Dr. Jonathan Stanfield, an independent
management consultant who conducted a review of the state
evaluation units in Washington, California, and Montana. This
report provides an illustration of a management consulting review
study in evaluation as well as offering insightful observations
about the nature of evaluation practice in state departments.

Nick L. Smith, Editor
Paper and Report Series

CONTENTS

SUMMARY

This study has reviewed evaluation practice within
state education agencies (SEAs) from a management
consultant's perspective, as part of the NWREL
Research on Evaluation Program's ongoing enquiry
into the utilization and effectiveness of educational
evaluation.

The study is based upon a review of literature,
discussions with the Program, and visits to SEAs in
California, Montana and Washington.

The main findings of the study deal with the current
practice of evaluation in SEAs, and with management
metaphors which throw light upon the evaluation unit
considered as a supplier of services within the SEA,
and upon the unit's clients who are its customers
for such services. Evaluation is not a homogeneous
function, but in actual practice embraces at least
four separate strands of activity, namely, research
studies, computerized information systems, testing,
and expert assistance.

The central theme of the study is that the basic
paradigm underlying SEA evaluation is in the process
of change, moving from the big study; advanced
methodology glamor of the late 1960s to an organ-
ization-based function supporting SEA management
in a variety of ways, alongside other staff special-
ists, e.g., in budget and finance. From this it is
possible to show why some of the current issues
raised by the evaluation community have such a
frustrating quality. These include the lack of

impact of evaluation on policy, the failure of
expectations as to what evaluation can deliver, and
the inapplicability of prescriptive methodologies
originating in academe.

The study suggests that in the long term the SEA
evaluation function will condense into two sets of
roles, namely, that of evaluation systems analysis,
and a consulting role.   This concept of an
evaluation systems analyst may be compared to other
applications systems analysts in data processing,
e.g., financial, retail, engineering systems analysts.
As in all jobs there will be ranges of seniority and
emphasis.   Some will work closely with computers,
and others with the design of data collection schemes
for specific evaluations.   Some will be entry level
positions, and others will be senior technical staff.
Because of their organizational context, the rate and
extent of such changes depend greatly upon the
management style of the host SEA.

The study then goes on to identify three possible
Program intervention strategies: one dealing with
clarification and confirmation of the new paradigm
underlying SEA evaluation, one which seeks to assist
the supply side of evaluation services, and one which
is ad   to the building up of informed and
unders   demand among evaluation clients.
Recomme   s are given for Program activities
in each of these categories.

8

1. INTRODUCTION

This is the final report on the study: <u>Management Review of Evaluation Practice</u>, carried out by the consultant for the Northwest Regional Educational Laboratory (NWREL), under the direction of Dr Nick Smith, Director, Research on Evaluation Program.

The study reviews evaluation practice within state education agencies (SEAs) from a management consultant's perspective. Accordingly, this report is written primarily as a consulting report to a client (the Program), rather than as a research paper addressed to an audience of scientific peers.

The study is related to two areas of recent Program activity. The first is an enquiry into SEA evaluation practice, which seeks to document what SEA evaluators really do and the nature of the settings in which they work. Thus, Smith (1980) concludes by suggesting that more attention should be paid to understanding how evaluation functions within organizational and social contexts rather than to questions of causal modeling and experimental design. The second area of Program activity is that concerned with the metaphor series of papers, in which potential new approaches to evaluation are explored as metaphors derived from practices in other fields, e.g., operations research, investigative journalism.

Originally, the study was intended to identify a range of candidate metaphors from the field of management, interpreted broadly, to assess their applicability to SEA evaluation practice, and to derive conclusions

from such application. This proved to be somewhat
too simple an approach due to the complex and varied
nature of SEA evaluation practice, which renders most
metaphors either superficial or of only limited
relevance. Thus, the study's results are derived
mainly from a direct analysis of SEA evaluation
activities, rather than through the indirect route of
metaphor.

The study is based upon two major inputs:

- a review of literature on SEA evaluation
  practice, largely furnished by the Program,
  as a means to developing an initial approach
  to the subject; and

- visits to a number of SEA evaluation units,
  as a means to understanding something of the
  richness and complexity of what actually
  happens in practice.

The main text of the report following has three
sections, dealing with:

- the primary findings of the study (not all
  original by any means);

- discussion of the central theme of the
  study, some current issues, long term trends,
  and types of intervention strategy open to
  the Program; and

- recommendations for further action.

- 2 - 10

There are a number of appendices where some of the material is treated in more detail.

The consultant would like to thank the Program Director, Dr Nick Smith, for his assistance in the course of the study, not only in matters of overall direction, but also in discussion of the planning, interim and draft final reports.

The views expressed in this report, however, remain those of the consultant, and are not necessarily those of the Program.

2. FINDINGS

Major findings of the study are presented below.
They are not all necessarily original; but are
presented here for their significance to the study,
and to provide some introductory basis for the
discussion in the next section. They fall into four
groups:

- evaluation in an SEA context, that is,
  direct findings on the nature of SEA
  evaluation itself;

- management metaphors, that is, those
  activities within organizations which
  appear to be of some relevance to
  understanding SEA evaluation practice;

- management consulting and evaluation,
  that is, the relevance of management
  consulting to evaluation practice; and

- issues of concern, that is, some of the
  issues that surface repeatedly in the
  literature and were confirmed by the site
  visits carried out by the consultant to
  individual SEAs.

2.1 Evaluation in the SEA Context

Findings on the nature of evaluation in the SEA
context, relevant to the present study, are given
overleaf. (They are discussed more fully in Appendix A.)

(a) SEA evaluation is an organizational function,
carried out by an organizational unit located in
an organizational structure.   It therefore differs
from the view implicit in much of the writing on
evaluation, i.e., a single, one time study,
organizationally independent.

(b) The SEA host organization is a distinctive kind
of organization.   It cannot readily be compared
to private business for the obvious public-
private reasons.   Nor does it resemble other
government agencies closely due to the open,
political nature of its activities, and the
atypical emphasis upon outcomes that other agencies
do not share.   It is this emphasis upon outcomes
that gives one of the primary justifications for
evaluation in SEAs, as presently conceived.

(c) SEA evaluation is new.   It dates, in its present
form, from the 1965 Elementary and Secondary
Education Act, which mandated evaluation for the
new federally funded compensatory programs.

(d) SEA evaluation is a diverse activity, varying in
title/scope, size of unit, organizational location,
and practice with regard to contracting work out
or carrying it out inhouse.

(e) SEA evaluation is not a single homogeneous
activity.   Four major types of activity are
practised under the general heading of evaluation
(although only the first of these may be regarded
as evaluation by the purists)

- 5 -   13

- ad hoc, one of a kind research analyses
  of issues and programs, either for
  retrospective program assessment or for
  prospective policy formation;

- information systems development and
  operation, containing data on schools,
  their operations, students, their
  performance, etc.;

- testing, or, more generally, routine
  obtaining of information on a population
  of interest through some instrument of
  enquiry, e.g., a questionnaire; and

- expert and technical assistance to LEAs,
  other SEA staff, legislative committee
  staff, etc. While this may consume a
  great deal of staff time in small bits,
  it is a general function of staff specialist
  groups in all organizations and is not
  specific to evaluation.

(f) SEA evaluation units are not the only units in
    SEAs that carry out evaluations. Some programs
    carry out their own, with or without the help of
    the evaluation unit.

(g) SEAs include other units whose work relates closely
    to, needs to be coordinated with, or overlaps, that
    of the evaluation unit. These include units with
    such names as management information center,
    financial audit, policy and planning, etc.

