DOCUMENT RESUME

ED 251 514 TM 850 021

AUTHOR Alkin, Marvin C.; And Others
TITLE Annual Report. Technical Reports. Evaluation Productivity Project.
INSTITUTION California Univ., Los Angeles. Center for the Study of Evaluation.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Nov 84
GRANT NIE-G-84-0112-P2
NOTE 145p.; For the individual papers included in this report, see TM 850 022-024.
PUB TYPE Collected Works - General (020) -- Viewpoints (120) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC06 Plus Postage.
DESCRIPTORS Administrator Characteristics; *Administrator Role; *Cost Effectiveness; Cost Estimates; Decision Making; Evaluation Methods; *Evaluation Utilization; Evaluators; Experimenter Characteristics; Institutional Characteristics; Leadership Qualities; Predictor Variables; *Program Evaluation; Resource Allocation
IDENTIFIERS *Center for the Study of Evaluation CA; Evaluation Costs; Evaluation Reports; Indirect Costs

ABSTRACT

After outlining the 1984 activities and results of the Center for the Study of Evaluation's (CSE's) Evaluation Productivity Project, this monograph presents three reports. The first, "The Administrator's Role in Evaluation Use," by James Burry, Marvin C. Alkin, and Joan A. Ruskus, describes the factors influencing an evaluation's use potential and outlines the steps the program administrator must take to increase the evaluation's use. The second paper, "Factors Common to High-Utilization Evaluations," by Ruskus and Alkin, identifies six of these factors: level of evaluator effort, evaluator's leadership behavior, user involvement, evaluation reporting, evaluator involvement in implementing recommendations, and user commitment to use. The final paper, "Reflections on Evaluation Costs: Direct and Indirect," by Alkin and Ruskus, describes how to identify and assign values to both direct and indirect cost components, and how costs should be considered in relation to benefits accrued from an evaluation. (BW)

DELIVERABLE - NOVEMBER 1984

EVALUATION PRODUCTIVITY PROJECT

Annual Report

Technical Reports

Marvin C. Alkin
Project Director

Grant Number

NIE-G-84-0112, P2

Center for the Study of Evaluation

Graduate School of Education

University of California, Los Angeles

The project presented or reported herein was supported pursuant to a grant from the National Institute of Education, Department of Education. However, the opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement by the National Institute of Education should be inferred.

TABLE OF CONTENTS

ANNUAL REPORT: EVALUATION PRODUCTIVITY PROJECT, 1984

This report outlines the results of the Evaluation Productivity Project's 1984 scope of work and sets the context for the deliverables we produced during this period. These deliverables reflect the following project activities:

- o Recommendations for Improving Evaluation Productivity
- o Organizing for Evaluation Use
- o Instances of High Evaluation Use
- o Costs of Evaluation

In addition, the report describes progress on our two 1984 planning activities which will lead to project deliverables in 1985:

- o Demonstration/Documentation of the Use Process
- o Preliminary Planning for Synthesis of the Knowledge on Evaluation Use

Recommendations for Improving Evaluation Productivity

As we reflected on our examination of the evaluation use process and the factors that influence use, we saw that the educational administrator's role in promoting use was extremely important: In addition to context-organizational factors requiring administrator attention, many human and evaluation factors that affect the use process will require administrative attention and influence to promote an evaluation's program-level use. In the past, leadership in these last two factor domains had not traditionally been associated with administrative responsibility.

We began to discuss this finding with other evaluation-use researchers and with researchers and practitioners in educational administration. The consensus was that our perception of the administrator's role in evaluation

use was accurate and important and carried implications that administration researchers and practitioners need to understand and act upon in their work.

Our next task was to begin informing the field about how the administrator can contribute to evaluation use and to formulate tactics administrators can apply as they provide leadership in the use process. We initially presented our thoughts on the issue at the annual meeting of the Evaluation Network/Evaluation Research Society in 1983. We made a more detailed presentation with concrete procedural suggestions to Division A, Administration, at the annual meeting of the American Educational Research Association in 1984. We produced a CSE Report on the topic.

The positive reception our work received at meetings such as those mentioned above, as well as administrator interest in understanding how to guide and promote an evaluation's use, and therefore its productivity, prompted us to seek publication in a manner that would impact the educational administrator. The outcome of this effort is that the paper we presented to AERA's Division A in 1984 -- The Administrator's Role in Evaluation Use (CSE Report No. 225) -- will appear in 1985 in a special edition of Studies in Educational Evaluation devoted to "Evaluation as a Management Tool." This volume, edited by Dr. Naftali Glasman, Dean of the Graduate School of Education at the University of California at Santa Barbara, has the potential for influencing administrators' perceptions of the purposes of evaluation and their role in increasing its productivity.

The paper cited above and its acceptance for journal publication constitute the deliverable for our 1984 scope of work component dealing with recommendations for improving evaluation productivity.

Organizing for Evaluation Use

As we described in previous reports to the NIE, during 1984 we began

to seek book-length, commercial publication of the Organizing for Evaluation Use handbook and workbook. The handbook summarizes the research on evaluation use, describes and exemplifies the factors influencing use, provides a school-district scenario portraying these factors and how they might be influenced to promote use, summarizes the procedures we devised for organizing an evaluation's use, and provides some examples and worksheets to aid the administrator taking responsibility for such organizing.

The handbook, on the other hand, encapsules the results of our research and is primarily devoted to providing detailed examples and procedures, and a series of worksheets to be used during the evaluation organizing process.

Earlier in 1984 we began negotiations with Sage Publications to produce the Organizing for Evaluation Use handbook and workbook. These negotiations, involving close cooperation between project staff and members of Sage's editorial and marketing units, were time consuming but extremely productive. They led to re-assessment of the market potential for a two-volume product, to reconsideration of the audiences for our work, and to numerous revisions and additions.

For example, as Sage and CSE staff discussed the two-volume market potential, it became apparent that Sage's market interest and CSE's concern for dissemination and application of findings might better be served by combining the two volumes -- the handbook with its extensive narrative and the workbook with its amplified procedures -- into one volume. Therefore, our conceptual scheme for organizing an evaluation's use and the procedures and worksheets we developed to apply that framework are now merged into one volume.

Further, as CSE and Sage discussed the audiences for the published

work, two principal considerations emerged. First, many of the factors discovered to have a bearing on use exert that influence regardless of the particular evaluation setting. Second, evaluation is increasingly being applied in many different social and other program settings. We decided, therefore, to broaden the proposed publication's scope.

To accomplish this end, productivity project staff developed a new scenario, running parallel with the original (which has been retained) but within a different context, to provide an additional bridge between our use-research findings and the procedures designed to promote evaluation use.

The revised work, entitled A Handbook for Evaluation Decision Makers, has been accepted by Sage for publication in 1985. The book and the successful contract with Sage Publications constitute the deliverable for this facet of our 1984 work.

Instances of High Evaluation Use

As we outlined in earlier progress reports, we have analyzed the evaluation descriptions and accompanying documentation submitted to us as part of the High Utilization Evaluation Competition initiated by CSE and Division H at the 1984 AERA annual meeting. We have produced a CSE Report which details the purposes and procedures of the competition, and which will stimulate future evaluator involvement in the competition (see previous progress reports outlining the procedures for continuing AERA recognition for "high utilization evaluation").

In addition, the paper describes the results of our high utilization evaluation analysis and profiles both the evaluations and the factors they share in common. The paper also discusses those factors appearing in the evaluation entries which have counterparts in our utilization framework, as well as possible additional indicators of utilization.

The CSE Report (No. 240), entitled Factors Common to High Utilization Evaluations, represents the deliverable for this component of our 1984 work.

Costs of Evaluation

Earlier CSE work (see, for example, Alkin & Solmon, The Costs of Evaluation, 1983) has suggested some issues involved in attempting to apply existing cost methodologies to educational evaluation. We have noted, however, that given the idiosyncratic nature of the evaluation use process, it may be difficult to generalize cost implications from one evaluation setting to another.

We have produced a CSE Report which provides some further perceptions on the costs of evaluation, focusing on direct and indirect cost implications of various evaluation factors. This report (No. 239), entitled Reflections on Evaluation Costs: Direct and Indirect, constitutes the deliverable for this component of our 1984 scope of work.

Demonstration/Documentation of the Use Process

As we proposed in our 1984 scope of work, the handbook cited above is being implemented in representative school districts throughout the country to demonstrate the use process. In each of the selected sites, a local administrator has assumed the responsibility (1) for organizing a program's evaluation for use, and (2) for coordinating the procedures we are planning for demonstration/documentation at each local site. Site selection proceeded as follows:

With the assistance of practitioners and researchers who nominated candidate school districts for participation in our demonstration activities, we compiled a nationwide sample of districts conducting a formal program evaluation. The sites were selected so as to provide

variety in features such as size, location, population, program emphasis, and evaluation needs. The sampling pool was sufficiently large to ensure that we would attain the eight to ten broadly representative sites that we proposed. Next, we conducted exploratory telephone conversations with district administrators (normally the superintendent) to discuss the project and possible district participation. Each administrator indicating interest received a formal letter of invitation and a description of the project including respective CSE and district responsibilities in the demonstration-documentation effort (see Appendix A for copies of these materials).

In each district accepting the invitation to participate, the handbook was sent to a designated administrator who is assuming responsibility for organizing the evaluation for use. These individuals range from superintendents, to deputy superintendents, to directors of the program being evaluated.

Through this process, the handbook is being implemented and the use process is undergoing demonstration in the target number of sites in different states covering the eastern, western, southern, and mid-western regions of the country.

Documentation of the use-demonstration process does not begin until 1985. However, CSE has begun to collect the baseline information necessary to that documentation effort. As described in our scope of work, the documentation efforts are geared toward describing the settings, factors of influence, use-enhancing strategies, and degrees of evaluation use in each setting. This information, and other information mentioned above, will be a part of the documentation deliverable provided to the NIE in 1985.

Preliminary Planning for the Synthesis of the Knowledge on Evaluation Use

As we described in earlier progress reports to the NIE, our preliminary planning for the 1985 CSE-LRDC synthesis of the knowledge on evaluation use is underway. The synthesis is planned to take place in a working meeting in which four or five evaluation-use papers will provide the nucleus for developing suggestions for improving evaluation practice and use.

Our preliminary planning to date has involved staff of CSE, LRDC staff, and other synthesis contributors. So far our plans include the following:

Four or five major papers will be prepared so as to provide the nucleus described above.

One of the papers will be prepared by CSE staff (Alkin, Burry, & Ruskus). This paper will distill our findings on evaluation use over the past decade, present the conceptual model of use that we developed from the research base, and offer implications for evaluation practice.

A second paper will be written by staff of LRDC (Cooley & Bickel), and will draw on that Center's technical assistance role in the Pittsburgh school's evaluation efforts. This paper will identify factors found to be relevant to use in the Pittsburgh setting and describe strategies adopted there to stimulate use.

A third paper, to be prepared by Michael Patton, will synthesize the evaluation use research, in education and other settings, that has evolved over the past decade.

A fourth paper, to be developed by Jean King, will focus on the implications of research findings for improving evaluation practice.

We plan to share these papers among the presenters and other selected experts who have contributed to evaluation use research, practice, or theory. This selected group, which will also participate in the synthesis, will include such experts as Michael Kean, Nathan Kaplan, and Ross Connor. In addition, our NIE project monitor, Susan Klein, has agreed to be a synthesis participant.

We plan to have the participants meet in California for a two-day working session in February of 1985. In this session, the contents of each paper, to be read in advance by all participants, will be discussed and weighed for breadth and accuracy.

Each paper will be summarized by the author(s) and critiqued by participants from the standpoint of their own research, knowledge, and experience. This critique is seen as a vehicle for suggesting additional areas the synthesis may address, for conceiving implications for the field, and for generating group-sanctioned suggestions for improving practice.

In this way, then, we will bring together in a working session about twelve evaluation use experts who will jointly contribute to the synthesis, its implications, and recommendations for future practice. A court stenographer will record the entire proceedings.

After the proceedings of the February meeting have been transcribed, we plan to initiate 1985 editorial work leading to a draft, book-length work in potentially publishable form. The issues outlined above are among the topics to be fleshed out as we continue or joint synthesis planning with LRDC.

Summary

During 1984 we have provided a series of recommendations for improving evaluation productivity. These recommendations (CSE Report No. 225) will appear in a special edition of Studies in Educational Evaluation in 1985.

Our work on organizing for evaluation use and the handbook guiding such organization has been brought to a successful conclusion with the receipt of the contract with Sage Publications to produce our Handbook for Evaluation Decision Makers in 1985.

We have documented instances of high evaluation use in a CSE Report (No. 240) entitled Factors Common to High Utilization Evaluations.

We have offered some additional thoughts on evaluation costs in a CSE Report (No. 239) offering Reflections on Evaluation Costs: Direct and Indirect.

The pre-publication version of the handbook is being implemented in school districts in different parts of the country as part of our efforts to demonstrate and document the use process. We are collecting baseline data in anticipation of our 1985 documentation needs.

Our preliminary planning for the knowledge synthesis working meeting is well underway.

During 1984, further, we have presented our findings at AERA and are continuing our dissemination activities to advance knowledge of the use process and to improve evaluation practice. For example, we have proposed three presentations -- two papers and a symposium -- for the 1985 annual meeting of the American Educational Research Association (see Appendix B). Our planned activities for 1985 offer additional prospects for formal publications.

APPENDIX A

Organizing for Evaluation Use:

Invitational and Descriptive Material

UNIVERSITY OF CALIFORNIA, LOS ANGELES

UCLA

BERKELEY · DAVIS · IRVINE · LOS ANGELES · RIVERSIDE · SAN DIEGO · SAN FRANCISCO

SANTA BARBARA · SANTA CRUZ

CENTER FOR THE STUDY OF EVALUATION
UCLA GRADUATE SCHOOL OF EDUCATION
LOS ANGELES, CALIFORNIA 90024

Dear :

Thank you for your interest in participating in CSE's project on administrative organizing for evaluation use. As I mentioned on the phone, all of the research on evaluation use agrees that in order to build locally relevant uses into a program's evaluation, a local administrator (e.g., a project director) needs to decide what these uses should be and organize the evaluation around them.

I've enclosed a CSE report describing our research on evaluation use, the factors known to affect use, and the process of organizing for use. The report introduces the product -- Organizing for Evaluation Use: A Handbook for Administrators -- that your designated administrator-organizer will receive from us and follow as he or she works with the program evaluator to plan for and follow up on uses for the evaluation.

I've also included an outline describing our project's overall structure and the role of the schools or districts working with us. CSE's role is not to influence how you evaluate your programs but rather to help ensure that these evaluations have a high use potential. To help with this, in addition to the Handbook, CSE staff will be available for telephone consultation, and I can make a short, orientation visit to participating sites in the fall of 198 .

We will ask participating schools and districts to help us document the use process. This will primarily consist of filling out an occasional (short) questionnaire and some telephone follow up, and asking your administrator-organizer to maintain a record of correspondence, meetings, and so forth, between him/herself and the evaluator. This documentation will also provide some information about the project director and the evaluator, and describe the project being evaluated and its overall context.

Please let me know as soon as possible if you and your district will work with us in this project, so I can send you the Handbook in time for your administrator-organizer to become familiar with its content before the school year begins.

Again, thank you for your interest. I hope we can work together in this project and jointly offer an important contribution to improved evaluation practice.

Cordially,

James Burry
Senior Research Associate

JB:klf
Encl.

ORGANIZING FOR EVALUATION USE

CSE has been conducting research on evaluation use for the past several years. This NIE-funded work has uncovered a variety of factors that affect the degree to which program evaluations and the information they provide are used for local program needs.

One of the central findings pinpoints the need for a local administrator (e.g., a superintendent; a program director) to share with the evaluator the responsibility for planning locally-important uses for evaluation, and for organizing the evaluation in ways that help ensure that these planned uses have a strong chance of taking place. Without this local leadership, it is extremely difficult, and often impossible, to make local use happen.

To meet that administrative-evaluative need, CSE has developed a product called Organizing for Evaluation Use: A Handbook for Administrators. This handbook summarizes the relevant evaluation-use research findings; describes the factors found to have an effect on evaluation use; exemplifies these factors at work in a variety of educational settings; and shows how an administrator might influence the factors to help promote evaluation uses.

With these examples as a constant referent, the local administrator using the Handbook decides on the uses/users for the evaluation, decides which of the factors are likely to have an influence on use in his/her setting, and then follows a series of step-by-step procedures, with planning and worksheets for each, to make sure that these factors have a strong chance of working for the intended uses. These organizing procedures begin while the evaluation is being planned and continue throughout the program's evaluation.

NIE has asked CSE to implement the Handbook in districts throughout the U.S. to demonstrate this evaluation-use process. We are looking for districts which:

- will be doing a program evaluation, 1984-1985;
- will designate a local administrator who will be responsible for working with the evaluator to organize the evaluation for use around locally-relevant needs such as program planning and improvement; curriculum evaluation; staff development; etc.