2.2 Management Metaphors

There are two sets of management metaphors of potential relevance to SEA evaluation. The first relates to the SEA evaluation unit's activities themselves - the unit plays the role of a supplier of services. The second relates to the SEA itself and the unit's clients, which play the role of customers for the unit's services. We shall call these supply metaphors and demand metaphors respectively. (They are discussed more fully in Appendix B.)

The Program's metaphor series of papers has identified a large number of potential supply metaphors. However; it is important to extend consideration to demand metaphors in order to do justice to the fact that evaluation is neither self-sufficient nor an end in itself, but increasingly shaped by the character of demand in its immediate environment.

2.2.1 Supply metaphors

One of the difficulties of identifying metaphors for SEA evaluation is that it encompasses a number of distinct activities, as already described. Since a single metaphor is therefore unsatisfactory, the consultant has sought metaphor(s) for each of the component activities. Although all four activities are represented below for completeness, metaphors for the last two turn out to be somewhat trivial. As can be seen, to distinguish metaphors from the activities themselves, different names are used. Mostly, the metaphors view evaluation as an activity of staff

specialists, not as an activity within the management
mainstream of the organization, that is, of the SEA.

(a1) <u>Operations Research/Management Science (OR/MS)</u>.
This is concerned with ad hoc analyses to support
management decision making, based upon appropriate
theoretical frameworks. At the technique level
it is usually based upon mathematical, financial
and physical science models, rather than upon
behavioral science, but from a management point
of view this is not an important distinction.

(a2) <u>Marketing/Market Research</u>. A second metaphor
for the research analyses element of evaluation
is marketing and market research, which are
traditional business functions. Recently they
have been seen as increasingly relevant to
government operations. In the present case,
for example, they deal with needs assessment,
market segments and segmentation, the success
of products and services as consumed by
customers, in fact, the whole machinery of market
feedback. This is a major part of what evaluation
is all about. They are also much more concerned
with human behavior than OR/MS.

(b) <u>Data Processing/Management Information Systems
(DP/MIS)</u>. This is concerned with the development
and operation of information systems to support
management on a planned or routine basis, both
through on-demand enquiry as well as scheduled
reports, and may include a variety of analytical
tools.

16

- 8 -

(c) <u>Surveys</u>. This is concerned with the periodic
survey of a population of interest, e.g., as in
market research or public opinion polling.

(d) <u>Consulting</u>. This may be internal or external
consulting, and extend from a casual telephone
enquiry to a mini project.

2.2.2 <u>Demand metaphors</u>

The two main metaphors here have to do with the
varieties of managerial decision making on the one
hand, and the necessity of information feedback from
the field on the other. These are, of course,
related, but the metaphors and their associated
literatures have different emphases, and draw
attention to different features of the SEP evaluation
environment.

(a) <u>Decision making categories</u>. For many years the
management community has categorized decision
making into three categories, concerned with:

- operational control;

- management control; and

- strategic planning.

The distinction between these categories is
sometimes blurred, and associated roles and
responsibilities may overlap. A useful indicator
of the level of decision making is the relation to

the budget.   Management control is concerned with
setting and meeting annual goals and budgets,
operational control is concerned with the actual
tasks that the budgets fund, and strategic planning
is concerned with coherent views of the future
extending over several budgetary periods.. These
categories have been used mostly in business
settings.   In government, strategic planning can
often be equated with broad policy formation.

The information required for each category is
different, and consequently different means of
supporting each category have arisen.   In the
present case, operational control is mostly to
be found at the district and school level.
Evaluation will support each type of decision
making differently.

(b) Feedback control.   Feedback control systems are
models for a wide variety of biological and social
processes, as well as underlying the industry of
engineering control from thermostats to missile
guidance.   One of man's distinctive attributes
is his ability to learn from the outcomes of his
actions.   If he cannot observe these outcomes,
because of extraneous factors, or if the rates of
change in the environment occur faster than he can
detect, then feedback is impaired, and his activities
are less likely to be reliably effective.

The same is true for SEAs.   If they cannot track
what is happening in the field in an accurate and
timely manner, then their ability to direct the
state educational enterprise is seriously impaired.

This metaphor is closely related to the preceding
one, but has an emphasis upon information, speed
of response, and noise, that is particularly
appropriate to the function of evaluation in an
SEA.

2.3 Management Consulting and Evaluation

One of the Program's interests is the extent to which
management consulting itself might be a metaphor for
evaluation. As identified earlier, there is clearly
a part of SEA evaluation practice that is a consulting
activity. It therefore shares with management
consulting a common concern with what might be called
consulting hygiene, that is, good client relationships,
timely reporting, etc. The consulting activities of
NWREL also share this concern.

In the view of the consultant, the consulting component
of SEA evaluation is mostly a form of technical
consulting rather than of management consulting. This
is mainly due to the fact that evaluators are not
usually asked questions for which the management
consulting approach is the most appropriate.
Technical consulting is narrower, more focussed,
concerned with the supply side, and "how to" questions.
Management consulting is broader, concerned with
demand and often with "what for" questions. Although
there can be a degree of overlap between the two,
nevertheless, the two types of consulting acquire
different styles, attract different staff, serve
clients differently, and generally operate in
different markets. Management consulting is discussed
further in Appendix C.

2.4 Issues of Concern

SEA evaluation today is characterized by a number of
issues that surface repeatedly, both in the literature,
and also in discussions with individual SEAs. These
issues are often only touched upon partially, but
when all the references are taken together, a picture
emerges of real issues that need to be explored,
understood, and responded to, if the future of SEA
evaluation is to match its early hopes. Three
issues are discussed more fully in the next section,
namely:

- evaluation does not have the impact on
  SEA management that its proponents believe
  it should;

- expectations of what evaluation as a
  technique can deliver are not met (and
  this is sometimes true of expectations
  of the educational process itself); and

- there is some degree of tension between
  methodologists and practitioners.

20

3. DISCUSSION

In this section we discuss a number of topics that we
believe are important to an understanding of SEA
evaluation today and to its future. These include
the three issues mentioned at the end of the last
section, as well as long term trends in SEA evaluation,
and types of intervention strategy open to the Program.

In considering these topics, the consultant has come to
hold the view that the dominant paradigm underlying
SEA evaluation is in the process of changing
significantly. This proposition will be discussed
first, since it is so fundamental that it sheds light
on most of the other issues, as well as giving an
orientation towards the future, and to possible
intervention strategies open to the Program.

3.1 SEA Evaluation: A Paradigm in Transition

Much evaluation literature treats evaluation as
synonomous with large, one of a kind social research
studies, carried out by independent contractors,
utilizing advanced methodologies, highly influential,
and lasting many months or even several years. This
is not characteristic of SEA evaluation.

Let us go back to 1965, and consider both the mission
and the methodologies of the new federal programs in
compensatory education. On the mission side they
were seen as a means of idealistic social change and
not just as education, part of the administration's
Great Society program. The poor and disadvantaged

would be helped in a demonstrable way. This led to
an emphasis on outcomes that other government spending
does not share, e.g., roads, health, defense. On
the methodology side, Robert McNamara was introducing
PPBS, and the climate favored a methodological
rationalism. The climate has since changed. Thus
evaluation acquired early on two set beliefs:
compensatory education will really change outcomes,
and this change in outcomes can be shown by an
appropriate methodology.

The first evaluations were large, one of a kind studies,
performed by independent contractors, from academe as
well as business, who considered themselves major
contributors to policy through their research - they
were the peers of the policymakers.

SEA evaluation, influenced by the mandatory requirements
of programs such as Title I, tended to form itself in
the mold of such early federal activity. (Of course,
some SEAs were too small to do very much in this
direction.) However, the realities of SEA evaluation
are mostly quite different, in terms of organizational
context, size and skills available, timescales, funding
level, and so on. The SEA evaluation unit, in fact,
plays a SEA management support role, and this role will
develop in different ways depending upon the host SEA's
legislative context, size, leadership, organization,
and allocation of responsibility between the variety
of specialist staff units that serve its management,
including evaluation.

The paradigm of a field of activity is the most basic
underlying concept of what it is about. It is an

organizing concept, which brings to mind certain
characteristic questions and approaches. Indirectly,
it directs research and development. It provides a
common basis for integrating the efforts of different
parties. The paradigm underlying SEA evaluation is
in the process of changing from the autonomous conduct
of a big study to the operation of a management support
unit in an organization.

Out of date paradigms lead to a widespread sense of
confusion and frustration in practice. That this is
so can be seen by considering some of the current
issues of the evaluation community, which are discussed
below. However, it should be noted that updating the
paradigm will not by itself solve all of the problems
facing evaluation in SEAs. There will remain difficult
tasks, except that they can now be approached directly
and pragmatically, without the confusing influence of
an out of date paradigm.

3.2 Issues in Evaluation

We now discuss three mainstream issues of the SEA
evaluation community in the light of the above.

(a) Impact. There is no a priori reason why
    evaluations should have the major, much less the
    determining, impact on policymakers that the
    literature so often implies. This notion arises
    out of the original paradigm underlying SEA
    evaluation, in which the social researcher played
    an influential role in the development of federal
    policy towards both programs and evaluations.

Consequently, when writers speak of lack of impact,
they are harking back to a golden age of evaluation
(which still exists, to some extent, in current
major federal evaluations). In the same vein,
they imply a degree of autonomy on the part of
the evaluator-researcher that is appropriate to
that time. Since SEA evaluation is a specialist
staff function, such autonomy is not appropriate
today in this setting. There can be creative
initiative on the part of the unit, but authority
and sponsorship to carry out studies, gather data,
develop systems, etc., should come from the clients
served. This is clear from the OR/MS and DP/MIS
metaphors.

Consideration of the type of decision making that
evaluation supports sheds further light on the
question of impact. Most evaluation impact is
assumed to be on policymaking, which corresponds
to the broad area of strategic planning in business.
This is the area where the inputs range the most
widely, may well be qualitative rather than quant-
itative, approximate rather than exact, pertain to
the future rather than to the past or present, etc.
In short, evaluation input is only one of a number
of inputs, and may well rank lower than others of
a political or financial nature.

The situation is compounded by technical limit-
ations. In many cases, evaluations produce "soft"
data which can be successfully challenged, for
example, in the adversarial setting of legislative
hearings, or even in the courts.

In contrast, routine inputs to program management decision making from a database with which some familiarity has been gained in the past may well produce rapid and immediate action, as, for example, when sales figures drop below some threshold in business.

Impact is also affected by the quality of evaluator-client communication. It appears that although lip service is paid to this, in most cases there is room for considerable improvement. In general, people habitually underestimate the amount of time and effort that must go into communicating (needs, capabilities, results, etc.). Consulting experience brings this home.

(b) Expectations. There is frequently a gap between what evaluation promises, or is perceived to promise, and what it can deliver. This, too, can be traced back to the first utopian days of evaluation, when major studies of national significance were commissioned by the federal government (and as some still are). Since then SEA evaluation tends to have been oversold, both by evaluators wishing to maintain their influence, and by clients wishing to get programs funded or legislation passed, and incorporating promises of evaluation without being completely clear that the results promised are in fact deliverable.

With the change of context from the federal to the state arena, a change of timescale has also taken place. State level clients want results fast. This leads to reluctance on the part of evaluators

to reduce client expectations, and there is
consequently frustration when they are not met.
However, it is likely that certain things just
cannot be found out in a hurry. Compare, for
example, dieting or drug research.. It may take
years to _prove_ results and causal links. There
are inherent limits to what can be achieved and
how fast. These comprise the current state of
the art and should be openly acknowledged: nothing
is gained by pretending that these limits do not
exist.

(c) _Methodologists and practitioners_. Evaluation is
a young discipline and bears all the hallmarks of
this condition. Typically there is an incomplete
separation of theory and practice, as, for example,
also occurred in information science, artificial
intelligence and machine translation. In all of
these, utopian claims were made for the theories
by academics, and their immediate prescriptive
application was urged. This immediate application
was a failure. Recrimination and pessimism
followed.

Some years later the situation has changed.
Academics pursue theoretical questions in one
reward setting. Practitioners pursue practical
questions in another. Both represent cumulative
knowledge bases, linked by examples of problems
flowing from practice to academe, and by results
from theory flowing in the opposite direction.

In the consultant's view, evaluation is just
coming to this point of separation and there is

understandably some reluctance on the part of the
people involved to lose touch with "the other side".
However, it will probably be better for all
concerned when the two groups have distinct
identities and a degree of mutual respect that
sometimes appears to be lacking at the present.
Such a separation represents both a responsibility
and an opportunity for SEA evaluators to develop a
body of professional expertise appropriate to their
setting.

3.3 Trends in SEA Evaluation

To predict the future of SEA evaluation is highly
speculative, but from the perspective developed in
this study possible directions of development can be
suggested.

The role of evaluation in supporting SEA long range,
or strategic, planning will decline since this is
likely to remain a permanently underdeveloped client
area. SEAs suffer such abrupt changes of social
priority and funding levels that any strategic posture
can be obsoleted almost overnight, and consequently
SEA management tends to be pragmatic in approach, not
often given to planning beyond the next budget.

Evaluation will therefore lose one of its earlier
claims to glamor. As the new paradigm becomes more
accepted, it will become clear that evaluation is one
of a number of support functions available to SEA
management. Evaluation will probably remain the
internal SEA authority on field data collection.

This role in understanding and designing feedback systems represents an important contribution to SEA management that is not presently duplicated by other SEA staff units, and consequently may represent a stable long term role for the evaluation unit.

Costs and response time will cause the reduction of major research studies to a minimum, except for those needed by existing or proposed categorical programs. With the advent of block grants there will probably be less of these.

Costs and response time will also favor the growth of the DP/MIS component of evaluation activity at the expense of the research study component. This mirrors the way in which DP/MIS has come to dominate OR/MS in organizational influence and investment. The growth of the DP/MIS component may initially be masked by its greater efficiency. This irony comes about because although it may be doing more and more, once systems are set up and running, costs are small, and even fall along with falling hardware costs. A secondary effect of the growth of operational systems, however, is "maintenance, drag" which affects both software and procedures, and which further reduces resources for major discretionary studies.

The actual organizational role of evaluation will depend upon individual SEA management. Some units will remain independent, others be absorbed by program management, others merged into an expanded DP operation, and yet others take over the department's DP activities.

28

While differences will always remain, not least because
of size, the consultant believes that evaluation
activities will eventually condense into two clusters:
an evaluation systems analyst role working closely
with integrated database systems (sometimes engaging
in major studies or data collection design projects),
and a consulting role, both internally and for LEAs
and legislative clients.

To expand upon the first of these roles, an evaluation
systems analyst may be compared to other applications
systems analysts in data processing, e.g., financial,
retail, engineering systems analysts. The results
of their work is to cause a computer system to support
some real world activity. They know enough about the
activity itself to be able to think in terms of the
system's impact on it, and enough about computing to
understand how to structure the data processing
requirements for effective implementation. They are
concerned with what information goes into the system,
where it comes from, what databases are maintained,
and what reporting or enquiry facilities are needed.

As in all jobs, there will be a range of seniority
and emphasis. Some will work closely with computers,
others with the design of data collection schemes for
specific evaluations. Some will be entry level
positions, and others will be senior technical staff.