We would like to select districts that vary in size and geographic location; we expect to select some sites whose evaluator is a regular district/school employee and others whose evaluator is hired externally; we are interested in having the Handbook used in both externally and locally-funded programs.

For each school or district participating, CSE will supply the Handbook, will make staff available for telephone consultation, and arrange for a short visit to each district in the Fall of 1984.

Each participating school or district will be asked to supply CSE with information (via short questionnaires; telephone follow-up) to help us document the use process at each site.

Project period is from July 1984 to August 1985.

James Burry, Senior Research Associate
(213) 206-1508

APPENDIX B

Proposed AERA Presentations on Evaluation Use

Proposal ID# _____

PAPER
Proposal Cover Sheet
1985 AERA Annual Meeting

Session ID# _____

1. Paper title Improving Evaluation Practice and Use Through Administrative Leadership

2. Presenting author Burry (last name) James (first name (no initials)) _____ (middle initial)

   Affiliation UCLA Center for the Study of Evaluation

   Please abbreviate your name (last name first) and address to 70 characters (including between-word spaces) for the Program address directory. Be sure to include ZIP code.

B	U	R	R	Y		J	A	M	E	S		C	N	T	R		F	O	R		S	T	U	D	Y		O	F		E	V	A	L	U
A	T	I	O	N		U	C	L	A		L	O	S		A	N	G	E	L	E	S		C	A	9	0	0	2	4					

3. Coauthors (list full name, institutional affiliation)

   Alkin, Marvin C
   UCLA Center for the Study of Evaluation

4. Subject descriptors: Indicate one- or two-word (maximum two-word) descriptors for your paper to be used in the subject index of the Program. (See Section III-A.)
   1. managing evaluations 2. using evaluations 3. administrative leadership

5. If you wish to have this paper grouped in the same session with other papers submitted to this division, please attach a separate sheet listing presenting authors and titles of the other papers. If you prefer assignment to a 40-minute poster or small roundtable format rather than a paper or critique session, please check here _____ Poster _____ Small Roundtable

6. Will this presentation require audio-visual equipment?
   (See A-V policy, Section II) No X Yes _____ (specify equipment) _____
   (Person signing this cover sheet will be invoiced for any A-V.)

7. Are you an AERA member? X yes _____ no. If no list AERA sponsor _____

I hereby certify that this paper has not been submitted to any other division or Special Interest Group in AERA, and if this paper is accepted and placed on the program, I promise to appear and deliver it. I also certify that this paper has not been previously published or presented at another professional meeting.

James Burry
Signature

8/13/84
Date

Be certain that all of the following are enclosed:
6 SETS OF MATERIALS, STAPLED TOGETHER, EACH SET CONTAINING ONE OF EACH OF THE FOLLOWING:
Paper proposal cover sheet 2–3 page summary 50–100 word précis (for 3 sets only)
PLUS:
Two 3" × 5" index cards with author's name, address and presentation (see Section III-B)
Four self-addressed stamped envelopes.

THIS INFORMATION MUST BE RECEIVED BY THE APPROPRIATE PROGRAM CHAIRPERSON BY AUGUST 15, 1984
First Class Mail is STRONGLY recommended
Air Mail is STRONGLY recommended for international mail

Improving Evaluation Practice and Use Through Administrative Leadership

James Burry and Marvin C. Alkin
Center for the Study of Evaluation, UCLA

This paper takes the perspective that an evaluation's potential for serving program-level decision-making needs requires administrator leadership and administrator-evaluator collaboration. That collaboration should guide specification of the evaluation's questions and the procedures selected to answer them. The paper presents an organizing framework that administrators can apply as they plan and manage their program evaluations and evaluation uses. The framework describes and exemplifies factors which can affect the use-management process and demonstrates tactics that administrators can apply to moderate these factors in the interest of use.

Improving Evaluation Practice and Use Through Administrative Leadership

JAMES BURRY AND MARVIN C. ALKIN
UCLA Center for the Study of Evaluation

James Burry
Center for the Study of Evaluation
145 Moore Hall
University of California, Los Angeles
Los Angeles, CA 90024

Objectives

This paper examines the development of a framework for organizing evaluations to increase their use at the local program level. The principal objectives of the paper are: (1) to describe a framework that administrators can use for organizing a program's evaluation to meet specified uses in light of factors which can influence these uses; (2) by reporting our observations on how the use-influencing factors identified in the framework are manifested and how administrators can deal with these factors to promote use; and (3) to demonstrate the use process to provide administrators with a means of managing their own program evaluations to meet specified uses.

Perspectives

A principal function of evaluation is to provide information to serve program decision-making needs. Administrative leadership and administrator-evaluator collaboration in specifying these decision needs and planning and conducting the evaluation around them is vital to that function.

A variety of factors -- human, context, and procedural -- exist in any evaluation setting and can influence the degree to which an evaluation can successfully be used to meet program needs. These factors include evaluator role, administrator commitment to use, range of questions asked of the evaluation, appropriateness of the information provided to program users.

An evaluation has an increased potential for being put to use to the extent that a program administrator understands the full range of these factors and the broad kinds of influences they can have; ascertains the specific effects they are likely to have in his or her own evaluation setting; and applies organizing procedures which pay attention to these effects so as to enhance the degree to which the evaluation and its uses can be successfully managed.

To assist the program administrator in this endeavor, the framework presented and analyzed places use-influencing factors into patterns which reflect a validated conceptualization of the evaluation use process and which can be managed to meet specified program-level uses.

Methods

The framework and its conceptual schema grew out of a decade of CSE research on evaluation use funded by the National Institute of Education. A series of evaluation case studies (Alkin, Daillak, & White, 1979), an evaluator field-observation study (Daillak, 1980), and an evaluation-user survey (Stecher, Alkin, & Flesher, 1981) formed the conceptual underpinning of the evolving framework. These studies

demonstrated the use process in actual evaluations, illuminated some of the factors representing various perspectives and degrees of influence on an evaluation's use potential, and uncovered the tactics adopted by evaluators and users to stimulate evaluation use.

It began to be apparent that some of the identified factors can potentially be moderated by an evaluator to stimulate the use process. However, it became equally clear, across all studies, that a program-level administrator is frequently in a better strategic position than the evaluator to stimulate evaluation use. But the extent to which a program administrator can stimulate use will depend upon his or her ability to determine which factors will have a bearing in the particular evaluation, to plan specific evaluation uses/users for the evaluation in light of the factors that will affect these uses, and to collaborate with the evaluator to organize the evaluation, its procedures, and its information reporting systems in ways that will influence the factors in the direction of the intended uses.

This conceptualization of the use process would clearly lead to high-utilization evaluations only if it could be supported by an organizing framework which recognized the constraints of school-level evaluations and the broader decision arena of educational administrators who would be asked to assume the use-organizing responsibility.

To determine what such a framework might look like, we spent a year analyzing the literature on evaluation use (Burry, 1983). This analysis uncovered a broad range of factors, each of which had been identified in empirical studies and further substantiated in numbers of theoretical/conceptual works, which can influence evaluation use.

Our next task was to determine possible relationships among the identified factors and to assess the extent to which individual factors or groups might be moderated in the interest of use. This investigation, which examined various factors and their relationships in detail, allowed us to construct a set of core factors which are likely to have a bearing in most evaluations and which can be managed in the interest of use by a program administrator who assumes responsibility for this task (Burry, 1983, 1984).

Having specified the core factors of interest, we constructed a four-stage use-organizing framework for administrators which was validated and refined on the basis of reviews by evaluation practitioners and theoreticians as well as on the basis of try-out in actual evaluations.

The research findings, organizing framework, core factors, and moderating tactics are now presented in a handbook for administrators for managing their program evaluations around a set of intended uses (Burry & Alkin, 1984; Alkin et al, 1984, in press). Final validation of the framework is currently taking place in a nation-wide sample of school- and district-level program evaluations in which the handbook is guiding the evaluation use process.

Data Source and Results

The studies described above, the careful review and try-out process, and the resultant framework all focus on the principal use-influencing factor previously described: The extent to which a program administrator assumes responsibility,

in collaboration with the evaluator, for organizing the evaluation around intended uses is one of the principal determinants of an evaluation's potential for use. The administrator's organizing handbook demonstrates this conceptualization of the use process.

Educational Significance

Research already cited has suggested that an evaluator working alone encounters many obstacles to use (Burry & Alkin, 1984). Research has also indicated that administrator leadership and administrator-evaluator collaboration is frequently required to realize an evaluation's potential for use at the local program level (Alkin et al, 1984).

The conceptual schema described here and its operationalization in the administrator's handbook will guide that collaboration. The handbook demonstrates the use process in various settings, and contextualizes an organizing framework that administrators can apply to influence evaluation practice and increase its use potential in their own programs.

Understanding the need for administrative leadership in the use process and demonstration of the forms such leadership can take will increase the ability of educational administrators, and the staffs they manage, to monitor and improve their instructional programs via improved evaluation practice.

REFERENCES

Alkin, M.C., Daillak, R.H., & White, P. Using evaluations: Does evaluation make a difference? Beverly Hills, CA: Sage Publications, 1979.

Alkin, M.C., Jacobson, P., Burry, J., White, P., & Kent, L. Handbook for evaluation decision makers. Beverly Hills, CA: Sage Publications, in press.

Burry, J. Synthesis of the literature on evaluation use. Los Angeles: UCLA Center for the Study of Evaluation, 1983.

Burry, J. Federal requirements affecting evaluation use. Los Angeles: UCLA Center for the Study of Evaluation, CSE Report No. 222, 1984.

Burry, J., & Alkin, M.C. The administrator's role in evaluation use. A paper presented to Division A at the annual meeting of the American Educational Research Assocation, 1984. Los Angeles: UCLA Center for the Study of Evaluation, CSE Report No. 225, 1984.

Daillak, R.H. A field study of evaluators at work. Los Angeles: UCLA Center for the Study of Evaluation, CSE Report No. 154, 1980.

Stecher, B.M., Alkin, M.C., & Flesher, G. Patterns of information use in school level decision making. Los Angeles: UCLA Center for the Study of Evaluation, University of California, CSE Report No. 160, 1981.

PAPER
Proposal Cover Sheet
1985 AERA Annual Meeting

Proposal ID# ______

Session ID# ______

1. Paper title Factors Common to High Utilization Evaluations

2. Presenting author Ruskus (last name) Joan (first name (no initials)) A (middle initial)

Affiliation Department of Education, UCLA

Please abbreviate your name (last name first) and address to 70 characters (including between-word spaces) for the Program address directory. Be sure to include ZIP code.

RUSKUS JOAN, DEPT. OF ED. 230 MOORE
HALL, UCLA, 405 HILGARD L.A. CA 90024

3. Coauthors (list full name, institutional affiliation)

Alkin, Marvin C., UCLA

4. Subject descriptors: Indicate one- or two-word (maximum two-word) descriptors for your paper to be used in the subject index of the Program. (See Section III-A.)
1. Evaluation Utilization 2. ______ 3. ______

5. If you wish to have this paper grouped in the same session with other papers submitted to this division, please attach a separate sheet listing presenting authors and titles of the other papers. If you prefer assignment to a 40-minute poster or small roundtable format rather than a paper or critique session, please check here ______ Poster ______ Small Roundtable

6. Will this presentation require audio-visual equipment?
(See A-V policy, Section II) No ______ Yes X (specify equipment) ______
(Person signing this cover sheet will be invoiced for any A-V.)

7. Are you an AERA member? ______ yes X no. If no list AERA sponsor Alkin, Marvin C.

I hereby certify that this paper has not been submitted to any other division or Special Interest Group in AERA, and if this paper is accepted and placed on the program, I promise to appear and deliver it. I also certify that this paper has not been previously published or presented at another professional meeting.

Signature: Joan A. Ruskus Date: August 10, 1984

Be certain that all of the following are enclosed:
6 SETS OF MATERIALS, STAPLED TOGETHER, EACH SET CONTAINING ONE OF EACH OF THE FOLLOWING:
Paper proposal cover sheet 2–3 page summary 50–100 word précis (for 3 sets only)
PLUS:
Two 3" × 5" index cards with author's name, address and presentation (see Section III-B)
Four self-addressed stamped envelopes.

THIS INFORMATION MUST BE RECEIVED BY THE APPROPRIATE PROGRAM CHAIRPERSON BY AUGUST 15, 1984
First Class Mail is STRONGLY recommended
Air Mail is STRONGLY recommended for International mail

-Precis-

This paper utilized data generated from the AERA Division H competition, Excellence in Evaluation Utilization, in identifying factors significant to high utilization. The variables were compared to an empirically derived utilization framework. In addition to validating factors already identified in the framework, several new indicators emerged which will be considered for expanding the factor framework. The study has practical implications for educational evaluators in directing them to aspects of the evaluation process that have high potential for maximizing utilization.

Factors Common to High Utilization Evaluations

RUSKUS, JOAN, UCLA

ALKIN, MARVIN C., UCLA

Ruskus, Joan
Dept. of Ed. - UCLA
405 Hilgard Avenue
L.A., CA 90024

-Summary-

Objectives

The purpose of this study was to distill the variables common to evaluations that indicated success in generating high utilization. The empirical basis of the analysis was the comparison of user validated high utilization evaluations with an empirically derived evaluation use framework. The unique feature of this study is that the variables identified emerge from a data base created by an AERA Division H competition for high utilization evaluation. These data are documented by users and evaluators as critical to actual utilizations.

The study has both a pragmatic and theoretical focus. The findings will provide practitioners with empirically validated information for conducting educational evaluations in a way that will maximize utilization, and they will contribute to theory building in evaluation utilization.

Perspective

The definition of use employed in this paper is one developed by Alkin (1975) and Patton et al. (1978). Use is not limited to the direct application of evaluation findings to immediate decision making, but rather it includes the supporting role often played by evaluation findings in the multivariate, incremental process of decision making. This perspective on utilization was further developed by the Center for the Study of Evaluation through a series of evaluation case studies (Alkin, Daillak, & White, 1979), an evaluator field study (Daillak, 1980), and an administrator user survey (Stecher, Alkin, & Flesher, 1981). Based on these empirical data, Alkin et al. (1983) identified and classified the factors affecting evaluation use into three interrelated categories -- human (evaluator and user), context (pre-existing evaluation bounds, organizational features, and project characteristics), and evaluation (procedures, information dialogue, substance of information, and reporting).

The data provided by the evaluators of the studies reviewed for this analysis were analyzed in the context of this factor framework. While a large number of the utilization factors were validated, several new indicators emerged which will be considered further for incorporation into the more comprehensive factor

framework.

Data Source

Twelve evaluation studies were used as the basis of this analysis. Summaries of these studies and supporting materials from relevant users were submitted to the AERA Division H Award for Excellence in Evaluation Utilization competition for 1983. While all of the evaluations dealt with educational programs, the majority (seven) focused on public school programs at the district level. One addressed a state-wide educational program, three were directed at university programs, and one targeted vocational education for a special needs population. Three of the studies specifically focused on teacher effectiveness.

Application for the competition required each evaluator to discuss unique factors which they judged to be significant in contributing to the ultimate utilization of their evaluation. These data were analyzed for the purpose of this study.

Method

Applicants' narrative data on unique factors contributing to utilization were analyzed by abstracting all statements made in the application and documentation. Data were given equal weight regardless of the applicant's placement in the utilization competition since all applications represented well documented instances of utilization. The indicators represented in the data matrix were then classified according to the utilization framework. Most indicators were easily subsumed under existing major factor categories, although some indicators presented extensions of existing factors or their interpretation. Frequencies were then tabulated by factor.

Results

The results confirmed many of the factors identified by Alkin et al. as significant. The citing of factors within the evaluator's control as most significant constituted one dominant theme in the findings. Evaluator characteristics accounted for 31% of use and evaluation factors accounted for 52% of use. Factors outside of the evaluator's direct control (user characteristics and contextual factors) accounted for only 17% of use. However, we should keep in mind that in a self report schema it is likely that personal rather than the contextual characteristics would predominate.

Among evaluator characteristics, willingness of the evaluator to involve users in the evaluation was the most frequently cited attribute. Two facets of the evaluation stood out as critical to use: evaluation procedures and evaluation reporting. The most frequently highlighted evaluation procedures were the evaluation plan, evaluation methods, and follow-up activities. The most frequently identified reporting features were the

frequency of reports, the format of reports, the readability of report narrative, and the dissemination of reports.

Several indicators were identified in this analysis which were either not included in the factor model or which expanded the meaning of existing factors. These indicators are noted along with the existing related factor in parentheses. They are: sense of urgency or need for the evaluation to improve an ailing program (external organizational feature); a structured evaluation plan (evaluation procedures); dissemination of the report (evaluation reporting); and follow-up activities after the evaluation is completed (evaluation reporting).

Educational Significance

Increasing the extent of utilization of educational program evaluations has clear implications for improving programs. This empirical validation of utilization factors has practical implications for practitioners. Evaluators should focus on assuring user's working involvement in the evaluation, designing an evaluation plan that accommodates the information needs of all users, utilizing evaluation methods that are technically valid but congruent with program practices, generating readable reports, and conducting follow-up activities.

The study has implications for evaluation utilization theory as well. First, the study validated the existing utilization framework and suggested modifications and needed amplification. The findings also highlighted a possible evaluator blind spot -- contextual variables. Exploration of factors outside of the evaluator's direct control and ways these factors can be successfully influenced to enhance use is an important step in promoting more widespread utilization of evaluations.

REFERENCES

Alkin, M.C., Evaluation: Who needs it? Who cares? Studies in Educational Evaluation, 1975, 1, 201-212.

Alkin, M.C., Daillak, R.H., & White, P. Using Evaluations: Does evaluation make a difference? Beverly Hills, CA: Sage Publications, 1979.

Alkin, M.C., Jacobson, P., Burry, J., White, P., & Kent, L. Organizing for evaluation use: A handbook for administrators. Los Angeles: UCLA Center for the Study of Evaluation, 1983.

Daillak, R.H. A field study of evaluators at work. Los Angeles: CSE Report No. 154, UCLA Center for the Study of Evaluation, 1980.

Patton, M.Q., Grimes, P.S., Guthrie, K.M., Brennan, N.J., French, D.B., & Blyth, D.A. In search of impact: An analysis of utilization of federal health evaluation research. In T.D. Cook and associates (Eds.), Evaluation Studies Review Annual, Vol. 3. Beverly Hills, CA: Sage Publications, 1978.

Stecher, B.M., Alkin, M.C., & Fleischer, G. Patterns of information use in school level decision making. Los Angeles: CSE Report No. 160, UCLA Center for the Study of Evaluation, 1981.

Proposal ID# _____

SYMPOSIUM
Proposal Cover Sheet
1985 AERA Annual Meeting

Session ID# _____

1. Title A Synthesis of the Knowledge on Evaluation Use

2. Organizer Alkin (last name) Marvin (first name) C (middle initial)

   Affiliation UCLA Center for the Study of Evaluation

   Mailing address CSE 145 Moore Hall UCLA Los Angeles, CA 90024

3. Chair (if different from organizer) Law (last name) Alex (first name (no initials)) (middle initial)

   Affiliation California State Department of Education

   Mailing address 721 Capitol Mall Sacramento CA 95814

4. Participants: please attach a separate sheet listing full name, institutional affiliation, mailing address and title of presentation for each participant. Please abbreviate the name (last name first), and mailing address (including ZIP code) of each presenter to 70 characters (including between-word spaces) for inclusion in the Program address directory. Only names and addresses 70 characters or less will be published in the directory.

5. Time length request: 1½ hrs X 2 hrs _____

6. Are you a member of AERA? X yes _____ no. If not, please list AERA sponsor _____

7. Subject descriptors: Indicate three one- or two-word (maximum two-word) descriptors for this symposium to be used in the subject index of the Program.
   (See Section IV-B)
   1. evaluation use 2. knowledge synthesis 3. improving evaluations

8. Will this presentation require audio-visual equipment?
   (See A-V policy, Section II) No X Yes _____ (specify equipment) _____
   (Person signing this cover sheet will be invoiced for any A-V equipment ordered.)

9. If you wish the symposium to be considered for cosponsorship, please suggest other division, SIG, or NCME _____.

I hereby certify that this proposal has not been submitted to any other division or Special Interest Group in AERA. All participants named have assured me of their willingness to participate and have certified that their papers have not been previously published or presented at other professional meetings.

Marvin C Alkin (Signature) 8/13/84 (Date)

Be certain that all of the following are enclosed:

6 SETS OF MATERIALS STAPLED TOGETHER, EACH CONTAINING ONE OF EACH OF THE FOLLOWING:
- Symposium proposal cover sheet
- 50–100 word précis (for 3 sets only)
- 2-3 page summary

PLUS
- Two 3" × 5" index cards with symposium title and name and address of organizer and chairperson
- Four self-addressed stamped envelopes
- Two copies of a list of all participants and their respective mailing addresses

THIS INFORMATION MUST BE RECEIVED BY THE APPROPRIATE PROGRAM CHAIRPERSON BY AUGUST 15, 1984
First Class Mail is STRONGLY recommended
Airmail is STRONGLY recommended for international mail

A Synthesis of the Knowledge on Evaluation Use

Law, Alex CSDE 721 CAPITOL MALL SACRAMENTO CA 95814
Alkin, Marvin CNTR FOR STUDY OF EVALUATION UCLA LOS ANGELES CA 90024
Burry, James CNTR FOR STUDY OF EVALUATION UCLA LOS ANGELES CA 90024
Ruskus, Joan CNTR FOR STUDY OF EVALUATION UCLA LOS ANGELES CA 90024
Cooley, William LRDC-UNIV OF PITTS 3939 OHARA ST PITTSBURGH PA 15260
Bickel, William LRDC-UNIV OF PITTS 3939 OHARA ST PITTSBURGH PA 15260
Patton, Michael UNIV OF MINN ST PAUL MN 55108
Kean, Michael CTB/MCGRAW-HILL 2500 GARDEN ROAD MONTEREY CA 93940
Klein, Susan NIE 1200 19TH ST NW WASHINGTON DC 20208

Chair	Alex Law
Presentations	Alkin, M.C., Burry, J., & Ruskus, J. Improving Evaluation Use: Implications for Practice
	Cooley, W., & Bickel, W. Evaluation Use: Pittsburgh Case Histories
	Patton, M. State-of-the-Art in Evaluation Use
Discussants	Kean, M. Klein, S.

A Synthesis of the Knowledge on Evaluation Use

This symposium synthesizes and contextualizes the knowledge of the evaluation use process which has evolved over the last decade. Each of the participants has been active in conducting research on evaluation use and/or conducting high-utilization evaluations. Presentations offer a framework for organizing evaluations to promote their use; discuss examples of use from actual evaluations, including the factors influencing use and strategies for moderating them to stimulate use; provide a state-of-the-art synthesis of the principal use findings in education and in other disciplines; and discuss implications for the practice and uses of evaluations.

A Synthesis of the Knowledge on Evaluation Use

The purpose of this symposium is to present a variety of current information reflecting knowledge of the evaluation use process and ways to improve evaluation practice and increase its use.

The principal objectives of the symposium are as follows:

- To present a framework for organizing evaluations for use in light of various influences on the use process;
- To provide examples, from actual evaluations, of the use process, as well as factors influencing use and strategies for moderating these influences to promote use;
- To provide a state-of-the-art synthesis of the principal findings on use, in education and in other disciplines;
- To discuss implications for the practice and uses of evaluation.

The symposium chair, presenters, and discussants bring a wide and varied background to the study and practice of evaluation. The UCLA Center for the Study of Evaluation presenters, Alkin, Burry, and Ruskus, have been conducting research on evaluation use for the past decade. The Pittsburgh Learning Research and Development Center presenters, Cooley and Bickel, have a long and successful history of conducting evaluations with high levels of use. Each of these two organizations has conducted their work under NIE auspices. Patton is one of the earliest evaluation use researchers whose work continues to influence the conduct of evaluations. The chair and discussants have published in the field of evaluation use.

The symposium summarizes, contextualizes, and updates a large and growing body of research on evaluation use. Another important contribution of the symposium, however, is the particular evaluation perspective it embodies. That is, the symposium demonstrates how evaluations should be conducted to meet particular uses in various settings. This demonstration has implications for both the providers and the users of evaluation information.

The three major presentations are as follows:

Improving Evaluation Use: Implications for Practice
Marvin C. Alkin, Project Director; James Burry, Senior Research Associate; Joan Ruskus, Senior Research Associate
UCLA Center for the Study of Evaluation

CSE staff have been conducting research on evaluation use in a series of empirical investigations over the past decade. For example, on the basis of evaluation case studies, evaluator field-observation studies, and

evaluation-user surveys, staff identified many factors which can act independently or in concert to have a marked effect on an evaluation's potential for use. These factors, consisting of human, context, and evaluation procedural considerations, require careful attention in order for an evaluation to be used.

How that attention might be applied has been the subject of recent CSE investigation. That investigation led to the development of a framework which can be used to organize an evaluation to meet specified uses given the factors likely to affect these uses in the particular setting.

This paper summarizes the CSE research leading to the organizing framework. In particular, it describes and exemplifies a set of core factors which, since they are likely to have a bearing in most evaluation settings, form the basis of the organizing framework. The paper describes how the organizing framework is used, offers procedures for improving evaluation practice and increasing its use, and discusses implications for the field.

Evaluation Use: Pittsburgh Case Histories
William Cooley and William Bickel, Senior Scientists
Pittsburgh Learning Research and Development Center

For the past several years, members of the evaluation unit at LRDC have been investigating strategies for improving the content and use of evaluation research in school districts. This goal has been pursued through collaboration with the Pittsburgh school system. The evaluation unit has taken on a number of Pittsburgh research tasks, ranging from evaluations of program impact to district-wide and program-focused needs assessments, designed to help managers establish district and program priorities. What has been learned about district use of evaluation research is being summarized in a series of case histories.

This paper draws upon several of the Pittsburgh case histories to discuss the issue of evaluation use. Specific examples of use are provided from individual cases. Factors that seemed to play a critical role in the use of research information are identified. One example of such a factor concerns the importance of taking a client orientation in the organization of evaluation research. The paper places emphasis on the discussion of research strategies that can increase the likelihood of use, and of actions that both researchers and managers can take to improve the utility of evaluation processes for school systems.

State-of-the-Art in Evaluation Use
Michael Q. Patton, Program Director, CAEP
University of Minnesota

This paper summarizes the principal findings on evaluation use that have evolved over the past decade. It complements the other symposium papers by reviewing the research on evaluation use in education, and supplements them by drawing from the research on use conducted within other disciplines. The author, as one of the early pioneers in the field, is in an ideal position to comment on and synthesize the current state of evaluation utilization research.

The paper covers the entire decade of research on evaluation use, beginning with some of the earlier research which the author himself conducted. Among this work was his 1975 analysis of evaluation utilization in health evaluation, and his 1978 Utilization Focused Evaluation which described the early research and summarized the then extant literature on use. From that vantage point, this paper provides the opportunity to update the evaluation use literature, both in educational settings and in other disciplines, and to suggest the implications of the findings.

The symposium chair, Alex Law, is head of the Office of Program Evaluation and Research, California State Department of Education.

The discussants are Michael Kean, Publisher, CTB/McGraw-Hill, and Susan Klein, Senior Associate, National Institute of Education.

The chair and discussants themselves bring another level of expertise to the topic of evaluation use, as represented in publications reflecting problems in district-level uses of evaluation; the administrative uses of evaluation; and federal involvement in evaluation use.

DELIVERABLE - NOVEMBER 1984

PRODUCTIVITY PROJECT

Marvin C. Alkin
Project Director

"The Administrator's Role In Evaluation Use"

Grant Number
NIE-G-0112
P-2

CENTER FOR THE STUDY OF EVALUATION
Graduate School of Education
University of California, Los Angeles

The project presented or reported herein was supported pursuant to a grant from the National Institute of Education, Department of Education. However, the opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement by the National Institute of Education should be inferred.

THE ADMINISTRATOR'S ROLE IN EVALUATION USE*

Center for the Study of Evaluation
Graduate School of Education
University of California Los Angeles

James Burry
Marvin C. Alkin
Joan Ruskus

INTRODUCTION

The theme of this journal issue -- evaluation as a management tool in education -- is both critical and timely. It is critical because, "as the shift continues from the federal to the state levels in the management of education programs, the states become more, not less accountable for them. SEAs and LEAs have become accustomed to the federal government not only requiring the evaluation of programs but also dictating methods of evaluation" (Council of Chief State School Officers, 1982). As federal control of evaluation diminishes, and as federal resources for these evaluations also diminish, it seems likely that state and local administrators will need to assume greater responsibility for their evaluations (Burry, 1984). Given scarce resources, they will need to think strongly about the best ways to commit people and money to ensure that their evaluations generate useful information. One primary index of that usefulness will derive from the extent to which evaluation becomes a tool for educational management and decision making.

*The research guiding this article was supported under a grant to the UCLA Center for the Study of Evaluation from the National Institute of Education, U.S. Department of Education. However, the findings and opinions expressed here do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement should be inferred.

The criticality of the issue is also matched by its timeliness. That is, to suggest increased administrator responsibility in evaluation without offering means to channel that responsibility would be less than satisfactory. Fortunately, recent research in evaluation use (Alkin et al, 1985 in press) has enabled us to develop a framework which helps administrators to take a more active, indeed proactive role, in organizing evaluations to increase their effectiveness as a management tool in educational decision making.

Recent attention to the quality of education in our schools (Boyer, 1983: Goodlad, 1983; National Commission on Excellence in Education, 1983) indicates that concern with excellence is pervasive. That pervasiveness suggests that educational decision makers will need information which accurately reflects the extent to which their educational systems are responding to the challenge for excellence. Can evaluation become a useful tool in the management of these systems? We believe that it can, if educational administrators take the kind of actions we discuss here to capitalize on the uses to which evaluation can be put and to ensure that these uses reflect their own system-level questions and needs.

Establishing a local focus for evaluation presents little in the way of technical difficulties but may run counter to present attitudes toward and expectations for evaluation. We will elaborate some of the relevant issues later in our discussion. Suffice it to say, for the present, that if evaluation is to address the demonstration of excellence (or any other desirable quality) then there should be open discussion of issues such as the definition of excellence, how it is to be judged, the best means of

demonstrating its attainment, and the uses of evaluation in addressing these issues.

Burstein (1984) has recently discussed some applications of evaluation in school improvement efforts. While these applications can play a potentially central role in the definition, judgment, and demonstration of excellence, that potential will not be realized until administrators come to recognize the ways in which evaluation information can be used for management purposes. Among these purposes are: pulse monitoring -- treating evaluation information as educational indicators of the extent to which the educational system is moving in the desired direction; student decision making -- using evaluation information to make accurate decisions about student progress and needs; program decision making -- drawing on evaluation information to monitor programs and services and to modify them as needed; informing educational policy -- using evaluation information to guide discussion of the status of educational systems and mechanisms for improvement; and long-range planning -- applying evaluation information in decision areas such as physical plant needs, teacher hiring and assignment, and resource allocation (Burstein, 1984, pp. 16-19).

As Burstein suggests, however, such management applications, which address instructional issues, support systems, and resource allocation at both policy and operational levels, face certain requirements. Among these requirements are commitment to the uses of evaluation information for "informed inquiry and educational change," and "a healthy and informed understanding of the limits as well as the possibilities of information-based decision making" (Burstein, 1984, p. 23). We will amplify these and other requirements as we proceed.

Enabling educational administrators to organize their evaluations to meet the uses outlined above, then, is the topic of our discussion. Recent research (Alkin et al, 1985 in press) has uncovered factors which influence the extent to which evaluation is likely to be put to use. One finding, as we shall see, is that the educational administrator, him- or herself, is a critical factor in the use process. That is, the extent to which the administrator actively influences the direction and course of the evaluation is a principal determinant of the likelihood that it will be put to use. Other research (Ruskus & Alkin, 1984) suggests the kind of administrative influence likely to promote system-wide use. Recognizing and promoting the uses to which evaluation can be put is the first step in establishing its application as a management tool.

Once accepted as a management tool, evaluation can provide a valuable resource for administrators who are interested in finding out how well the system they are responsible for is running, and deciding whether it could be improved. We have suggested that these decision needs can involve, for example, monitoring student and program decision making, policy setting, and long-range planning. To meet these kinds of needs, an evaluation should be planned around questions reflecting the system's context, operations, and expectations. It must be conducted in ways to ensure, first, that these questions are answered and, second, that the answers can actually be put to use in making decisions about whether the system should continue to run as is, if it needs to be modified, what these modifications might entail, and what kinds of policies and resources might be required either for maintenance or modification.

Although there are potential obstacles facing the administrator who wants an evaluation emphasizing such practical uses, there are also organizing principles that can be applied to overcome these obstacles. As we discuss the administrator's role in organizing an evaluation we will suggest ways to strengthen its potential for use. For purposes of discussion, we focus our remarks primarily at the level of a discrete education program such as, for example, Chapter I, bilingual education, mathematics, language arts.

Organizing for Evaluation Use

To have a high potential for use, an evaluation needs to be carefully planned, organized, conducted, and communicated to likely users of the information it provides. This kind of evaluation rarely happens by chance; someone has to take the responsibility to make it happen. Certainly, an evaluator can and should take some of the responsibility for organizing an evaluation for use. However, it has become clear that the role an administrator (e.g., a superintendent; a Chapter I program director) takes with regard to the evaluation has a marked effect on its use potential. We will demonstrate a framework, then, that administrators, working in cooperation with their evaluators, can apply to gain tactical influence over the direction the evaluation takes. That influence is intended to increase the evaluation's potential for use in program management and decision making.

Evaluation Purpose

The framework we propose reflects a particular evaluation perspective. That is, we define evaluation as a means of providing information that can be used to make decisions about programs. These decisions might

stem from questions about whether the program could be improved, they might reflect matters of resource allocation and monitoring, they might stem from questions about whether the kinds of attitudes people have about the program could be improved.