Staffing will reflect the growing maturity of the
field. First generation staff, the staff of the first
ten years, typically have a variety of backgrounds.
Second generation staff will have college backgrounds
in evaluation. However, because of the likelihood of

a growing separation between academe and practice; college trained evaluators may take some time to settle in to the practitioners' world.

A more specific problem with staffing is related to the fact that evaluation has the four components mentioned above. The personal skills and temperament required for each are different, ranging from precision and methodological sophistication in the case of research analyses, to the more pragmatic, results-oriented approach of expert assistance. In some cases, this may lead to tension between the quest for technical excellence and the desire to provide effective help promptly. In large SEAs, this may be alleviated by the fact that staff numbers will permit some degree of specialization. In small SEAs, however, a degree of versatility will be required. If the trends outlined above do take place, then these considerations will be less relevant, since the skills and temperament required for systems analysis and consulting are more similar.

3.4 Intervention by the Program

In the light of the above, the question arises as to what the Program can do to bring about constructive developments in the SEA evaluation field. There are at least three different types of strategy, and each leads to a number of recommendations in the next section.

(a) Clarification. Work to clarify the picture proposed here of a new paradigm underlying SEA

evaluation, i:e., that of a management support
function, so that it can provide a realistic
basis for cooperative activities by all parties
concerned.

(b) Supply side help.   Work to build up the
capability of the evaluation unit to offer
services to the SEA, through tools, techniques,
meetings, etc.

(c) Demand side help.   Work to develop an understand-
ing among SEA clients and top management of the
role and value of evaluation in leading and
managing an SEA today.   This may seem somewhat
out of scope, but the concept of demand-pull can
be as important in developing services as the more
traditional supply-push.   The support of the state
superintendent of education will do more for the
role of evaluation in an SEA than the latest
techniques.

4. RECOMMENDATIONS

Recommendations to the Program arising from the study
are given below, divided into three groups, reflecting
the different intervention strategies outlined at the
end of the last section.

4.1 Clarification

(a) Clarify and validate the proposition that the basic
paradigm underlying SEA evaluation is changing,
along the lines outlined above, and work towards
its acceptance by the various parties involved.

(b) Related to this, develop a model of SEA management
and function, differentiated into, say, three size
groups.   This would include policy and program
development, decision making, legislative and
executive relations, budget preparation and key
information flows.   While reasonably comprehensive
in general, the model would focus on developing a
representation of the role of evaluation in SEA
affairs.

(c) Develop from these a number of detailed case
studies reflecting a range of practical
evaluation requirements in realistic environments.
These could serve as the basis for suggesting
research in academe, for sharing approaches between
practitioners, and for discussing information needs
with clients.   They would be more detailed than
some of the case studies appearing in the literature.

4.2 Supply Side Help

(a) Extend the work of the Research on Evaluation
Program into a more detailed review of the scope
of SEA evaluation units and their activities,
trends, job descriptions, inter-unit relations, etc.

(b) Develop a taxonomy of evaluation methodologies
organized by such attributes as response time,
accuracy, cost, respondent burden, etc.   Be
realistic on any limitations.   Identify profile
of desired methodologies, i.e., gaps in the
taxonomy, and commission research to fill the
most important gaps.

(c) Consider a metaphor study to explore the marketing/
market research metaphor further, as a resource
service for SEA evaluation units, rather than as a
means of extending methodology.

(d) Carry out a study to examine the information
systems that SEA units are developing with
attention to capability, cost, hardware,
transferability, etc.

(e) Organize peer group meetings for evaluators on
specific topics to share experience.   These would
be working meetings.   A newsletter, occasional
journal, or information exchange might also be
organized.

4.3 Demand Side Help

(a) Develop detailed life cycle studies of policy and program development in order to identify the potential role of evaluation, as a means to helping clients perceive this role.

(b) Carry out some kind of market research enquiry into clients' needs for evaluation, perhaps using the Delphi technique.

(c) Organize peer group meetings for program managers and top policymakers, even state superintendents (different meetings), addressed by prestigious experts on information and management, e.g., Peter Drucker. These would require careful preparation, choice of location, and personal invitation from the Director of NWREL.

34

APPENDICES

A. Evaluation Practice in SEAs

B. Metaphors from Management

C. Management Consulting

D. SEAs Visited

E. References

A. EVALUATION PRACTICE IN SEAS

This appendix describes evaluation practice in SEAs
as the consultant has come to understand it in the
course of the study. It is based mainly upon
materials provided by NWREL, as well as visits to a
number of SEAs, and will probably contain little new
to those familiar with the subject.

This description is selective, reflecting the special
interests of the study. That descriptions of
evaluation are frequently, perhaps inevitably,
selective has been amply documented by the metaphor
series of reports sponsored by NWREL during the last
few years, e.g., Guba (1979), Page (1979). The
present study is one more example of how the conceptual
framework and analytical tools of the reporter influence
what is perceived as significant in the field, and how
it is described and structured. (What is sought
ideally, of course, is an approach to evaluation which
has substantial, and lasting, explanatory power and
the ability to suggest and support constructive
strategies in the field.

SEAs are organizations intermediate in the public
education hierarchy, between the federal government,
acting through the US Department of Education, and
local districts. They have relationships with each
of these levels, as well as with their own state
legislatures at state level. These relationships
may be summarized as follows:

> - federal relationships: SEAs are responsible
>   to the USDE for the funds flowing into the

state for specific programs, e.g., Title I;

- state relationships: SEAs are responsible
  to the state legislature, often through an
  intermediary state school board, for
  managing the state educational enterprise,
  to the extent that the SEA has an effective
  leadership and/or control function in the
  state; and

- local district relationships: SEAs have an
  oversight role, often on behalf of state
  and federal funding sources, to see that
  funds are spent as intended and education
  is carried out to appropriate standards,
  as well as a technical assistance role and
  a policy or guidance role.

None of these relationships, however, are strictly
executive, in the sense of one level having the power
to control lower levels by direction, but there are
implicit sanctions through funding that can be
strongly influential. It should be noted that
although the federal relationship is highly visible
and has influenced evaluation deeply, only a relatively
small amount of funds for education actually come
from federal sources.

For all practical purposes, evaluation in SEAs in its
present form dates from 1965 and the passing of the
Elementary and Secondary Education Act (ESEA). This
increased significantly the amount of money flowing
into states from federal sources for education, and
brought with it more than corresponding requirements

for, evaluation and accountability or monitoring.
On the one hand, the worth of the federal programs
had to be established in relation to the national
policy arena, and on the other hand, the money had
to be shown to be spent appropriately to the intent
of the program planners.   In the language of
evaluation, outcomes and process were both to be
evaluated.

It should be noted that 1965 was also the year in
which Robert McNamara's Programming, Planning and
Budgeting System (PPBS), originally instituted in the
Department of Defense, was promoted more widely as an
aid to effective public management in general.
Evaluation, as originally conceived, and PPBS shared
the same utopian rationalist expectations, which many
now believe to have been inappropriate.   The mood of
the times has changed. .

Most of the states now have evaluation units or
departments with specific staff designated as
evaluators.   However, there is a great difference
between states in such matters as:

- the title of the unit - some combination
  of research, testing, evaluation, planning,
  accountability may be used;

- the size of the unit, which may range from
  0 to 70+; and

- contracting practice, ranging from mostly
  inside to mostly outside.
  (Caulley and Smith, 1978)

Perhaps size is the most important factor relating
to competence, and it is not surprising that this
should vary considerably, from SEAs that are
reluctantly forced by the federal requirements to
some minimal level of evaluation, to those that pride
themselves on greater sophistication than the federal
guidelines.   This range of skill has been described
in Boruch and Cordray (1980, pp3-13,14) under the
headings of exemplary, compliance-oriented and
emergent SEAs.