Using evaluation information to assess and perhaps influence participant attitude is likely to make an important contribution to evaluation's potential as a management tool. Some programmatic changes will require modification of participants' attitudes about the program or its evaluation before the change can be implemented successfully. That is, an information-based administrative decision to make some curricular change, to redirect resources, to reassign staff, will require staff support of the proposed change.

That support will require acceptance of the information driving the change, and frequently use of that information by the people who, in addition to the administrator, will play some role in accepting and implementing the change. Therefore, administrator ability to successfully use evaluation as a management tool will require collegial support and information use at various system levels.

For an evaluation to meet its potential as a management tool in a particular decision area, then, the administrator needs to identify other potential information users whose support of the decision is necessary. For each user or user group, the administrator will need to determine what questions and concerns they have with respect to the given decision area, and then make sure that the evaluation applies procedures and reporting techniques which are appropriate to the users and their questions.

The extent to which a program administrator takes responsibility for identifying the intended users, determining their questions about the program, shaping the evaluation procedures for answering the questions, deciding what kinds of information will be collected, and ensuring that the information is effectively communicated can profoundly affect the degree to which the information can successfully be put to use.

Evaluation Use

By use we mean applying evaluation information to the resolution of the kinds of problems, questions, or concerns we have alluded to above. To be sure, evaluation can have other, perhaps unintended consequences, but we do not emphasize them in this paper.

There are many potential users of evaluation information. In a school setting, for example, there might be a variety of programs in operation, such as: a Chapter I program; a state-funded bilingual program; a remedial math or language arts program designed for students in need of specialized instruction.

Each of these programs might be evaluated and each could have a variety of potential evaluation users. For example, let's assume that a district superintendent wanted to have more productive evaluations, wanted to be able to use evaluation information as a management tool in district operations. The administrator might then consider ways to organize the evaluation to meet his or her questions and needs and those of other potential users. These users, in addition to the superintendent, might include the people responsible for program operation, for instance, such as the director, other administrators, curriculum developers, instructional

staff, and funding agencies. Other users might consist of parents, advisory councils, and community organizations with an interest in the program. Since each of these groups can have professional and personal interests in the program and its evaluation, each is a potential user of the information it provides. A central concern in organizing for evaluation use, therefore, is the selection of the intended users of the evaluation.

Evaluation information can be used in a variety of ways. For example, let's follow the case suggested above and assume that the superintendent is concerned about the instructional content and methods used in a remedial mathematics program. As a responsible manager, the superintendent has questions about how students are selected for the program, the extent to which teachers are implementing the program as planned, the extent to which building principals support teachers as they attempt to implement the program, the extent to which resources earmarked for the program are actually used in the program, whether or not the program seems to be beneficial for the students. The superintendent wants the program's evaluation to provide answers to these questions so that he or she can make information-based decisions about maintaining the program as is, modifying the program, maintaining, increasing, or reducing its level of resources.

Now the evaluation may ultimately find that the program seems to be running quite well. On the other hand, it may pinpoint problems and suggest areas for change. Regardless of the evaluation findings, staff in the program are likely to have different conceptions about the program. Some may enjoy working in the program, think it's a good one, and would

like to see it being continued; other staff may take the opposite point of view. Staff are also likely to differ in their expectations for the evaluation. Some may want information to help them as they carry out their responsibilities in the program; others may think that evaluation does not provide the kinds of information they need; some may have no expectations for the evaluation.

In short, program staff, as potential implementers of the superintendent's decisions, are potential evaluation information users. They can differ in the extent to which they have questions about the program, in the kinds of questions they have, and their disposition toward using evaluative answers. For some staff, asking them to make changes may create a problem for the superintendent; for others, asking them to continue current practice may create a problem.

To help preclude these possibilities and to promote the kind of support we mentioned earlier, involving staff and other potential users in the evaluation, finding out their questions and concerns, and determining the kinds of information they are likely to accept and use, are crucial.

F rs Affecting Use

In any setting, there are many factors that can have an effect on evaluation use and therefore on its potential as a management tool. By factor, we have in mind any characteristic or element present in a given situation that can affect the extent to which the evaluation is used. These factors stem not only from the conduct of the evaluation, but also from the surrounding social, political, organizational, administrative, and programmatic context. Factors potentially affecting an evaluation's use,

for example, include the kind of role the evaluator chooses, the intended users' views about the program being evaluated, the various requirements for the evaluation, and its proposed methods.

If these factors are accepted as givens, they can reduce or negate the evaluation's use potential. For instance, if an intended group of users firmly believes that a program could not possibly be improved upon, it may be difficult to convince them to modify their view, no matter what the evaluation findings might reveal. On the other hand, if the evaluation is structured and organized around intended users and kinds of uses, and if the possible effects of various factors on the use potential are planned for, then the evaluation's likelihood for use can be greatly increased.

Later we will describe the full range of factors that have been shown to affect an evaluation's use and discuss an organizing framework administrators can follow to minimize negative factor influence and strengthen positive factor influence. The organizing framework, as well as the associated operating terms we have discussed above, grew out of our research on use over the past several years (Alkin et al, 1985 in press; Burry, 1983).

BACKGROUND ON EVALUATION USE

For a good number of years, the terms use or utilization have been cropping up in the evaluation literature. Up to about the mid 1970's, however, discussions of use relied fairly heavily on impressionistic and anecdotal information. There was a lot of talk reflecting what people thought use looked like, with explanations often relying on speculation (Rossi, 1972; Mann, 1972; Cohen & Garet, 1975).

Around the mid 1970's the picture began to change. Then we began to see the results of systematic research on use, research trying to discover what use actually means, whether or not it occurs, and what works for it or against it (Alkin, 1975; Patton et al, 1975).

To a great extent, the careful study of use grew out of the kinds of promises made for evaluation. For example, evaluation was to be an important tool for decision making and for improving policy and practice. All the evaluator had to do, it was thought, was to provide valid data. People would see the light and use the information provided; decision making would be more rational and policy and practice would improve.

By now we know this was a naive view. Certainly, information validity, especially when that term is mutually agreed upon by evaluator and potential user, can contribute to use. But so long as evaluation and its use were (1) seen as the sole responsibility of the evaluator, and (2) expected to produce quick, observable, and rational decisions in action, the promise was not met.

One of the things explaining the seeming lack of use was that for a long time many people thought that information received was necessarily put to use, and put to use quickly. When that did not bear up in practice, it was assumed that no use was taking place.

As the research was to show, however, use was occurring, though in a form quite different from and perhaps more modest than had been expected (Alkin et al, 1974; Patton et al, 1975). We began to understand that evaluation processes and evaluation information usually accumulate over time before they are finally put to use. And even when they are used in

making a decision, that decision may also have been influenced by other kinds of information and forces outside of the evaluation. This kind of use can and does take place and when it does it can help to improve educational decision making and practice.

However, there is something else that helps explain lack of use. That is, for use to take place, we had thought, such technical factors as the quality of the evaluation's procedures would be important. And that is true. Procedural soundness can certainly contribute to use, but so can other factors, factors that are somewhat removed from the technical realm.

For example, one early CSE finding (Alkin, 1975) showed that the stance taken by the evaluator with respect to a program's social context can affect the evaluation's use potential. Concurrent research (Patton et al, 1975) pointed up the contribution to use of the "personal factor" which is typified, for instance, when someone takes direct responsibility for trying to make use happen.

Until recently, that "someone" was usually taken to be the evaluator, the "provider" of information. Our research, however, as it has amplified the "personal factor" and discovered others contributing to use, demonstrates that the role of the potential "user" of information, such as an administrator, is just as important as that of the evaluator in promoting use. In many situations the evaluator him- or herself will lack the power, prestige, political sensitivity, or contextual understanding necessary to promote use. Our work has shown that use will frequently require the influence of a program administrator who does possess these and other attributes.

CSE Research on Use

Drawing on the early studies mentioned above (Alkin, 1975; Patton et al, 1975), we conducted several empirical studies of evaluation use. Among these were: (1) evaluation case studies; (2) an evaluator field study; and (3) a user survey. These studies contributed to our synthesis of the knowledge on use and led to a practical handbook for administrators who wish to organize their program evaluations for use.

The evaluation case studies: The case studies (Alkin, Daillak, & White, 1979) focused, over a period of two years, on five different programs with required evaluations. These cases provided detailed descriptions of school-level program implementation and evaluation, and how the evaluation process unfolded in each program. Our analyses uncovered the people who shaped the evaluation process, how it was used in each case, how it fitted in with other school operations, and how it influenced decisions about the program. Further, by identifying some of the factors promoting these uses, we were able to develop a conceptual framework to guide our future study of use.

The evaluator field study: Drawing on the emerging framework, Daillak (1980) spent a year as a participant-observer working closely with three school-district program evaluators in the belief that observation and analyses of evaluators -- the providers of information -- at work would illuminate conditions of use. By observing these evaluators at work Daillak was able to elaborate some of our previously identified factors, particularly those reflecting the evaluation's organizational setting, as well as the kinds of tactics that evaluators adopted to increase their use-enhancing effect.

The user survey: The user survey (Stecher, Alkin, & Flesher, 1981) took place over the course of a year in 22 schools in the district in which the field study had previously been conducted. Our concern here was to characterize the role of a particular information user, the program administrator, in terms of the nature of the decisions typically confronting administrators, and to uncover how and what kinds of information come to shape these decisions.

The interviews provided a picture of the kinds of decisions -- programmatic and other -- school administrators need to make to do their jobs, the ways that they use evaluation and other information -- to pinpoint a need, to amplify a previous conclusion -- as they form these decisions, and the broad strategies they adopt to stimulate others to use information in their programmatic responsibilites.

Synthesis and handbook: To help synthesize the knowledge on use we developed an annotated review of the relevant empirical and conceptual-theoretical literature, drawn from educational and other settings, (Burry, 1983), and a handbook for the administrator-user who plans to build use into his or her program evaluation (Alkin et al, 1985 in press). All of our work to this point illustrated the importance of user-evaluator collaboration in promoting use given various factor impacts. The handbook therefore clusters factors into patterns which reflect the stages of the use process and which can be influenced to promote use.

Factors Affecting Evaluation Use

On the basis of the work described above, we identified and classified the individual factors affecting evaluation use into three related cate-

gories -- human, context, and evaluation procedure or methodology. How these factors interact together determines the extent to which evaluation is likely to be used.

Figure 1 lists the three kinds of factors. Those in the human category reflect evaluator and user characteristics that have a strong influence on use. Included here are such factors as people's attitude toward and interest in the program and its evaluation, their backgrounds and organizational positions, and their professional styles.

Context factors include the kinds of requirements and fiscal constraints the evaluation faces, and the relationships between the program being evaluated and other segments of its larger organization and surrounding community.

The evaluation factors refer to the actual conduct of the evaluation, and include how the evaluator and users work together, the procedures used in the evaluation, and the quality of the information it provides.

The factors in each of the three groups have a demonstrated importance to use, and many of them will require administrative influence to promote use. In the next section of the article, therefore, we offer a series of observations drawn from the empirical studies of use. These observations help define each of the factors in Figure 1 and suggest the kinds of influence they may have, as a precursor to discussion of factor interaction patterns and administrative organizing to promote use as a management tool.

Observations Drawn from Empirical Studies

With respect to the human factors affecting use, an evaluation's use potential is likely to increase to the extent that:

Figure 1: Factors Affecting Use

I. Human Factors

A. Evaluator Characteristics
1. commitment to use
2. willingness to involve users
3. choice of role
4. rapport with users
5. political sensitivity
6. credibility
7. background and identity
   a. gender
   b. title

B. User Characteristics
1. identity
   a. range of potential users
   b. organizational positions
   c. professional experience levels
2. interest in the evaluation
   a. views about the project being evaluated
   b. expectations for the evaluation
   c. predisposition toward the evaluation
   d. perceived need
   e. perceived risks
3. commitment to use
4. professional style
   a. administrative and organizational skills
   b. initiative
   c. openness to new ideas or change
5. information processing
   a. preferences for particular forms
   b. how information is processed

II. Context Factors

A. Pre-existing Evaluation Bounds
1. written requirements
2. other contractual obligations
3. fiscal constraints

B. Organizational Features
1. intraorganizational
   a. role of central/district office
   b. interrelationship between unit and central/district administration
   c. institutional arrangements
   d. unit level autonomy
   e. sources of information beyond evaluation likely to be in use
   f. perceived institutional risk
2. external features
   a. community climate
   b. community influence
   c. role of other agencies

C. Project Characteristics
1. age/maturity
2. innovativeness
3. overlap with other projects

III. Evaluation Factors

A. Evaluation Procedures
1. methods used
   a. appropriateness
   b. rigor
2. dealing with mandated tasks
3. used of a general model

B. Information Dialogue
1. amount and quality of interaction between evaluator and users

C. Substance of Evaluation Information
1. information relevance
2. information specificity

D. Evaluation Reporting
1. frequency of information provided
2. timing of information
3. format of presentations
   a. oral presentations
   b. written reports
   c. statistical and narrative data

1. The evaluator --

   - is personally committed to seeing his or her work put to use, and actively makes efforts to facilitate the use of information;
   - is willing to involve users in the evaluation through cooperative planning and conduct of the evaluation and its uses;
   - recognizes that alternative evaluation roles exist, chooses a role that is appropriate in the given setting, and focuses on serving program needs and questions in addition to any external requirements;
   - develops rapport with users by earning their trust in an atmosphere of harmony and agreement;
   - is politically sensitive to the program and understands the relationship among formal and informal power sources, opinion makers, decision making processes, and the function of evaluation as one of the inputs to these processes;
   - establishes credibility in terms of technical competence and personal and professional manner.

2. The users --

   - are clearly identified so that the evaluator understands the range of organizational positions and professional experience levels -- administrative vs. operational, sole or shared decision-making authority, familiarity with evaluation -- which are represented among the users and which bear on their potential for using information;
   - view the project in such ways that they would be willing to modify these views, if warranted;
   - have specific expectations for the evaluation -- determining the program's efficiency, understanding its processes, assessing its outcomes -- which are translated into questions and concerns that the evaluation will address;
   - are predisposed to accepting the evaluation's findings, which may be because they
   - have a high perceived need for evaluative answers to their questions, and
   - perceive the risks of the evaluation as outweighed by the potential benefits. In addition, they

- are personally committed to using evaluation information as their questions and concerns are answered, and
- have sufficient administrative and organizational skills to act on information, to get things done. They will
- take the initiative to use evaluation information in their own area of responsibility and, if necessary, to stimulate others to follow their example. Further, they
- are open to new ideas or change that stem from the findings, even if these findings suggest they need to modify their original views of the project. And, as the evaluation process unfolds, their positive interest in the evaluation remains high, because they
- ask for and receive the kinds of information they prefer to use -- narrative, descriptive, or some combination, through the kinds of processes -- oral reports, written reports, detailed or summary treatments, they are most comfortable or routinely familiar with.

With respect to the context factors affecting use, an evaluation's use potential is likely to increase to the extent that:

1. The pre-existing evaluation bounds --
   - are characterized by a guided harmony rather than by conflict and tension. The evaluation's written requirements -- legal codes, federal/state requirements -- permit sufficient flexibility so that the evaluator can respond to such other contractual requirements as those set by program administrators or operators.
2. The organizational features --
   - are marked by amicable co-existence in an atmosphere stressing discussion and the negotiation of problems and needs;
   - facilitate the central/district office -- often the evaluation sponsor -- role in balancing broad system concerns with those of the individual units, such as the schools who are subject to evaluation;
   - permit sufficient unit level autonomy so that unit (e.g., a school) questions receive a fair share of the evaluator's attention as he or she addresses a variety of broad organizational and unit questions of interest;
   - promote frank discussion of the perceived institutional risks and, where there is a question of whether the evaluation benefits will outweigh the risks, consider the possible outcomes and resultant actions the organization might take;

- are free from undue or negative influence from the surrounding community or other agencies.

3. Program characteristics --

- are clearly defined on such dimensions as age/maturity, innovativeness, and overlap with other programs because these characteristics have a bearing on the kinds of procedures the evaluator should select and the kinds of information he or she should provide in order to stimulate use.

With respect to the evaluation factors affecting use, the use potential is likely to increase to the extent that:

1. The evaluation procedures --

- are appropriate to the particular project. A selected procedure must be appropriate as a method for addressing the given question, and also appropriate in the context of the project;

- address the matter of rigor from the dual standpoint of accepted standards of evaluation practice and the users' conception of what constitutes rigor;

- deal with mandated tasks -- funding agency requirements, central office needs, unit level questions -- in a balanced manner so that no single point of view is seen to dominate;

- reflect the viewpoint that no single evaluation model is inherently superior; instead, evaluation is seen as a tool for decision making and the selection of evaluation procedures is guided by the decision-making process.