Evaluation units perform services for a number of
different clients, for a variety of purposes:

  - federal government: these evaluations are
    required by law, and are often specified in
    some detail.   They are intended to provide
    the federal government with information on
    the funded programs from a variety of points
    of view, e.g., descriptive, summative,
    outcomes/impacts;

  - state legislature: may direct evaluations
    to be performed both to inform policy
    making activities and to monitor new
    programs, sometimes going direct to SEA
    evaluation unit staff, without going through
    the SEA administration;

  - SEA top management: both in response to
    external requirements and in response to
    its own autonomous development of policies
    and initiatives, requires the support of
    the evaluation unit in providing specific
    information and analyses;

- SEA program managers: these require
  evaluation assistance to discharge their
  program management responsibilities, both
  within their own offices, and in order to
  show funding agencies (both state and
  federal) how the money is spent; and

- local districts: these require technical
  assistance in carrying out local evaluations,
  both for their own purposes, as well as to
  meet requirements arising from state of
  or federal funding.

Although these appear to be neat and clean distinctions,
visits to actual SEAs showed the consultant that SEAs
are complex organizations without strong organizational
boundaries, and evaluation units are caught up in a
variety of client relationships which can exert
extremely taxing pressure on evaluation staff.

Turning to what evaluators actually do, evaluation is
sometimes treated as a monolithic, indivisible activity.
This is not the case in practice, and a number of
distinct activities are performed in providing the
services above. This fact makes a single definition
of evaluation, such as many writers propose, a mis-
leading goal, since such definitions tend to reflect
only a subset of the field of evaluation practice.
These activities include:

- research analyses: such as special analyses,
  enquiries, studies, which entail some study
  design, data collection, analysis and
  reporting. These are mainly one of a kind

projects conducted on an ad hoc basis and
serving either some high level prospective
activity, e.g., policy and planning, or
some previously mandated requirement for
retrospective evaluation;

- information system activities: such as
  the development of systems giving information
  on school operations, test scores, etc.,
  that can be used to generate periodic
  reports, as well as respond to on-demand
  questions;

- testing: the actual management of a testing
  program, from requirements to actual results;
  and

- expert assistance: for example, advice to
  local districts on testing, testimony to a
  legislative committee, inputs to SEA policy
  development, etc.

As in the case of types of client above, these may be
somewhat too exact to be used without some qualification.
Nevertheless, they do seem to span the range of activ-
ities that evaluation units perform, and are used as
indicators of that range in the course of the study.
In practice, as might be expected, there is some
overlap between them. For example, testing may
result in data being fed into an information system,
or expert assistance may involve advice on a research
analysis for an SEA colleague.

In the evaluation literature a number of issues are
currently being discussed.   These include:

- whether the dominant realities of evaluation
  are political or technical.   In the former
  point of view, evaluation is intimately part
  of the political process and its success
  will be partly political.   In the latter
  case, methodological and communications
  improvements will lead to success.   These
  contrasting viewpoints are illustrated by
  two recent national studies, Rand (Pincus,
  1980) representing the political reality
  and Northwestern University (Boruch and
  Cordray, 1980) the technical.

- closely related to these is the notion of
  some methodological grail, which, when
  discovered, will lead to an educational
  utopia.   This grail is generally assumed
  to be some form of advanced social science
  research methodology.

- there is also concern over the impact of
  evaluation studies on policy and decision
  makers, a feeling that it is too low, and
  should be higher.   This overlooks the
  threat of change that serious evaluation
  poses to program managers and policy
  makers alike.   These people often have
  enough difficulty getting agreement on
  political grounds, and do not particularly
  want to have their investment in compromise
  and negotiation put at risk by consideration

of new information from retrospective
evaluations.   These are consequently
treated as pro forma reports, and should
not be expected to have impact.

In reviewing the literature provided by NWREL the
consultant has some additional observations to make:

- scope of evaluation: there appears to be
  some confusion over the scope of evaluation
  as the term is used in the literature, e.g.,
  does it really include testing or information
  systems, or are these lesser activities
  reluctantly carried out by evaluators, whose
  primary calling is research;

- autonomy/organizational role: what can be
  inferred from the writings of some evaluators,
  especially in relation to limited impact on
  policy, is that they conceive themselves as
  having a certain professional autonomy to do
  research on their own initiative in order to
  provide input and guidance (even direction)
  to policy makers;

- rates of change: a factor additional to the
  political nature of the evaluator's environ-
  ment is the rapid rate of change in policies,
  program goals, guidelines, social expect-
  ations, etc., not to mention staff and
  organizational change.   These changes
  occur with a rapidity that approaches and
  sometimes is faster than the timescale
  required to carry out rigorous evaluations.

In the limit, this timescale factor can make evaluations almost irrelevant to decisions; and

- <u>robustness of evaluations</u>: as these factors are taken into account, together with the presence of a good deal of environmental noise, that is, extraneous factors of various kinds, and the often adversarial nature of policy making discussions, it is a matter of question whether or not it is in fact possible to do robust evaluations today, that is, evaluations whose results cannot be successfully challenged. If this is the case, it should give evaluators some serious concern.

B. METAPHORS FROM MANAGEMENT

The aim of this section is to describe and discuss
some metaphors that may prove helpful in understanding
present evaluation practice, and in supporting
constructive developments in the future.   These
metaphors relate to evaluation at two levels:

- _supply metaphors_, relating to evaluation as
  a supplier of services; and

- _demand metaphors_, relating to the SEA and
  evaluation unit's clients, who play the
  role of customers for the unit's services.

In considering metaphors from management, the very
real question of differences between business and
government arises.   One treatment of management that
tries to provide a single perspective for viewing
these two contexts is Buchele, (1977).   This is also
of interest because of his strong advocacy of systems
management, with its emphasis on feedback and infor-
mation, which is quite relevant to the SEA context.
However, SEA policy making is quite unlike business
strategic planning, due, to the rapid change possible,
in funding and goals imposed by external forces, e.g.,
the state legislature.

Two readable writers on the systems approach are
Churchman (1968), who gives a balanced treatment of
the systems approach that is aware of the social
complexity of the real world, and Boguslaw (1965),
who criticizes some of the utopian tendencies of the
early systems analysis movement.

B.1 Supply metaphors

As remarked in the preceding section, evaluation
is not a monolithic activity, but has a number of
distinct elements in practice (even if "pure"
evaluation is only concerned with the first), namely:

- research analyses;

- information systems;

- testing; and

- expert or technical assistance, which can
  be associated with any of these.

B.1.1 Metaphors for research analyses

Research analyses are carried out by specialist units
of various kinds within organizations. These include
a whole family of evaluation methodologies (Swanson,
1975), some of which can be carried out externally,
e.g., by a consultant for a client. In fact, a
number of these are the kinds of analyses carried
out by internal consulting units in organizations,
e.g., operations research, management audit, while
others are associated with budgeting activities, e.g.,
cost-benefit analysis and programming, planning and
budgeting (PPBS), and yet others with product
optimizing activities, e.g., value analysis/engineering.
Of these, operations research has already been treated
in the metaphor series of reports (Page, 1979), and
value analysis/engineering is reminiscent of the

special thrust of formative evaluation.

Research analysts are also employed by public organ-
izations in a variety of other fields, e.g., budget
analysis, policy analysis. Another example of a
research analyst is the library or information
analyst. Market research analysts may conduct
original research analyses or routine surveys.
Investigative journalists also function in a similar
way without the organizational background, see, for
example, the metaphor series of reports (Guba, 1979).
These research analysts are characterized by
specialized training and skills, theirs is technical
work, not managerial or administrative. They carry
out ad hoc studies rather than routine activities,
and are concerned with questions of methodology.