2. Information dialogue --

- reflects purposeful, guided sharing of ideas between evaluator and users;

- is ongoing, in sufficient amounts to stimulate or maintain user interest in the evaluation, with quality growing out of collegiality and reciprocity.

3. Evaluation substance --

- is relevant from the users' standpoint because it constitutes pertinent answers to the questions they have raised; and

- is specific by focusing its content on the needs and interests of the particular user or user group.

4. Evaluation reporting --

   - is marked by frequent and well-focused provision of information;
   - is timely in that it reflects program chronology and meshes with important events stemming from the program's decision needs;
   - uses whatever variety of presentation formats -- oral, written, statistical/narrative, formal or informal -- that is appropriate to the range of users and their evaluation interests.

Factor Interactions

The preceding observations begin to suggest that factors are likely to interact to affect use. Here we will discuss a few possible interaction patterns to illustrate the kinds of phenomena the administrator might need to consider as he or she organizes the evaluation for use, primarily because many of the factors are beyond the evaluator's control.

For example, to help promote program-level use, the evaluator should address questions relevant to the program, questions of interest to program staff. The extent to which the evaluator is successful will depend, in part, on the various requirements for the evaluation, such as those set by a funding agency, and whether any particular requirement is allowed to dominate. But it will also depend on users' interest in the evaluation and their commitment to applying its findings. However, users' predisposition to make this application can be affected by perceived institutional risk, pressures from the program's community, and the timing at which reports are provided, to mention but a few of the possibilities.

Many of the factors and interactions suggested above may not be amenable to evaluator influence. For example, while the evaluator may commit him- or herself to use, the associated user commitment, which also contributes to the application of results, is properly in the administrator's sphere of influence.

In short, to the extent that the factors mentioned above are subject to influence in a given setting, many are in the administrator's domain and are therefore perhaps more amenable to his or her influence. And this influence, if necessary, can cut across all three factor categories, not only the context/organizational category traditionally associated with administrative responsibility.

We suggested earlier that to be able to use evaluation as a management tool, as a decision-making tool, it is critical that those people who may be affected by a particular deision be involved in the decision-making process. What do the kinds of interactions noted above, then, suggest for the administrator-organizer trying to increase an evaluation's potential as a management tool?

First of all, by very virtue of his or her entry into the use process, the administrator becomes one of the factors influencing use. Continuing CSE research on factors promoting high evaluation utilization has suggested kinds of evaluator behavior which promote use. These behaviors offer clues to the kind of overall demeanor that the administrator-organizer might adopt, first of all, to create an <u>atmosphere</u> conducive to evaluation use.

We recently analyzed several evaluations whose high utilization levels were documented as part of an AERA award to recognize such evaluations (Ruskus & Alkin, 1984). Many of the factors cited tended to confirm those displayed in Figure 1. Five of these factors, each of which was cited as a use-promoting characteristic, suggest how professional style can have a bearing on use. These five factors are level of effort, leadership behavior, user involvement in the evaluation, involvement in implementing recommendations, and commitment to use.

Evaluation users frequently cited the high levels of evaluator effort that contributed to their use of results. From the standpoint of the equity theory of motivation (Adams, 1965; Weick, 1966) it seems likely that when level of evaluator effort is deemed to be high, users demonstrate high utilization.

Leadership was another factor cited as contributing to evaluation use. Social psychologists such as Likert (1961) and Stogdill (1974) suggest that leadership may be seen as originating new ideas; mixing with other participants; acting on behalf of other participants; reducing conflicts; organizing; communicating; recognizing participants' efforts; stimulating participants to achieve; and helping them carry out their duties.

Involving the potential users was another frequently cited factor in the highly utilized evaluations studied. Beyond the idea that users are likely to use information when they play a part in generating the information, participant management theory (Likert, 1967) suggests that supportive relationships, group decision-making, and shared organizational objectives contribute to commitment to carrying out organizational policy and decisions.

Evaluator involvement in implementing recommendations also played a role in the highly utilized evaluations. Such behavior, taking place after the report was generated, can run the gamut from interpreting implications of a recommendation to making concrete suggestions about areas in need of improvement and possible means of promoting such improvement. In this vein, sociological theory on the management of change (e.g., Keen & Scott

Morton, 1978) suggests that evaluation needs to be concerned with introducing the need for change, striking out in the direction of change, and integrating the change into existing frameworks.

User commitment to use, finally, seemed important in all the highly utilized evaluations. On the basis of themes identified in the marketing literature (e.g., Rogers, 1962), the users in the evaluations studied can be typified as "early adopters" who were (or became) highly disposed to try out new ideas.

Now, in several important ways, the administrator trying to promote evaluation use is assuming evaluation-like responsibilities. To the extent that such is the case, then administrator efforts in promoting use, in providing leadership to other potential users and involving them in the evaluation and in implementing its recommendations, will help to stimulate their commitment to use. That commitment, however, is likely to be short-lived unless the evaluation, or at least part of the evaluation effort, is focused on user concerns.

Establishing a User Focus

Previously we outlined some functions of evaluation that would enhance its relevance and use as a management tool. We have suggested that, to the extent an evaluation is to serve multiple audiences -- funding source, program director and staff -- then the needs of these various audiences need to be recognized and kept in proper balance. To be used as a management tool at the local system level at which the evaluation is conducted, the evaluation must be organized so that, in addition to satisfying any other requirements, it identifies, addresses, and answers local-level questions.

We mentioned earlier that while establishing a local focus presents few technical difficulties, it may encounter some attitudinal barriers reflecting conflicting sets of evaluation expectations. For example, let's resume the case of the superintendent responding to the issue of excellence in education. What problems might he or she encounter in attempting to establish local needs as one of the foci of the evaluation? In attempting to get evaluation information that has local management relevance?

Consider the kinds of management concerns we alluded to earlier (Burstein, 1984). In terms of drawing upon evaluation to provide indicators of progress, who is to decide what these indicators are to be? Are there conflicting viewpoints? Can they be reconciled? Can they legitimately differ in various settings?

In regard to making decisions about students, what kinds of decisions are to be made? Are decision needs -- such as comparison versus individual diagnosis -- in competition? Is one kind of measure deemed superior to another? Is that viewpoint based in fact or does it grow from tradition?

With respect to program decision making, does the evaluation have to address multiple audiences? Are there potential conflicts between them? Can the evaluation reconcile external accountability concerns and local concerns about program monitoring and improvement?

With respect to other possible management applications of evaluation -- long-range planning and policy formulation -- it is unlikely that evaluation can, or should, be used at the local level unless it first has local relevance on the other three issues outlined above. Can evaluation come to have that local relevance?

While the reduction in federal control of evaluation that we alluded to earlier is intended to increase local -- SEA and LEA -- responsibility, can we assume that the federal intention is accepted at these levels? Further, can we assume that SEAs and LEAs responding to the possibility of assuming greater control of their evaluations will be in agreement on basic issues such as evaluation purposes, emphases, and procedures? If not, will one point of view dominate and thus reduce the evaluation's relevance for the other? What are some of the issues confronting our school district superintendent who wants to make sure the evaluation will serve his or her management and decision concerns?

One primary issue, as we have already suggested, is that the evaluator's ability to focus on one set of needs may be constrained by other factors in the setting. It may be that one set of needs, requirements, or dominant attitudes causes the evaluator to adopt a certain role and collect certain kinds of information which, in turn, may cause the superintendent, and his or her colleagues, to view the evaluator and the evaluator's work with something less than enthusiasm. We believe that the current situation with respect to evaluation foci requires the superintendent's attention.

First, some of the superintendent's potential evaluation users may not be convinced that changes in federal requirements will actually reduce external supervision and control. That is, while federal supervision may decrease, the state may continue or initiate, or be perceived by LEAs to be continuing or initiating, policies which offset LEA attempts to direct their evaluations toward LEA matters.

Second, some of the superintendent's potential evaluation users may be so thoroughly immersed in the business of administering a previously required test, perhaps a commercially published, norm-referenced test, that they are unresponsive to the possibility of developing a more locally relevant test, perhaps a criterion-referenced test of a particular content area.

Third, decreased federal control is accompanied by decreased federal funds, and with reductions in resources, local school districts may be unable to supply sufficient evaluation expertise across the various content areas they either need to, or would like to, evaluate. An evaluator may need to take responsibility for simultaneous evaluations of Chapter 1 and 2 programs and of other programs such as bilingual education.

Related to the above issues is the emphasis on technical procedures that still exist in the regulations accompanying some programs. This potential problem area, though it is of particular concern in bilingual programs, is seen to some extent in the "sustained effect" provision in the Chapter 1 regulations. Coupled with dwindling local resources and evaluation expertise thinly stretched, a school district facing multiple program evaluation needs of a technical nature may find it difficult to comply.

Further, the "objective measures" mentioned in the Chapter 1 requirements may not be uniformly understood. Owing to historical precedent, the evaluator of an LEA program may believe that a test must be norm-referenced in order to be considered objective and may continue to use this kind of test even when it serves no relevant local purpose. If this is the case,

the evaluation results are likely to be seen by program staff as having little practical value for them.

In another LEA, also concerned with the matter of objectivity, the pros and cons of various kinds of tests may be discussed at length without ever resolving the matter to the satisfaction of all potential users. So much time may be spent on the debate over the test question that the evaluator has little time left for planning and conducting an evaluation around more important issues.

In districts meeting the kinds of problems outlined above, that is, where testing issues are difficult to resolve, where there is limited expertise to balance technical adequacy and local relevance, there is likely to be some negative effect on factors promoting use: anxieties and sense of risk may dominate; program staff may believe their questions and concerns are receiving insufficient attention; the evaluator's credibility is likely to suffer; commitment to evaluation use and perception of usefulness will decrease.

Problems such as those outlined here do appear to warrant the attention of the superintendent in the case we are using for illustration. For example, we found in a recent exploratory study (Burry, 1984) of directors and staff members of school district research and evaluation units a general agreement concerning the kinds of problem areas described above.

For example, there was a general concern about the force of historical precedent. Because of earlier state and district preferences (which grew out of earlier federal emphases), district evaluators were still required to administer norm-referenced tests for reporting purposes. But informa-

tion produced by such tests was of limited use to them in carrying out their responsibilities. Further, the evaluators also agreed that these tests did not address teachers' interests or provide the kinds of information needed to monitor and adjust classroom instruction. Most evaluators felt that a good deal of their limited time and resources were given over to generating information which was state required in response to Chapter 1 regulations.

A few districts were trying to distill some locally-useful information from a norm-referenced test. But the evaluators felt that too much time was involved in having the tests scored and then returned for analysis and interpretation so that they would be of some use to curriculum specialists and classroom teachers. Consequently, these efforts detracted from the time they needed to generate information more specific to instructional needs.

Several districts, in addition to administering and reporting the results of a norm-referenced test, used district-developed objectives-based measures and/or the tests accompanying curriculum materials. (See Burry, et al, 1982, for a discussion of how widespread this practice is, as well as some of its implications.) While districts found the information from these additional tests instructionally useful, they asked why they had to conduct what was, in effect, a parallel evaluation. That is, for the reasons suggested above, they felt they should use a "respectable" norm-referenced test for external reporting purposes; given the limitations of the information provided by such tests, however, they felt at the same time that they had to resort to other devices for locally useful information.

The question raised by the evaluators was as follows: if such district-level efforts are necessary to produce formative data for district purposes, cannot these efforts also be used to satisfy external (summative) requirements? They asked this question even though they were aware that federal requirements impose no particular kind of test; perhaps they (or their superiors) continue to believe that anything other than a standardized test is unacceptable.

This situation created a double dilemma. On the one hand, the evaluators realized that, to be locally useful, the evaluation should provide different types of information for different groups of users and that such an effort takes time: time to identify the needs and questions of various potential uses, time to develop or select appropriate tests or design other data collection procedures, time to win user support for the evaluation. On the other hand, the felt need to run a separate evaluation for external purposes also takes time; time that might be better spent addressing local questions and needs.

Finally, while some districts were attempting to increase their evaluation's local relevance, the evaluators stated that they had a hard time convincing other staff, such as resource specialists and teachers, to become involved in the conduct of the evaluation. Because of what they knew or believed about previous evaluations, personnel were hesitant about raising their own evaluation questions, reluctant to participate in the process of devising ways to answer them, and unwilling to believe that the evaluator would want to help them in the task of carrying out their day-to-day responsibilities.

In essence, then, then evaluators we interviewed are describing how some potential evaluation users, the kinds of people whose support will be needed in order for the superintendent to be able to use evaluation to help manage his or her district's operations, may act with respect to the evaluation, its conduct, and its uses. They suggest some of the ways the factors we introduced earlier may interact to affect evaluation use. They suggest some likely factor patterns that any administrator will need to consider as he or she begins the task of organizing for evaluation use.

ADMINISTRATIVE ORGANIZING TO PROMOTE USE

Figure 2, which is excerpted from the handbook we described earlier (Alkin et al, 1985, in press), places the factors that we believe are central to use in most evaluation contexts into a pattern which will facilitate organizing for use. In this pattern the factors are grouped to reflect stages in the process of planning for and conducting an evaluation to maximize its use potential.

The factors and their potential influence on use should be considered from the standpoint of the intended users/uses, gathering information that will help the evaluator focus on these users/uses, and gathering information that will help the administrator-organizer ensure that factors in the setting do not impede that focus.

The administrator-organizer may want the evaluation to provide information that he or she, and other potential users, can apply to one or several decision concerns -- broad monitoring issues, and/or student or program decisions, and/or planning and policy needs. The major tasks, then, are to decide on the users/uses on which the evaluation should focus,

FIGURE 2:

Factor Pattern For Evaluation Use

A. Setting the Stage

Pre-existing evaluation bounds
User identity
Program characteristics
Intra-organizational features
External features

↓
↓

B. Identifying/Organizing the Participants

User interest in evaluation
User commitment to use
Evaluator characteristics • background/identity • commitment to use • willingness to involve user in evaluation • choice of role • political sensitivity • credibility
Evaluation procedures—plan
User professional style(s)

↓
↓

C. Operationalizing the Interactive Process

Evaluation procedures—execution
Substance of evaluation information
Evaluator commitment to use
Information dialogue—formative
User information processing preferences

↓
↓

D. Adding the Finishing Touches

Evaluation reporting
Evaluator characteristics (selected)
Information dialogue—summative
User commitment to use

↓↓

Excerpted from Alkin et al. Organizing for Evaluation Use: A Handbook for Administrators. Los Angeles: UCLA Center for the Study of Evaluation, 1983.

asking questions in light of the situational factors that might influence use, and using the answers to these questions to help guide both the evaluation process and the administrative tactics devised to help ensure that the evaluation stays on target.

The administrator who assumes the use-organizing responsibility can use the factor pattern in Figure 2 -- with any appropriate emphasis, addition, or deletion of factors given the particular context -- while he or she considers the program, its evaluation, the setting in which it takes place, and the intended uses. This consideration involves anticipating the effects that a particular evaluation direction, once taken, is likely to achieve. It consists of asking oneself a series of questions in light of the listed factors with the intention of determining how the program embodies each factor; that is, deciding whether that embodiment is likely to have a positive, neutral, or negative effect on the intended uses, and then devising strategies to strengthen or maintain positive effects while minimizing negative effects. These strategies may then be implemented by the administrator and/or the evaluator or some other potential user.

For example, assume that the superintendent in our case illustration goes through the process described above and, using the scheme suggested in Figure 2, asks him- or herself the following question about the first organizing issue -- setting the stage: "As part of the intra-organizational features, is there any perceived staff risk that might hinder my using the evaluation to plan future instructional offerings in response to changing student enrollment patterns? After due deliberation and discussion with potential staff users, the superintendent discovers that some

teachers feel that the evaluation poses a risk to them and that the degree of risk is likely to outweigh possible benefits. Other teachers either feel that the evaluation presents little or no risk or that benefits to be accrued outweigh any possible risks.

Now, this user group's reluctance to accept and apply evaluation findings is one manifestation of a two-part problem. First, their acceptance and application of information may be important to a larger decision area, one that may need to be made consensually by all users. Given the reluctance of one segment of the decision-making group to participate, it may be that the resultant decision concern is never fully resolved.

Second, those with the sense of risk may advance beyond reluctance to participate to outright attempts to convince others of potential dangers. If they are successful, then initially receptive users may later opt to remove themselves from the evaluation effort and, further, may attempt to thwart the entire effort.

In such a situation, the superintendent organizing for evaluation use would need to ask other questions in order to determine: the reason for the sense of risk on the part of one user group; whether or not that perception is justified; the extent to which the group in question may attempt to convince others of the imminent risk; the likelihood of success. He or she would then need to devise appropriate strategies given the answers to the preceding questions.

For example, it may be that the sense of risk is unjustified or has become magnified, perhaps on the basis of some previous evaluation experience. In this situation, the superintendent would need to convince the

hesitant group that this perception is unjustified so that the evaluator's credibility does not suffer and the necessary user group involvement is achieved.