The most suggestive metaphors are those of operations
research/management science (OR/MS) and marketing/
market research. These are discussed at greater
length below.

(a). Discussion of the OR/MS metaphor. · OR/MS, as well
as more general forms of consulting, have special
requirements in performing services for management.
These include the negotiation of workable spec-
ifications for tasks, good communications (both in
mutually consistent views of what is being asked
for by management as well as in clear, timely
reporting), and mutual education. Management
should be kept abreast of the kinds of realistic
methodological capabilities that OR/MS has to
deliver results with, and OR/MS should understand
what the organization's goals are and how they

are being pursued. When it commits
resources to tasks, OR/MS is responsive to
management's sponsoring authorization. OR/MS
staff should not indulge in private projects,
nor believe that their technical expertise in
some way makes them superior to general management.
Nevertheless, an explicit awareness of this support
service role need not rob OR/MS staff of high
professional aspirations and conduct.

From what the consultant has been able to learn
of evaluation, it seems that good evaluation
should share many of the characteristics of good
OR/MS practice in the area of dealing with clients.
In general, seeking good communication with clients
seems to be acknowledged as a goal, although one
that is not always achieved in practice. However,
there is some (implicit) reference in the liter-
ature to the professional autonomy of the evaluator
in initiating studies, which violates the essential
support role. It is not surprising that such
studies lack impact.

It takes time, both staff time and elapsed time,
to carry out analyses. If requirements or
activities are changing rapidly, then long
running projects cannot be undertaken. If the
environment is noisy, that is, there are many
extraneous factors, then methodologies that rely
upon exact data will not work. The OR/MS
professional responds to limitations inherent
in the context of the study by selecting an
economic and appropriate methodology capable of
giving results on time. If none appear to be

available, he advises management, but does not
necessarily expect the organization to stop
functioning until one is developed, or for long
enough for one to be applied.

In some evaluation literature, in contrast, there
appears to be the assumption that methodology comes
first, and that policymakers should only propose
programs that existing methodology can analyse.
This is not a position that SEA evaluators can
afford, even if appropriate to academe.   On the
other hand, there is some truth to this position,
a variant of "do not promise what you cannot
deliver".   In the case of the SEA program manager
preparing his budget, the caution is "do not
promise what your evaluation unit cannot document".

In the same direction, a mature field of applied
professional skills, such as OR/MS, knows that
the gulf slowly widens between academe and the real
world.   The academic view of the subject is pure,
exact, permitting sophisticated methodologies in
simplified and abstracted settings.   The real
world is pragmatic, oriented towards useful results
rather than theoretical purtiy, and constrained by
time and cost.   In the same way, professional
evaluators do not do the same kind of work as
academic evaluators, and should not regard the
writings of the latter as prescriptive.   On the
other hand, the latter should not feel slighted
by this, since each has their own legitimate field
of endeavor.   Ideally, feelings of mutual respect
should obtain between academics and practitioners,
but, from experience in other fields, this is easier
said than realised.

(b) <u>Discussion of the marketing/market research metaphor</u>.
In contrast to OR/MS, marketing functions are more
closely integrated into an organization's mainstream
and central planning. Marketing staff work
closely with production and finance to assure
that the right amount of the right products are
produced at the right time and the right cost.
They are concerned with the consumption of existing
services as well as planning for new ones through
needs assessment. They carry out comparative
studies to find out ways of improving products
and services. Their situation differs from
evaluation, however, in the existence of revenue
in the business environment as a concrete measure
of worth.

The concept of market segementation seems a key
one for evaluators and educators generally.
How does one decide to split up student populations
into distinct groups to be served by different
programs? How does program effectiveness depend
upon the number of segments?

Looking at the central role of marketing, evaluators
may be able to learn from the way in which marketing
inputs are widely used in organizational decision
making, and how they are able to deliver effective
inputs from the same messy environment, within the
disciplined schedule of the organizational planning
cycle.

50

B.1.2 Metaphors for information systems

Data processing and management information systems
(DP/MIS) are now integrally part of any modern
organization. DP/MIS is, like educational evaluation,
a relatively young activity, with first practical
applications of computers dating from the mid 1950s,
and the MIS movement from the mid 1960s. MIS is
characterized by the ability to develop, maintain
and exploit a database of information relevant to
the organization's activities and responsibilities.
Over the years the emphasis in DP has moved away from
ad hoc projects with short life to the systematic
development of long running applications, based upon
a growing database of machine readable information.

Mathis (1976), Rasp (1980) and Cronin (1980) all
describe how evaluators deploy such systems, within
their own organizational contexts.

DP capacity planning is part of any serious planning
at the organizational level, since it is now clear
to all concerned that all applications take machine
and staff resources, and future applications must
be planned for. Of course, if the organization's
requirements change faster than the typical applications
development cycle, then DP can no longer give current
support to the organization and there will be a degree
of waste in the applications that become obsolete
before the investment in their development can be
recovered.

DP/MIS departments have some role in educating their
organization as to what the current technology can

achieve, i.e., what is possible.   However, they do
not indulge in independent research (except in the
very largest and then budgeted in advance), but
develop specific applications for specific clients
after careful cost-justification;

Because DP/MIS is a new organizational function, there
is an awareness that there is a degree of development
associated with·it as a field, as well as within any
particular organization.   This has been described by
a stage model (Strassman, 1976) in which four distinct
stages are identified:

- <u>initiation</u>, characterized by interest and
  excitement in the novelty of DP;

- <u>expansion</u>, characterized by demonstration
  of applications in a variety of organizational
  settings;

- <u>control</u>, characterized by a desire to manage
  a variety of DP projects more efficiently,
  a narrow, bottom-line approach; and

- <u>maturity</u>, characterized by the management
  of DP as an effective contributor to the
  organization's purposes, a key resource
  component for long range planning, and a
  broader bottom-line approach.

The longer DP/MIS has been active in an organization,
the greater will be the volume of applications software
and databases.   In turn, the maintenance requirements
will rise, competing for scarce DP/MIS staff time that

- 44 -

can be applied to new developments. Thus, long term
DP/MIS activities are accompanied by long term
maintenance "drag", a phenomenon that is receiving
considerable attention in the computing world.

More recently, as an outgrowth from the MIS movement,
there has been the desire to develop better aids to
decision-making than MIS typically provide. This
effort is known by the name "decision support systems",
and is aimed at taking what is known in advance about
the structure of the decision and developing computer
models working with relevant databases to provide a
more intimate and powerful support than is possible
with a general purpose database and an enquiry
language. This represents an interesting move to
narrow the gap between the ad hoc and the routine
information system, and clearly may represent an idea
of relevance to evaluation. Decision support systems
are well described by McCosh and Scott Morton (1978).

Discussion of the DP/MIS metaphor. DP/MIS is only a
little older than SEA evaluation. Its role within
organizations during this time has changed a great
deal, and this change has been described by the
stage model above. It seems likely that SEA evaluation,
at least that element that is concerned with information
systems, will pass through some similar progression of
stages. There is an inherent, experience-based
learning curve, compounded of both organizational
and technical factors, which probably cannot be
speeded up. This appreciation should make us more
patient when considering the effectiveness of evaluation
within SEAs. Time will bring increased maturity.
Something of this may be already seen in the evolution

- 45 -

of the scope of evaluation units in some SEAs.

A major issue in many organizations has been the
relationship between OR/MS and DP/MIS.   In the early
days, OR/MS tended to have the lead, because they
were there first, and had the technical skills to
operate the early computers.   Progressively, however,
DP/MIS has become the dominant service in most
organizations. Reasons for this include such things
as the highly visible investment in hardware, the
desire to use prepared information to answer urgent
questions rather than ad hoc studies, and the initial
capture of data in machine readable form from the
operational environment, rather than having to pay
extra for its later entry into the management system.

For reasons such as these, it seems likely to the
consultant that the information systems component of
the evaluation function will come to dominate the
analysis and survey components. This may lead to a
forced reappraisal of individuals' career goals and
their relative importance to the SEA.

The leadership of the DP/MIS function in organizations
is another issue of interest by analogy to evaluation.
There is some controversy over whether or not this
should be a person of general management orientation,
or a technical expert.   Technical expertise is
increasingly less important. More important is an
understanding of the goals of the organization and
the ability to deliver services to support management,
in reaching these goals.   This may suggest that the
leadership of SEA evaluation units (larger ones, that
is, since smaller ones may need all the technical

expertise that they can get) will pass into the hands of general managers, especially in those cases where information systems become dominant as foreseen above. One factor which may work against this move is the fact that the top echelon of management in an SEA is either elected or appointed, and may not have skills of a general management nature adequate to taking over and running an evalua██n unit without its own traditional technical leadership.

DP/MIS is also committed to systematic capability or resource planning, integrated with other planning activities of the organization, both in the short and long term. This grows out of the increasingly indispensable role that it plays in organizational effectiveness. It seems likely that some components of SEA evaluation will come to practise similar planning as they become more intimately part of effective SEA operations, and are perceived as such.

B.1.3 Metaphors for testing

A number of different disciplines engage in surveys. These include public opinion polling specialists, and market researchers. These correspond roughly to the testing activity that is often part of evaluation.

These activities are characterized by a degree of routine application of some enquiry tool, e.g., a questionnaire, statistical questions of sampling size, population homogeneity, reliability, confidence limits, and so on. Because of the degree of routine, results may be accorded a significant role in decision

making in advance, since there is prior experience of
the form of the results and their past utility.

Discussion of the survey metaphor. The main point
-that arises from this metaphor is that since surveys
are not exciting, state of the art activities in
general, but rather routine in nature, a premium will
be placed on value for money, i.e., efficiency.

B.1.4 Metaphor for expert assistance

The expert or technical assistance component of SEA
evaluation activity is a form of consulting. This
can range over a wide variety of subjects and levels
of effort. Consulting is therefore a metaphor for
expert assistance, but the two are almost synonomous,
and so this metaphor does not really extend our under-
standing of evaluation. It will not be discussed
further.

B.2 Demand Metaphors

The two main metaphors here have to do with the
variety of managerial decision making on the one hand,
and the necessity of information feedback from the
field on the other. Although closely related, they
have different emphases.

56

B.2.1 Decision making categories

Anthony (1965) introduced the classic distinction of
management decisions as falling into the categories of:

- operational control;

- management control; and

- strategic planning.

Lucas et al (1974) discuss this framework along with
related approaches of Simon and Scott Morton.   Simon
uses the terms intelligence, design, choice to describe
decision making at the broadest level, and Scott Morton
emphasizes the degree of structure in a decision,
speaking of structured, semi-structured and unstructured
decisions.   The information systems that support each
of these classes of decision will be different, and
this is reflected in the literature and practice of
MIS.

In actual fact, these three decision categories are
not absolutely distinct, but tend to form a continuum
in practice.   Anthony describes a number of information
characteristics for each of the categories.   These
are shown overleaf.

Evaluation, broadly considered, may support each of
these types of decision making, but not in the same
way.   Further, because of the SEA setting, strategic
planning is not such a clearly defined activity as it is
in business, because of the difficulty of developing
and applying coherent strategy in an extremely

Information requirements for decisions in:

Operational control	Management control	Strategic planning
1. very detailed data	moderately detailed data	aggregate data
2. related to a specific task	related to achievement of organization's	related to establishing broad policies
3. frequently reported	regularly reported	infrequently reported
4. historical data	historical and predictive data	predictive data
5. internally generated	mostly internally generated	externally generated
6. very accurate	accurate within decision bounds	accurate in magnitude only
7. repetitive	exception reporting	unique to problem being considered
8. often non-financial	mainly financial	often non-financial

The Anthony Framework

changeable environment. Additionally, operational
control is more likely to be practised at the local
district level than in the SEA. Failure to
distinguish between these different requirements is
likely to lead to confusion.

Kling (1980) and Keen (1981) both treat the organ-
izational context of information systems in a manner —
that extends Anthony's rather general categories
to more realistic political environments, such as are
found in SEAs. Research shows, for example, that
much human information processing tends to be simple,
experiential and non-analytic, yet fairly effective.
The information dependencies of SEA management are real,
but not so direct or determining as is sometimes
suggested.

B.2.2 Feedback Control

A simple abstract model of information and decision
making has been proposed by Yovits and Ernst (1968),
and discussed in a number of papers since, e.g.,
Yovits et al (1981a,b). This control systems model
is highly abstract, but it does remind us of important
features of general systems, including feedback, time
constants and noise, all of which are relevant to
considering the systems character of an SEA, within
which the evaluation unit functions. These features
may go some way to explaining some of the inherent
limitations on what evaluation is able to achieve in a
practical situation. For example, if feedback is not
employed, or if it is employed but time constants or
noise are unfavorable, then the organization cannot

follow an effective course of action with any
confidence.

Another notion presented by this model is that of a
transducer, a means of converting information from
one medium into another.   In a sense, evaluation
acts as a transducer, or sensor, in extracting
information in useful structure from the unstructured
social environment in which education actually takes
place.

60

C. MANAGEMENT CONSULTING

One of the aims of the Program in sponsoring the present
study is to use it as an example of a study carried out
from a management consulting perspective. This is in
addition to the substantive results of the study.
Therefore it seems useful in this appendix to give a
brief description of management consulting, together
with some sources for further reading. This may be
of interest both to NWREL, which does a significant
amount of consulting itself, as well as to evaluators,
who, as we have seen, also play a consulting role in
their own environment.

Generally, management consulting is concerned with
specific assistance given to an organization by an
outside consultant and characterized by extensive
on-site activity on the part of the consultant,
although the term is often used more broadly. One
definition, for example, of classic management
consulting reads:

> "The service provided by an independent and
> qualified person or persons in identifying
> and investigating problems concerned with
> policy, organization, procedures and methods;
> recommending appropriate action and helping
> to implement these recommendations."

Reasons for employing a consultant include: their
special skills, the need for intensive professional
help on a temporary basis, or the need for an
intelligent and impartial outside viewpoint.

The present study, therefore, is not a typical
management review, since it is indirect, based in part
upon literature, with visits to SEAs for field input
and checking of hypotheses.   The study has been
sponsored by NWREL, but is concerned with SEAs, who do
not work for NWREL.   This naturally limits their
participation and interest in the study.   If the SEAs
themselves had commissioned the study, at a senior
level, then more time would have been spent with them,
and the results and recommendations would be directed
to them.   However, this would not have been the usual
form of a multi-client study, which more usually deals
with market research, technology assessment, industry
profiles, etc.

As consulting has developed, some firms have come to
specialize, or work in specialized settings, e.g.,
business or government, a particular sector such as
forestry or education, a particular management
function such as marketing or personnel.   Larger
consulting firms have diversified extensively to
include such services as traditional management
consulting, technical services in economics or data
processing, executive recruitment, training and
education, technology research, applied research
studies, policy analysis, and so on.   In addition,
major CPA firms have also diversified into management
consulting, as have also other service companies, such
as, for example, large computer service firms.