To the extent that the superintendent him- or herself encounters difficulty in minimizing sense of risk, then it may be possible to enlist trusted and respected staff members from among the more receptive users to help convince their colleagues that, in this particular setting, the risk factor is unwarranted and that participation in the use process is justified and important to the larger institution.

Keeping the above potentially inhibiting factor example in mind, and the kinds of question-raising process and associated strategy formulation the superintendent considered, we will now suggest a few possible questions, and how they might be addressed, for factors in each of the four stages in the use process, as depicted in Figure 2. These questions are intended to guide administrative organizing for evaluation use, and their answers, as with those of all the factors displayed, should inform the administrator's selection of strategies to build use into the evaluation.

Setting the Stage

Setting the stage involves determining, before the evaluation planning process begins, the kinds of factor interactions likely to affect use in a given setting. While these factors may be set to some extent, they are not necessarily "givens." Note in Figure 2 that this determination considers possible effects stemming from the pre-existing evaluation bounds, the potential users identified, program characteristics, and intra-organizational and external features.

Questions that the administrator-organizer might raise here could include, for example:

- Who are the intended users of the evalution information?
- Are the pre-existing evaluation bounds such that there may be potential conflict, real or perceived, between program expectations and other requirements?
- How is the program best characterized with respect to its maturity, innovativeness, and overlap with other programs?

Now, let's narrow the focus a little and add some context before we go any further. Suppose that our district superintendent had a programmatic concern to resolve. In the district, enrollment in math classes in some high schools has been dropping off sharply in the last two or three years; in others, math enrollment is staying relatively constant, even increasing a little. Board and parental concern with students' technical literacy is on the rise. The superintendent would like the required evaluation of the math program to help explain the different enrollment patterns and discuss what might be done about it.

Who might the interested stakeholders, and hence potential evaluation users, be? At the least: the funding agency and the board; building principals, math department chairs, math teachers; parents and students; district office math specialists.

What might be some possible conflicts among these users? Should the superintendent consider this question in light of program maturity/ innovativeness to help illuminate the possibility of conflicting expectations? Very definitely. Consider the following:

- The board and the funding agency expect the district to continue reporting the math program results in terms of student scores on

the norm-referenced test that has been in use in the district for the last five years; parents are used to seeing these results discussed in the local newspaper; some parents want these scores to go up; others are asking why their children did not have to take the test.

- The superintendent, after meeting with math teachers in both reduced- and maintained/increased enrollment schools begins to get the distinct impression that the district's "math program" does not look the same across all schools. In some schools, especially those with high enrollment levels, innovativeness seems to be the defining feature. But innovativeness seems to differ in these schools. In a school or two, teachers rely heavily on tests they have developed themselves to make decisions about instruction; they treat the norm-referenced test as something that has little relevance for them. In some of the low-enrollment schools, teachers stress the importance of the norm-referenced test to their students and emphasize its content in their instruction. A few teachers in each kind of school do not fit the general pattern.
- The superintendent would like to explore these differences in the next year's program evaluation. First, he would like to have the norm-referenced test requirement waived for that year. In its place, he would like to conduct intensive observational studies of high school classroom math practice to find out if different teacher approaches to math instruction and/or math assessment might help explain different enrollment patterns.

- Will the funding agency accept this plan? Will the school board? How about teachers and parents? If the norm-referenced test cannot be waived, could it instead be administered on some sample basis fo the coming year? Might the evaluation apply and analyze the norm-referenced test and also conduct the intensive observational study? Will resources permit this?
- If the evaluation is permitted to emphasize the observational component, and discovers that a certain instructional approach seems to be more effective than others in attracting and maintaining student enrollment, how might the board react to this finding? How might teachers respond? How might it be received by the funding agency?

While questions such as these might all be considered in tne initial organizing step -- setting the stage -- some of them may not be resolved until later on in the use process.

Identifying/Organizing the Participants

After setting the stage for evaluation planning has taken place, a series of questions which amplify user characteristics such as interest in the evaluation and commitment to its use, as well as questions reflecting relevant evaluator characteristics, should be raised. This process should result in the formulation of the evaluator's role and the evaluation procedures, carefully matched to users' interests, expectations, and professional styles, which will be used.

Among the questions that ought to be considered at this stage are:

- Are the intended users committed to use and, if so, is their commitment rhetorical or real?

- What do the intended users expect from the evaluation; are these expectations likely to affect their desire or ability to apply information?

- What would be the most appropriate role for the evaluator to take with respect to the program, and will the evaluator be willing and able to assume this role?

- What kind of evaluation procedures will provide the best match with users' professional styles?

Let's pause for a contextual breather again. Continuing the superintendent's scenario, what are some of the issues of concern in this second organizing stage?

Let's start by thinking about the evaluator for a moment. Though a highly-skilled professional, is there anything in his or her personal comportment that would cause anxiety among teachers whose classrooms were being observed? If so, would the evaluator accept a carefully phrased suggestion about classroom entry?

Now let's take up a possible politico/methodological problem. Let's assume the observational component was sanctioned by the funding agency. Let's also assume that the district needs all the resources it can get to continue its math offerings. Let's also assume that a well-defined classroom practice did seem to account for student interest in math. Should the superintendent recommend that this particular approach implemented district-wide, is it likely that the agency, although they sanctioned the observational study, would find its results to be credible? Is it possible that their understanding of observational data would affect their refunding decision? Would the board's reaction to the superintendent's recommendation be influenced by their financial concerns? How might teachers whose practice will be affected respond to the recommended change?

These kinds of questions need to be resolved before evaluation procedures are selected and put into operation.

Operationalizing the Interactive Process

Up to this point, the administrator-organizer has been anticipating future evaluation actions and effects; in this third stage the carefully planned evaluation procedures are put into effect. The central factor in this group, execution of evaluation procedures, will temper all other factors grouped here.

Among the questions that should be considered are:

- What is the most effective data-collection schedule, and are there any possible impediments to this schedule?
- Do any of the proposed procedures require any special arrangements and, if so, with whom?
- For each intended user, what particular kinds of information and in what kinds of format will be deemed relevant?
- What kinds of dialogue, via what techniques, will best match users' routine information processing styles?

While the evaluation process is underway, the superintendent in our scenario, or any other organizer, would constantly monitor the process. He or she would ensure that the evaluation is proceeding in light of how previously raised questions were answered; determine if any unanticipated factor influence is beginning to emerge; determine if an expected influence is less than anticipated, and if resources might be safely shifted to another factor of concern.

Adding the Finishing Touches

This activity is the final phase in maximizing the potential for evaluation use. The group of factors of interest here represents that

point in the evaluation process where most, or all, of the evaluation information has actually been collected. That information must now be communicated in such a way that the designated users will actually apply the information.

Among the questions the administrator-organizer should consider here are:

- What combination of written and oral reporting will most enhance use of information?
- At what time(s) should these reports be provided?
- After the reports are provided, will any final arguments be needed to convince users to act on the information?

Finally, note that the answers arrived at in any one stage will influence questions and organizing strategies stemming from a subsequent stage. Further, the process is cyclical and permits specifications proposed at an earlier stage to be modified (e.g., stressing/de-emphasizing one of the evaluation questions) in light of subsequent planning, conduct, and emerging receptivity toward the evaluation and its use.

CONCLUSION

We have suggested here that evaluation can serve a variety of educational management questions and outlined some of the question areas. We have stressed that an administrator's ability to use evaluation as a management tool depends not only upon his or her own perception of evaluation but also on the perceptions of other potential evaluation users in the system. Evaluation's contribution as a management tool is affected by the degree to which evaluation comes to be accepted and used throughout the various levels of the system.

Research has uncovered a variety of factors influencing an evaluation's use potential. These factors reflect human considerations -- such as people's attitudes toward and expectations for the evaluation; context considerations -- such as an evaluation's requirements within a particular setting; and evaluation considerations -- such as procedures used and means of communicating information. The research demonstrates that an evaluation's use potential, and therefore its application as a management tool, can be greatly enhanced if someone takes responsibility for organizing the evaluation to meet specified needs, for particular users, in light of the factors operating in the given setting. The research also suggests that a program-level administrator is in the most strategic position to assume this responsibility.

It seems evident that administrators and evaluators must come to know more about each other's operational needs and viewpoints. To the extent that administrators and evaluators share responsibility for setting an evaluation's foci and purposes, and ensure that the evaluation addresses these purposes, the evaluation's decision-making power and relevance are increased.

When a program evaluation is being considered, therefore, the administrator needs to decide on its various audiences, determine their questions and information needs, and anticipate the uses they are likely to make of the information. At the same time, the administrator needs to consider the factors existing in the given setting that are likely to influence these uses. The framework we suggest can be applied to organize the evaluation so as to have a high potential for meeting the intended uses in light of various factor influences.

REFERENCES

Adams, J.S. Injustice in social exchange. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 2). New York: Academic Press, 1965.

Alkin, M.C., Kosecoff, J., Fitzgibbon, C., & Seligman, R. Evaluation and decision making: The Title VII experience. CSE Monograph No. 4. Los Angeles: UCLA Center for the Study of Evaluation, 1974.

Alkin, M.C. Evaluation: Who needs it? Who cares? Studies in Educational Evaluation, 1975, 1, 202-212.

Alkin, M.C., Daillak, R.H., & White, P. Using evaluations: Does evaluation make a difference? Beverly Hills, CA: Sage Publications, 1979.

Alkin, M.C., Jacobson, P., Burry, J., Ruskus, J., White, P., & Kent, L. A handbook for evaluation decision makers. Beverly Hills, CA: Sage Publications, 1985 (in press).

Boyer, E.L. High school. New York: Harper & Row, 1983.

Burry, J. Synthesis of the literature on evaluation use. Los Angeles: UCLA Center for the Study of Evaluation, 1983.

Burry, J. Federal requirements affecting evaluation use. CSE Report No. 222. Los Angeles: UCLA Center for the Study of Evaluation, 1984.

Burstein, L. The use of existing data bases in program evaluation and school improvement. CSE Report No. 229. Los Angeles: UCLA Center for the Study of Evaluation, 1984.

Cohen, D.K., & Garet, M. Reforming educational policy with applied social research. Harvard Educational Review, 1975, 45(1), 17-43.

Council of Chief State School Officers Consolidation Evaluation Task Force. Memorandum: Final report and recommendations. Council of Chief State School Officers, Washington, D.C.: 1982.

Daillak, R.H. A field study of evaluators at work. Los Angeles: CSE Report No. 154, UCLA Center for the Study of Evaluation, 1980.

Goodlad, J.I. A place called school. New York: McGraw-Hill, 1983.

Keen, P.G.W., & Scott Morton, M.S. Decision support systems. Reading, MA: Addison-Wesley, 1978.

Likert, R. New patterns in management. New York: McGraw-Hill, 1961.

Likert, R. The human organization. New York: McGraw-Hill, 1967.

Mann, J. The outcome of evaluative research. In C.H. Weiss (Ed.), Evaluating action programs: Readings in social action and education. Boston: Allyn and Bacon, 1972.

National Commission on Excellence in Education. A nation at risk: The importance of educational reform. Washington, D.C.: U.S. Government Printing Office, 1983. No. 065-000-00172-2.

Patton, M.Q., Grimes, P.S., Guthrie, K.M., Brennan, N.J., French, D.B., & Blyth, D.A. In search of impact: An analysis of utilization of federal health evaluation research. In T.D. Cook and associates (Eds.), Evaluation Studies Review Annual, Vol. 3. Beverly Hills, CA: Sage Publications, 1978. Also, Minneapolis, University of Minnesota, 1975.

Rogers, E.M. Diffusion of innovations. New York: The Free Press, 1962.

Rossi, P.H. Booby traps and pitfalls in the evaluation of social action programs. In C.H. Weiss (Ed.), Evaluating action programs: Readings in social action and education. Boston, Allyn and Bacon, 1972.

Ruskus, J.A., & Alkin, M.C. Factors common to high-utilization evaluations. CSE Report No. 240. Los Angeles: UCLA Center for the Study of Evaluation.

Stecher, B.M., Alkin, M.C., & Flesher, G. Patterns of information use in school level decision making. Los Angeles: CSE Report No. 160, UCLA Center for the Study of Evaluation, 1981.

Stogdill, R. Handbook of leadership. New York: The Free Press, 1974.

Weick, K.E. The concept of equity in the perception of pay. Administrative Science Quarterly, 1966, 11, 414-439.

DELIVERABLE -- NOVEMBER 1984

EVALUATION PRODUCTIVITY PROJECT

Factors Common to High-Utilization Evaluations

Marvin C. Alkin
Project Director

Grant Number
NIE-G-84-0112, P2

CENTER FOR THE STUDY OF EVALUATION
Graduate School of Education
University of California, Los Angeles

Factors Common to High-Utilization Evaluations

Report Prepared by:

Joan A. Ruskus
Marvin C. Alkin

Center for the Study of Evaluation
University of California, Los Angeles

November 1984

FACTORS COMMON TO HIGH-UTILIZATION EVALUATIONS

Joan A. Ruskus
Marvin C. Alkin

CSE Report No. 240
1984

CENTER FOR THE STUDY OF EVALUATION
Graduate School of Education
University of California, Los Angeles

The project presented or reported herein was performed pursuant to a grant from the National Institute of Education, Department of Education. However, the opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement by the National Institute of Education should be inferred.

TABLE OF CONTENTS

This paper reports on the factors that characterize high-utilization evaluations. It is based on materials submitted to an AERA Division H competition that was instigated and organized by the Evaluation Productivity Project of UCLA's Center for the Study of Evaluation (CSE). This project, which is funded by the National Institute of Education (NIE), has been at the forefront of research on evaluation utilization.

The paper is organized into three sections. The first section outlines the background of the study: the purposes and procedures of the Division H competition, and the conceptual framework used in analyzing the data. The second section describes the analytic methods, summarizes the results of the analysis, and specifies the six factors that seem to distinguish evaluations whose results are used by decision-makers. The final section profiles the evaluators themselves. The data are viewed and interpreted from a multidisciplinary perspective which draws upon theories from psychology, sociology, organizational behavior, management, and marketing.

BACKGROUND

The primary purpose of the Division H competition was to recognize and reward excellence in promoting evaluation use. Its

secondary purposes were to make both evaluators and potential users more use-oriented and to encourage practitioners to engage in those kinds of behaviors that seem to promote evaluation utilization.

Competition Procedures

Evaluators entering the Division H competition were required to submit an application form (see Appendix A) in which they discussed those features of the evaluation that, in their judgement, had contributed significantly to its ultimate utilization. Evaluation utilization was verified by means of materials submitted by evaluation users, who were asked to indicate the extent of use and to specify those aspects of the evaluator's performance that had contributed to use (see Verification Form in Appendix B). Over 30 users (administrators of the target programs) submitted such materials.

Twelve evaluation studies were entered into the competition. While all of them dealt with educational programs, the majority (seven) were directed to public school programs at the district level. Of the remaining five, three addressed university progr ns; one, a statewide educational program; and one, a vocational education program for a special-needs population. Three of the twelve studies focused specifically on teacher effectiveness. Most of the evaluators had been retained as private consultants, although four were employees of the organizations housing the program being evaluated.

The review process involved over twenty judges, all of them Division H members, who were selected either because they had contributed to the literature on evaluation use or because they were among the leading evaluation practitioners. Each evaluation study, along with the documents submitted as supporting evidence, was reviewed by three judges, who rated applicants on four dimensions: (1) conception of evaluation use, (2) extent of evaluation use, (3) degree of direct link between the evaluator's efforts and subsequent use, and (4) uniqueness/creativity of the evaluator's efforts to promote use (see Reviewer Rating Form in Appendix C). In addition, to ensure against rater bias, each rater ranked the applicants whom she or he had reviewed. Ratings were then totaled, and from these total scores, four finalists were identified; a fifth finalist who had received relatively high rankings was added to this group. The competition winner, and two honorable mentions, were chosen by a subset of raters at a full discussion session of the 1984 AERA convention.

Conceptual Framework

According to the definition employed in this paper, which was developed by Alkin (1975) and Patton et al. (1978), the term "use" is not limited to the direct application of evaluation findings to a specific decision but rather refers to the broader supporting role that evaluation findings often play in the complex, ongoing process of decision-making. As Alkin (1982) points out, the best definition for utilization is one that recognizes the gradual, incremental influence of evaluation.

This view of utilization was further developed by CSE through a series of evaluation case studies (Alkin, Daillak, & White, 1979), an evaluator field study (Daillak, 1980), and an administrator user survey (Stecher, Alkin, & Flesher, 1981).

On the basis of these empirical data, Alkin and his colleagues (1985, in press) have developed a "utilization framework" that classifies the factors affecting evaluation use into three interrelated categories: human factors (evaluator and user characteristics); context factors (pre-existing evaluation bounds, organizational features, and project characteristics); and evaluation factors (procedures, information dialogue, substance of information, and reporting). Within each of these factors, specific elements that influence utilization have been identified. For example, critical evaluator characteristics include commitment to use, rapport with users, and credibility. Critical user characteristics include interest in the evaluation, commitment to use, and information-processing preferences. Pre-existing evaluation bounds, a context factor, encompasses written requirements, contractual obligations, and fiscal constraints. Evaluation reporting, an evaluation factor, includes the frequency and timing of reports and the mix of statistical and narrative data in reports. Many other elements are subsumed under the various factors (for a complete listing, see Alkin et al. (1985, in press)).

ANALYSIS OF APPLICANT DATA

The comments of users and evaluators were analyzed by reading through the application and verification materials and abstracting all statements pertaining to use. These statements were then classified according to the three categories of utilization factors: human, context, and evaluation. (User data are summarized in Appendix D, and evaluator data in Appendix E.) In some instances, the results confirmed the importance of the factors already specified in the utilization framework described above. In other cases, the results suggested that certain factors not explicitly identified by previous research need to be further explored, with a view to elaborating the framework.

Almost half of both the users' and the evaluators' comments pertained to human factors, and nearly all of the remainder referred to evaluation factors. Context factors were rarely mentioned, probably because the competition emphasized the evaluator's contribution to enhanced utilization.

The most frequently cited evaluator characteristics were choice of role and willingness to involve users in the evaluation. The users' interpretation of choice of role is a unique one and will be discussed in more detail below. Also important were the evaluator's credibility and rapport with users.

The evaluation factors most frequently mentioned were procedures and reporting. Both users and evaluators recognized the importance of sound methodology, user-oriented designs, and

follow-up procedures. Of the several reporting features cited as contributing to evaluation utilization, a good balance between statistical and narrative data was seen as most relevant.

While the factors cited by evaluators tended to confirm the utilization framework already developed, users' comments suggested some additional factors that may influence utilization. Because of their significance from the users' perspective, six factors merit further discussion for the insight they give into utilization:

- o Level of evaluator effort
- o Leadership behavior
- o User involvement
- o Evaluation reporting
- o Evaluator involvement in implementing recommendations
- o User commitment to use

Level of Evaluator Effort

Users frequently said that the exceptional level of effort which the evaluator put into the evaluation contributed to their utilization of the results. The following comments illustrate this point:

> The evaluator (E) has been willing to expend time and energy beyond the typical work day to do the necessary work.
>
> The sustained efforts of E over the last three years have insured that there is a broad commitment to this evaluation.
>
> E's drive to see the project through and willingness to reach out and help contributed to making the evaluation useful.

Clearly, users appreciate the effort put forth by the evaluator, not only because it results in a higher-quality evaluation and a better evaluation report but also because it signifies a commitment to the evaluation and a concern about its outcomes. It would seem, then, that when the evaluator manifests a high level of energy, the likelihood of evaluation use increases.

The equity theory of motivation (Adams, 1965; Weick, 1966) suggests that something besides simple appreciation is at work here. This social comparison theory views human relationships as transactions involving inputs and outputs. In negotiating exchanges in the work environment, people seek a balance between inputs and outputs. According to this theory, the evaluator's level of effort constitutes an input, and utilization of the recommendations advocated by the evaluator constitutes an output. Therefore, when level of evaluator effort is judged to be high, users are motivated to demonstrate high utilization.

Leadership Behavior

A number of the evaluator characteristics noted by users as contributing to utilization can be characterized as leadership behaviors. And while some of these behaviors are already represented in the utilization framework (i.e., choice of role), it seems reasonable to restructure the framework slightly by subsuming these elements under the "leadership behavior" rubric. Such an alternative conceptualization may provide additional

insights into the dynamics of evaluation utilization.

A sizable portion of the management literature deals with the definition of leadership and of what constitutes leadership behavior (see Fiedler, 1967; Hollander, 1978; House & Baetz, 1979; Katz & Kahn, 1978; Vroom & Yetten, 1973). Perhaps the most thorough of these discussions comes from social psychologists like Rensis Likert (1961) and Ralph Stogdill (1974), who have identified nine dimensions of leadership: initiation (originates new ideas and new practices), membership (mixes with the group), representation (acts in behalf of the group), integration (reduces conflicts between members), organization (structures the work of members), communication (provides information to members and shows awareness of affairs pertaining to the group), recognition (expresses approval of group members), production (sets levels of effort for greater achievement), and consideration (helps members and explains procedures). These dimensions were epitomized in the actions of evaluators, as reported in a number of statements made by users:

> E's information caused us to initiate actions that might never have been taken without his impetus. (Initiation)
>
> E has the ability to develop rapport and trust with key program personnel. This trust is reflected in the way those she interviews and surveys open up to her with total confidence. (Membership)
>
> The information E has been able to provide for us was just the documentation we needed to justify program decisions. (Representation)
>
> E worked with program personnel and upper administration to resolve conflicts, facilitating a process where all points of view were heard. (Integration)

> The evaluation report contained practical management plans based on the findings. (Organization)
>
> E's evaluation report was primarily useful in telling our story. He was capable of making explicit our commonly held (but taken for granted) assumptions about what we do and why we do it. (Communication)
>
> E has provided the impetus for the development of goals and objectives to strengthen and improve programs. (Production)
>
> E provided the staff directly connected to the program . . . with much positive feedback. (Recognition)
>
> E explained what could and could not be measured, what instruments would be used and why. (Consideration)

User Involvement

The evaluator's willingness to involve users -- already identified in the utilization framework as an important evaluator characteristic -- deserves further examination because of the frequency with which it was mentioned, both by users and by evaluators, as a determinant of utilization. Users expressed the importance of their own involvement as follows:

> As superintendent, I was involved in the total process enough to follow through on the report.
>
> E developed "ownership" from the beginning of the evaluation process so that participants actually thought of it as "their" evaluation.
>
> During planning, E reached out to program participants for their involvement, assuring broad-based ownership of the entire evaluative process.

It makes intuitive sense that users will be more likely to use information that they asked for or played a part in generating, but is there more to it than that? Participant management theory (Likert, 1967) maintains that user involvement

is a critical component of effective management. Participant management is characterized by supportive relationships, group decision-making, group methods of supervision, and organizational objectives that reflect the needs and desires of all shareholders in the organization. According to its proponents, not only does this method of management make for better decisions, but it also guarantees that people will be more committed to carrying out these decisions. Evaluators who strive to build rapport with users, who involve users in the design of the evaluation and in data collection, and who consider the information needs of all users are following the principles of participant management, whether they are aware of it or not. While this theory validates the inclusion of users in decision-making, it does not really explain why this technique is so powerful.

One possible explanation is that involving users in the process of evaluation changes their attitudes about what the program should be and how it should operate. This new attitude is reflected in the recommendations that emerge from the evaluation. Thus, acting on the recommendations is consistent with their attitudes, whereas failure to use evaluation findings is likely to result in a state of cognitive dissonance (Festinger, 1957) which must eventually be resolved. User involvement is really a technique for attitude change that, once accomplished, motivates behavioral change.

Evaluation Reporting

While the utilization framework identifies several elements of evaluation reporting -- notably, content of reports (Alkin, Daillak, & White, 1979) and style (Brown, Braskamp, & Newman, 1978) -- as critical to utilization, users in the Division H competition noted another critical reporting element: thoroughness. This element is implicit in several of the utilization framework factors, being most closely related to substance of evaluation information, which includes information specificity. Nonetheless, users' views on the importance of the evaluator's thoroughness in reporting merit further examination. The following comments are illustrative of these views:

> The depth and thoroughness of the first-year evaluation enabled me to spot quickly and accurately the problems with the project so that they could be remedied the next year.
>
> The thoroughness and detail on what went well, what did not, and why, gave me a wealth of information and insight which I could utilize, in detail.

This preference for detailed information, rather than just summaries and generalities, underscores an important point. Evaluators must demonstrate their thoroughness, but at the same time, they must know the individual users well enough to tailor their presentation of information to the cognitive styles and preferences of the users. The users quoted above obviously prefer that full and precise data be included within the report. Other users, with less of a need for fine detail, prefer a summary of evaluation highlights. Evaluators must satisfy both groups by providing comprehensive but readable reports along with

concise executive summaries. They should at all times preserve the impression that comprehensive data are available as back-up.

Style and format are two dimensions of evaluation reporting specifically included within the utilization framework. Report style is the manner in which the evaluation "message" (i.e., the findings) is executed. Marketers have identified several successful message execution styles: slice-of-life, mood or image, technical expertise, scientific evidence, and testimonial evidence (Kotler, 1980). The users in this sample tended to prefer styles that were literate, conversational, warm, and down-to-earth rather than overly formal and jargonistic. One user commented:

> E had the ability to translate facts and figures and charts that would normally baffle the layperson's mind into interesting material.

Another noted:

> All of the numbers, statistics, data were surrounded by literate prose with appropriate quotations from Alexis de Tocqueville.

Format, another component of message execution, can also make a difference in message impact. Format elements include the size and dimensions of the report, and its use of color, illustration, and other graphic elements. The importance of format was acknowledged by many users. The following comment is typical:

> E's reports have been exceptionally well accepted. Her formatting and special touches, such as attractive customized covers and pertinent cartoons, make the reading enjoyable and interesting.

Evaluator Involvement in Implementing Recommendations

Users' comments also focused on the evaluator's active role in the actual, hands-on implementation of the study recommendations. These evaluators went well beyond the conventional role of encouraging utilization -- they made certain of it. The following statements illustrate this proactive role:

> E took leadership and responsibility during discussions and in preparing proposals and plans for the Board of Education.
>
> E has been remarkably successful in working with academic units throughout the campus in institutionalizing outcome information on an ongoing basis.
>
> E has held workshops for teachers and staff and is always available to help solve problems.

Clearly, evaluator involvement in implementing recommendations may take many forms. In some instances, the evaluator conducts workshops on the findings as a step in potential implementation. Or the evaluator may interpret the action implications of particular recommendations. A similar phenomenon occurred in a study of Title I evaluators (Alkin, Stecher, & Geiger, 1982): One evaluator helped to attain utilization by "suggesting changes in the program, planning next year's workshop, developing a dissemination plan, and creating a meaningful attendance policy" (p. 2). In essence, this evaluator trained school district personnel to use data in making school decisions by giving them practice with the process.

Sociological theory on the management of change provides insight into how post-report evaluator behavior influences use.

Keen and Scott Morton (1978) hold that the change process consists of three stages: unfreezing, which serves to disturb the current stable equilibrium and introduces the need for change; moving, which involves striking out in a new direction; and refreezing, which requires integrating the change into existing behavioral frameworks to recreate a whole, natural entity. Management theorists (Katz & Kahn, 1978) claim that, while many change agents are successful in the first and second stages, they fail to realize the critical importance of the last stage. The evaluation process can be viewed as the first stage, the evaluation recommendations as the second stage, and the implementation of recommendations as the third stage. It is no wonder that evaluator involvement in this final stage of the change process characterizes high-use evaluations.

User Commitment to Use

User commitment to use, which was identified as an important user characteristic in all of our prior work, also turns out to be important in this study. Although users themselves did not often cite this factor -- perhaps because they were asked to focus on the evaluator in their documentation -- evaluators referred to it frequently, as the following comments show:

> The people involved were ready for the evaluation and the changes.
>
> The leadership of the Chancellor and the Provost . . . was one of the unique factors that contributed to my success in promoting the use of evaluation data.

To some extent, commitment to use is a personality characteristic of users and thus is already established by the time the evaluator enters the scene. The marketing literature (Rogers, 1962) places consumers on a continuum, depending upon their willingness to adopt new products or new ideas: early adopters, early majority, late majority, and laggards. The users in this sample were clearly early adopters. Given their predisposition to try new ideas, they may have utilized any evaluation findings they were given. On the other hand, our earlier research (Alkin, Daillak, & White, 1979) shows that the actions of the evaluator can positively influence the users' predisposition to use.

PROFILE OF A HIGH-UTILIZATION EVALUATOR

This study provides some insight into the characteristics and behavior of those evaluators whose work has a high probability of being utilized. High-utilization evaluators have personality traits -- concern, warmth, patience, integrity, openness, tact, willingness to listen -- that make them attractive to users from the outset. But they are also true leaders. Motivated by the desire to see their work utilized, they systematically enlist the participation of all potential users. They often violate the conventional image of the evaluation consultant by becoming involved in program operations and program improvement. In all of their leadership activities, however, these evaluators remain conscious of users' concerns and problems. They frequently assume the role of teacher, coaching

program personnel in evaluation techniques. Their enthusiasm is not lost on users, who perceive such evaluators as investing a great deal of effort in the evaluation process and as being truly concerned about improving the program being evaluated.

When preparing the evaluation report, high-utilization evaluators are careful to respond to the cognitive styles of the various users and to translate quantitative data into interesting information about the program. Once the evaluation report has been delivered, they continue their active involvement in the program by providing specific guidance in the implementation of the recommendations.

REFERENCES

Adams J.S. (1965). Injustice in social exchange. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 2). New York: Academic Press.

Alkin M.C. (1982). Parameters of evaluation utilization/use. Studies in Educational Evaluation, 8, 153-155.

Alkin M.C. (1975). Evaluation: Who needs it? Who cares? Studies in Educational Evaluation, 1, 201-212.

Alkin M.C., Daillak R.H., & White P. (1979). Using evaluations: Does evaluation make a difference? Beverly Hills, CA: Sage Publications.

Alkin M.C., Jacobson P., Burry J., Ruskus J., White P., & Kent L. (1985, in press). Handbook for evaluation decision-makers. Beverly Hills, CA: Sage Publications.

Alkin M.C., Stecher B.M., & Geiger F.L. (1982). Title I evaluation: Utility and factors influencing use. Northridge, CA: Educational Evaluation Associates.

Brown R.D., Braskamp L.A., & Newman D.L. (1978). Evaluator credibility as a function of report style: Do jargon and data make a difference? Evaluation Quarterly, 2 (2), 331-341.

Daillak R.H. (1980). A field study of evaluators at work (CSE Report No. 154). Los Angeles: UCLA Center for the Study of Evaluation.

Festinger L. (1957). The human organization. New York: McGraw-Hill.

Fiedler F. (1967). A theory of leadership effectiveness. New York: McGraw-Hill.

Hollander E.P. (1978). Leadership Dynamics. New York: The Free Press.

House R.J., & Baetz M.L. (1979). Leadership: Some generalizations and new research directions. In B.M. Staw (Ed.), Research in organization behavior. Greenwich, CT: JAI Press.

Katz D., & Kahn R. (1978). The social psychology of organizations (2nd ed.). New York: John Wiley & Sons.

Keen P.G.W., & Scott Morton M.S. (1978). Decision support systems. Reading, MA: Addison-Wesley.

Kotler P. (1980). Marketing management (4th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Likert R. (1967). The human organization. New York: McGraw-Hill.

Likert R. (1961). New patterns in management. New York: McGraw-Hill.

Patton M.Q., Grimes P.S., Guthrie K.M., Brennan N.J., French D.B., & Blyth D.A. (1978). In search of impact: An analysis of utilization of federal health evaluation research. In T.D. Cook and associates (Eds.), Evaluation Studies Review Annual (Vol. 3). Beverly Hills, CA: Sage Publications.

Rogers E.M. (1962). Diffusion of innovations. New York: The Free Press.

Stecher B.M., Alkin M.C., & Flesher F. (1981). Patterns of information use in school level decision making (CSE Report No. 160). Los Angeles: UCLA Center for the Study of Evaluation.

Stogdill R. (1974). Handbook of leadership. New York: The Free Press.

Vroom V.H., & Yetton P. (1973). Leadership and decision making. Pittsburgh, PA: University of Pittsburgh Press.

Weick K.E. (1966). The concept of equity in the perception of pay. Administrative Science Quarterly, 11, 414-439.

APPENDICES

Appendix A

AERA DIVISION H
AWARD FOR EXCELLENCE IN EVALUATION UTILIZATION

Evaluator: _____

Program evaluated: _____

School district/Organization: _____

Address: _____

_____

1. Briefly describe the program you evaluated.

2. What use was made of the evaluation information which you produced? Please describe in detail and provide documentation of use, if available. Do not send the evaluation report itself.

3. What were the unique factors in this evaluation contributing to making use happen?

Appendix B Verification Form

AERA DIVISION H
AWARD FOR EXCELLENCE IN EVALUATION UTILIZATION

Dear _____:

Your evaluator, _____, who worked on the evaluation of _____, is a nominee for an AERA
(name of program or project)
award for excellence in the area of promoting evaluation use. We would appreciate your comments to assist our panel in judging the performance of this nominee on the above-noted program evaluation.

1. In what way or ways was the evaluation information provided to you about the above program useful?

2. In your judgment, what aspects of the evaluator's own performance during the evaluation helped make the evaluation useful to you?

Mail this form directly to:

Dr. Marvin C. Alkin
Graduate School of Education
Center for the Study of Evaluation
UCLA
Los Angeles, CA 90024

Name: _____

Position: _____

Date: _____

Please return by March 1, 1984.

Appendix C

AERA DIVISION H
AWARD FOR EXCELLENCE IN EVALUATION UTILIZATION

REVIEWER RATING FORM

Please rate the applicant's submission, including documentation and validation, on the following criteria. Use the corroborating evidence submitted in support of the application to assist your judgments.