Thus management consulting has a degree of breadth and
diversity in practice which is not always represented
by works on the subject.   Many of the hallmarks of
good management consulting practice are shared by

other specialized consulting services. The most
important of these is the relationship between client
and consultant. This includes the way in which the
the consulting task is specified or negotiated and
the consulting firm (or individual) selected, the
expectations on each side, the maintenance and
development of the relationship through the period of
service, the focus upon mutually agreed results, the
timely delivery of agreed products (considered in the
widest sense), and so on.

Within the framework established by a client paying
real money for a real service, i.e., the client
wants certain results or outcomes for his organization,
features of management consulting that are distinctive
include:

- flexibility of means to achieve ends, i.e.,
  the consultant is pragmatic and may select
  one of a number of different approaches as
  appropriate. There is no a priori limit-
  ation to a specific technique, as in the
  case of hiring a specialized consulting
  service;

- related to this, the consultant may bring
  in additional technical expertise as
  required (with the prior agreement of the
  client), serving as a kind of knowledge
  broker between such resources and the
  client;

- the emphasis upon benefits for the client,
  and with the objectives of the client
  organization (just what is it really trying
  to do?), rather than solely at the unit
  level, i.e., a holistic approach to the
  client;

- a diagnostic phase of learning, listening,
  thinking, prior to developing specific
  approaches to achieving solutions;

- related to this, the awareness that the
  expressed needs of the client may not be
  the real needs, and the ability to uncover
  the latter and work towards them within the
  framework of trust already established;

- a desire to produce the maximal beneficial
  change for the client, which leads to the
  advocacy of recommendations that are
  realistic and achievable rather than idealized
  and unachievable;

- the communication of results through reports,
  discussion, presentation, with simplicity,
  clarity, and emphasis upon key elements,
  with the sole criterion: is it clear to the
  client:

- a framework which explicitly includes the
  period after service, the client's world
  downstream of the final report, when the
  changes recommended are likely to come about
  and have their effects; and

- a willingness to continue service into the
  implementation period should the client
  wish.

Although these are some of the features that character-
ize the central practice of management .consulting,
i.e., advising on, or working with a client to produce,
beneficial change, most can also be identified in
the other types of services that management consulting
firms undertake for clients, such as training or
applied research.

Despite the importance of management consulting in
supporting the practice of business and government,
and often the leading edge of these, it does not have
a correspondingly high level of literature. Much of
this is due to the confidential nature of consulting
assignments, so that they do not contribute to the
open literature. On the other hand, part may also
be due to the fact that academic faculty are often in
competition as consultants in the outside world, and
are somewhat scornful of the pragmatic and eclectic
approaches employed in the real world. Also,
the client-centered approach of the professional
consultant is at variance with the theoretical approach
of the faculty member, whose primary allegiance is
often, through necessity, to the advancement of
knowledge through publication, rather than to client
well-being.

Having said this, some useful sources include:

- Bell, Chip R, and Nadler, Leonard (eds)
  The Client-Consultant Handbook, Houston, TX,
  Gulf Publishing Co., 1979.

- Fuchs, Jerome H.   Making the Most of
  Management Consulting Services, New York,
  AMACOM (Division of American Management
  Associations), 1975.

- Hunt, Alfred.   The Management Consultant,
  New York, Ronald Press, 1977.

- Kubr, M (ed).   Management Consulting: a
  Guide to the Profession, Geneva, International
  Labour Office, 1977.

Trade associations and professional societies in the
US include:

- Association of Consulting Management
  Engineers, Inc. (ACME)

- Association of Management Consultants

- Institute of Management Consultants

- Society of Professional Management
  Consultants, Inc.

66

D. SEAS VISITED

Visits were carried out to three SEAs, coordinated
by the directors of the evaluation units:

California	Dr Alex Law
Montana	Dr William Connett
Washington	Dr Alfred Rasp, Jr

The consultant is grateful to these directors and
their SEA colleagues for making their time so
generously available for interview and discussion.

67

E. REFERENCES

Anthony, Robert.   "Planning and Control Systems: a
Framework for Analysis", Boston, Graduate School of
Business Administration, Harvard University, 1965.

Boguslaw, R.   The New Utopians: a Study of System
Design and Social Change, Englewood Cliffs, N.J.,
Prentice-Hall, 1965.

Boruch, Robert F., and Cordray, David S.   "An
Appraisal of Educational Program Evaluations: Federal,
State and Local Agencies", Evanston, IL, Northwestern
University, 1980, a report prepared for the US Depart-
ment of Education.

Boyd, H.W., and Massey, W.F.   Marketing Management,
New York, Harcourt Brace Jovanovich, 1972.

Buchele, Robert B.   The Management of Business and
Public Organizations, New York, McGraw-Hill, 1977.

Caulley, Darrel N., and Smith, Nick L.   "Field
Assessment Survey", Portland, OR, NWREL, 1978.

Churchman, C. West.   The Systems Approach, New York,
Dell Publishing Co, 1968.

Cronin, Joseph M.   "Educational Research and Evaluation
in the 80's: a state perspective", Springfield, IL,
Illinois State Board of Education, paper presented at
the Annual Meeting of the American Educational Research
Association (64th, Boston, MA, April 7-11, 1980).
ED 193 276

Guba, E. G. "Investigative Reporting", paper prepared for NWREL, 1979.

Keen, Peter G. W. "Information Systems and Organizational Change", Communications of the ACM, Vol 24, No 1 (1981) pp24-33.

Kling, Rob. "Social Analyses of Computing: Theoretical Perspectives in Recent Empirical Research", Computing Surveys, Vol 12, No 1 (1980); pp 61-110.

Lucas, Henry C., Jr., Clowes, Kenneth W., and Kaplan, Robert B. "Frameworks for Information Systems", Infor, Vol 12, No 3 (1974) pp 245-260.

Mathis, William J. "Using Product and Process Results For Research and Planning: the New Jersey Model", paper presented at the Annual Meeting of the American Educational Research Association (60th, San Francisco, CA, April 19-23, 1976). ED 124 579

McCosh, Andrew M., and Scott Morton, Michael S. Management Decision Support Systems, London, MacMillan Press, 1978.

Page, E. B. "Educational Evaluation through Operations Research", paper prepared for NWREL, 1979.

Pincus, John (ed). "Educational Evaluation in the Public Policy Setting", Santa Monica, CA, Rand, 1980.

Rasp, Alfred, Jr. "Evaluation and Decision Making in Washington State", paper prepared for NWREL, 1980.

Smith, Nick L.   "The Context of Evaluation Practice
in State Departments of Education", Portland, OR,
NWREL, 1980.

Strassman, Paul A.   "Stages of Growth", Datamation,
October 1976.

Swanson, Rowena W.   "Design and Evaluation of
Information Systems", in Annual Review of Information
Science and Technology, Vol 10, Washington, DC,
American Society for Information Science, 1975,
pp 43-101.

Yovits, M. C., and Ernst, R. L.   "Generalized
Information Systems: Some Consequences for Information
Transfer", Columbus, OH, Department of Computer Science,
Ohio State University, 1968.

Yovits, M. C., Foulk, C. R., and Rose, L. L.
"Information Flow and Analysis: Theory, Simulation and
Experiments. I. Basic Theoretical and Conceptual
Development", Journal of the American Society for
Information Science, Vol 32, No 3 (1981) pp 187-202.

Yovits, M. C., Foulk, C. R., and Rose, L. L.
"Information Flow and Analysis: Theory, Simulation and
Experiments. II. Simulation, Examples and Results",
Journal of the American Society for Information
Science, Vol 32, No 3 (1981) pp 203-210.

70