1. Conception of evaluation use:

_____ 1 the applicant showed little or no understanding of evaluation us.

_____ 2

_____ 3 the applicant showed an acceptable level of understanding of evaluation use

_____ 4

_____ 5 the applicant showed full understanding of evaluation use

2. Degree of evaluation use:

_____ 1 no indication of evaluation use

_____ 2

_____ 3 some indication of evaluation use

_____ 4

_____ 5 strong indication of evaluation use

3. Degree of direct link between evaluator efforts and subsequent use:

_____ 1 no link was demonstrated between the evaluator's efforts and subsequent use

_____ 2

_____ 3 a minimal link was demonstrated between the evaluator's efforts and subsequent use

_____ 4

_____ 5 a strong link was demonstrated between the evaluator's efforts and subsequent use.

4. Uniqueness/creativity of evaluator's effort at promoting use:

_____ 1 the evaluator showed no creativity or originality of effort

_____ 2

_____ 3 the evaluator showed a minimum of creativity or originality of effort

_____ 4

_____ 5 the evaluator showed a significant amount of creativity or originality of effort

5. Please rank the entries you have read.

Rank	Entry No.
____	_____
____	_____
____	_____
____	_____

Name of Reviewer _____

Date completed _____

Appendix D

Frequency of User Comments by Utilization Framework Categories[1]

Category	Factor	Element	Frequency
Human Factors	Evaluator Characteristics		
		Commitment to use	3
		Willingness to involve users[2]	6
		Choice of role -- leadership[2]	13
		Rapport with users	7
		Political sensitivity	5
		Credibility	9
		Background and identity	6
		Level of perceived effort[2]	6
	Total Category Frequency		55
Context Factors	Project Characteristics	Innovativeness	1
	Total Category Frequency		1
Evaluation Factors	Evaluation Procedures	Methods	18
		Use of a general model	1
	Information Dialogue	Amount and quality of interaction	6
	Substance of Evaluation Information	Information relevance	9
		Information specificity	7

[1]These data are based on the comments of 34 users.

[2]This variation of the existing factor was newly discovered in this research.

Appendix D (cont'd.)

Category	Factor	Element	Frequency
Evaluation Factors (cont'd.)	Evaluation Reporting		
		Frequency of information provided	1
		Timing of information	3
		Style of oral presentations	1
		Format of reports	3
		Mix of statistical/narrative data	6
	Total Category Frequency		55
	Total Frequency		111

Appendix E

Frequency of Evaluator Comments by Utilization Framework Categories[3]

Category	Factor	Element	Frequency
Human Factors	Evaluator Characteristics		
		Willingness to involve users	7
		Choice of role	3
		Rapport with users	3
		Political sensitivity	2
		Credibility	2
	User Characteristics		
		Interest in the evaluation	2
		Commitment to use	4
		Professional style	1
	Total Category Frequency		24
Context Factors	Organizational Features		
		External features	2
	Total Category Frequency		2
Evaluation Factors	Evaluation Procedures		
		Methods	11
		Use of a general model	1
	Information Dialogue		
		Amount of interaction	1
	Substance of Evaluation Information		
		Information relevance	4
		Information specificity	1

[3]These data are based on the comments of 12 evaluators.

Appendix E (cont'd.)

Category	Factor	Element	Frequency
Evaluation Factors (cont'd.)	Evaluation Reporting		
		Frequency of information provided	2
		Timing of information	1
		Style of oral presentation	1
		Format of reports	2
		Mix of statistical/narrative data	3
	Total Category Frequency		27
	Total Frequency		53

DELIVERABLE -- NOVEMBER 1984

EVALUATION PRODUCTIVITY PROJECT

Reflections on Evaluation Costs: Direct and Indirect

Marvin C. Alkin
Project Director

Grant Number
NIE-G-84-0112, P2

CENTER FOR THE STUDY OF EVALUATION
Graduate School of Education
University of California, Los Angeles

Reflections on Evaluation Costs: Direct and Indirect

Report Prepared by:

Marvin C. Alkin
Joan A. Ruskus

Center for the Study of Evaluation
University of California, Los Angeles

November 1984

REFLECTIONS ON EVALUATION COSTS: DIRECT AND INDIRECT

Marvin C. Alkin
Joan A. Ruskus

CSE Report No. 239
1984

CENTER FOR THE STUDY OF EVALUATION
Graduate School of Education
University of California, Los Angeles

The project presented or reported herein was performed pursuant to a grant from the National Institute of Education, Department of Education. However, the opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement by the National Institute of Education should be inferred.

Acknowledgements

The authors wish to thank James S. Catterall for his review of this manuscript. His insight regarding the delineation of direct and indirect costs was most helpful. Special thanks are also due to Andrea S. Larson for her exceptional word-processing skills.

Table of Contents

The Evaluation Productivity Project, scheduled to run through November 1985, was initially intended to explore the relationship between costs and the extent of utilization of evaluation. The original plan called for the development of a set of case scenarios during FY 1984, to be used in the empirical research of FY 1985. Each of these scenarios was to be "costed" in terms of its direct and indirect costs. Subsequently, the plan was altered to give greater emphasis in FY 1985 to additional synthesis and dissemination of the project's longstanding research on evaluation utilization. This change in emphasis, along with the decision not to undertake new empirical research, rendered the planned costing exercise pointless. Nonetheless, it seems appropriate to set down here some of the reflections on evaluation costs that emerged during the initial stages of work. (See also Alkin & Solmon, The Costs of Evaluation, 1983.)

This paper, then, summarizes our views on the costs of evaluation and, in particular, our hunches about the kinds of costs associated with factors known to affect utilization. The first section deals with general issues involved in identifying and valuing cost components. We will then discuss the appropriate costs of evaluation and consider costs in relation to

benefits. The final section spells out the cost implications of high-utilization factors.

Identifying and Valuing Evaluation Costs

Before we consider how the costs of evaluation can be identified and valued, a distinction must be made between direct and indirect costs. "Direct costs" are simply the direct cash outlays necessary for initiating and implementing the evaluation. Examples include salaries, travel expenses, and the cost of test booklets and data processing. Because direct costs are easy to identify, they are often regarded as synonymous with the costs of evaluation.

Like other educational program activities, however, evaluations involve indirect as well as direct costs. While our discussion of costing focuses on the direct costs of evaluation, some consideration of the definition of indirect costs, and of general issues related to their identification and inclusion in the overall "cost" package, seems warranted.

Most authorities (Alkin & Stecher, 1983; Catterall, 1983; Haggart, 1983; Morell & Weirich, 1983; Sanders, 1983; Solmon, 1983) recognize the importance of these less obvious costs. But there is confusion in the literature over exactly how indirect costs should be defined. As a result of both this ambiguity and also their tendency to be less visible, indirect costs are often overlooked or underestimated when the costs of evaluation are computed.

"Indirect costs" can be defined in one of three ways. According to one point of view (Morell & Weirich, 1983), indirect costs are those dollar costs not specifically incurred by the evaluation project but shared across projects within the larger organizational structure: clerical time, facilities, communication, and even, in some instances, the salaries of project staff. This definition is probably the most straightforward. Indirect costs of this type are often calculated as a percentage of the direct costs of the evaluation and are included as a separate line item in the overall budget.

According to a second point of view (Catterall, 1983), indirect costs are the opportunity costs of the evaluation: the use of equipment that could be used in other ways; volunteer time that could be spent on other projects; pupil time that could be spent in learning instead of testing; and the time the state agency spends in legislating, developing, and monitoring the evaluation. Almost every party involved in an evaluation could be engaging in other, possibly more valuable, activities. Thus, the opportunity cost is the best alternative use of a resource used in the evaluation. Interpreted in this way, indirect costs can be included in the cost package as the dollar equivalent of alternative uses of evaluation resources, though Catterall suggests that these indirect costs are often better presented in their natural units (for instance, as the number of hours students could spend learning) rather than their dollar equivalents.

Yet a third point of view (Solmon, 1983) holds that indirect costs are the side effects of the evaluation. For example, when a project concentrates on achieving immediately observable objectives in anticipation of an evaluation, certain long-term project goals may be slighted or even sacrificed altogether. Thus, the evaluation can be said to entail indirect costs. (It should be noted that Solmon considers opportunity costs separately from indirect costs.)

As mentioned above, indirect costs in the first sense -- shared costs of projects within an organization -- are relatively easy to calculate, since they represent dollar costs for materials, personnel, and so forth. Calculating indirect costs in the second sense -- opportunity costs -- is somewhat more difficult, since one must determine alternative uses of resources and their corresponding values. It is indirect costs in the third sense -- as side effects -- that are most difficult to calculate; and we are not aware of systematic attempts to include such indirect costs as part of a total evaluation cost package.

Several authors (Catterall, 1983; Haggart, 1983; Levin, 1983) have outlined the procedures necessary to identify the costs associated with an evaluation. According to Haggart, the process begins with a definition of the scope of the evaluation. The scope depends on the extent of the education intervention, the level of decision-making involved (the higher the level, the greater the scope of the evaluation), the purpose of the evaluation, and the complexity of the evaluation design. These aspects of the evaluation provide the context within which

appropriate cost categories may be identified.

Most authorities agree on the specific cost categories to be employed. For example, the cost category systems presented by Alkin and Stecher (1983), Morell and Weirich (1983), and Sanders (1983) are highly similar, differing only in detail. Drawing on the input of a large sample of professional evaluators, Alkin and Stecher (1983) delineate seven typical cost categories: professional staff, clerical staff, external consultants, materials/supplies/telephone, data processing, facilities, and travel. Similarly, Morell and Weirich (1983) break down evaluation costs into the following categories: personnel (evaluation staff, consultants, program staff, and subjects), technological capabilities (data processing, telephone services, and duplication) travel, office space and furnishings, supplies, and dissemination of findings (graphic artists, printers, and audiovisual specialists). The most detailed listing of cost categories is that provided by Sanders (1983): evaluation staff salary and benefits, consultants, travel and per diem, communication, printing and duplication, data processing, printed materials, office supplies, subcontracts (outside of consulting), and overhead. Sanders expands on the types of costs frequently involved in data processing: systems design, data coding and checking, data storage and retrieval, computer programming, computer use for manipulating or analyzing data, and computer-based bibliographic searches.

After the cost categories have been identified, the next step is to determine values for each of these categories. One

method for valuing "ingredients" is described by Levin (1983), who offers the following hints about costing:

Personnel

Salaries and fringe benefits (percentage of salaries): Value determined by marketplace prices. When personnel costs cannot easily be ascertained, use estimates of market value for similar services.

Facilities

Rented/leased facilities: Value determined by annual cost of expenditure.

Owned facilities: Value determined by cost for similar space or annual cost (taking depreciation and interest on remaining undepreciated value into account).

Equipment

Purchased: Value determined in the same way as for facilities, depending on whether leased or owned.

Donated/Borrowed: Value determined as cost of leasing or renting similar equipment.

Supplies

Value determined by adding estimated expenditures to estimated value of contributed supplies.

Client Inputs

Service: Value determined by total expense associated with service or by market price of service provided.

Note that these valuing procedures can be adapted to any of the cost-category frameworks described above.

Using a variation of Levin's ingredients approach, Catterall (1983) offers an analysis of the costs of testing that can be generalized to evaluation. The first step is to do a cost inventory. Again, the cost-category frameworks described above constitute alternative ways of completing the cost inventory. The second step is to total the costs, which involves determining the actual monetary figure for each of the cost categories. (Levin's procedures outlined above provide a methodology for this step.) The last step in Catterall's process is to locate the cost: that is, to decide who will pay for a particular component -- the sponsor, a government agency, private party, the clients or subjects. An examination of Catterall's variation of the Levin Ingredient Chart (Levin, 1975; see Figure 1) provides insights into the distributed features of evaluation costs. This chart is particularly informative with respect to indirect costs. Note, for example, that most of the cost-associated columns (particularly columns four, five, and six) will usually list indirect costs, although there may be some direct costs related to contributed private inputs. The last row in the chart, client time and other client inputs, presents a set of costs, all of which are likely to be indirect. In addition, many facilities and equipment costs would be considered indirect if accounted for in the costs of conducting evaluations.

Entity Bearing Costs

(1) Ingredients	(2) Total Cost	(3) Cost to Sponsor	(4) Cost to Other Levels Government or Agencies	(5) Contributed Private Inputs	(6) Imposed Student & Family Costs
Personnel					
Facilities					
Material & Equipment					
Other (Specify)					
Client Time & Other Client Inputs					

TOTAL: $_____

Source: Adapted from Levin, 1975; p. 101.

Figure 1. Illustrative Framework for Cost Accounting in Educational Programs.

Appropriate Costs of Evaluation

Clearly, evaluations carry both direct and indirect costs, and both can be calculated. But just what costs are "appropriate" (i.e., what level of resources should be devoted to asking and answering questions about how well programs work or what they achieve)? This question can be approached in two ways. At a macro level, one can compare the cost of evaluation with the cost of other program elements. That is, one can ask what the total costs of an evaluation should be relative to the

entity being evaluated. At a micro level, one can look at the way in which direct costs are allocated among the various budget categories and at the nature of the indirect costs associated with the conduct of the evaluation. Here we might ask what particular sorts of costs are worth incurring, given the nature of our evaluative interest.

As noted in The Costs of Evaluation (Alkin & Solmon, 1983), the literature provides very few guidelines with respect to the macro level. Rusnell (1979) indicates that the evaluation cost should amount to 10 percent of program cost, a figure that had been recommended during the early years of the federal Title VII programs under the Elementary and Secondary Education Act (ESEA). More specifically, in the operation of these bilingual/bicultural programs during the early 1970s, it was suggested that funds for evaluation activities should constitute 8 percent and funds for audit activities should constitute 2 percent of programs costs. More recently, Drezek and his associates (1982) surveyed 55 LEAs and 14 SEAs and reported the proportions of funds allocated for evaluation in the various program proposals of those agencies. As expected, the percentages varied substantially from one program to another; the median low was 1.5 percent and the median high was 5.5 percent. In addition, respondents to the survey recommended a 4-8 percent range as desireable.

Guidelines at the micro level are even more scarce. In a study designed to gather data on the direct costs of the typical evaluation, Alkin and Stecher (1983) asked a nationwide group of

evaluators to consider the costs associated with two types of evaluation (process/implementation and outcome/summative) and with three budget levels ($25,000, $10,000, and $4,000). Further constraints were imposed on the hypothetical examples, to insure that the cost estimates would be comparable. Overall (that is, for both types of evaluation and at all three budget levels), the average shares recommended for each category were these:

Professional staff	70%
Clerical/secretarial staff	16%
External consultants	2%
Materials, supplies, telephone	5%
Data processing	3%
Facilities	0%
Travel	3%

The specific recommended cost breakdowns varied widely, depending in part on evaluation type and on total budget level. Many of these differences are easily explained. For example, the percentage of direct costs allocated for travel was four to five times higher in implementation/process evaluations than in outcome/summative evaluations. Process evaluations typically require more site visits and more consultation with people in the field than do summative evaluations.

Total budget level made a difference with respect to three of the seven cost categories. The first was data processing: "Data processing expenditures rose dramatically as the total evaluation budget increased from $4,000 to $10,000 and continued to rise in dollars (but not as a percentage of the total budget) when the total budget increased to $25,000" (Alkin & Stecher, 1983, p. 7). One would expect data-processing costs to be low in a total evaluation budget of $4,000, since most of the analytic

work connected with such a low-cost evaluation would probably be done on a hand calculator. Moreover, the percentage of the budget allocated to data processing would probably not increase beyond a certain level, due to economies of scale: The initial outlay is high, but incremental costs for added units are relatively low.

The second cost category where the proportionate allocation increased with the size of the total budget was travel: 0 percent of the $4,000 budget, 2-3 percent of the $10,000 budget, and 5-8 percent of the $25,000 budget. Similarly, the proportion allocated for the third cost category -- clerical and secretarial staff -- rose as the size of the budget increased, from only 5 percent at the $4,000 level to 20 percent at the $25,000 level. These systematic differences make clear the extent to which indirect dollar costs are likely to be present in various evaluation budgets. The amount of data preparation, reporting, and just plain bureaucratic red tape seems to grow as the scope of the evaluation grows. Thus, a greater proportion of secretarial time is required at higher budget levels. Another potential explanation -- and one that may be more relevant -- is that, at small budget (or direct cost) levels, secretaries and clerical personnel are not easily divisible into arbitrary smaller budgeting units. Thus, these services are often provided by other projects or even by the school district itself. As a result, the secretarial costs of small evaluation projects tend to be disregarded when costs of evaluations are contemplated. They actually represent an indirect cost which may, on the basis

of our findings, be fairly substantial in small-scale evaluations.

Our research on the costs of evaluation has convinced us that it is difficult, if not impossible, to generalize about the direct or indirect costs associated with evaluations of different types. Idiosyncratic features affect the costs of evaluation and the extent of direct and indirect costs. These features can be grouped into four categories which, though occasionally overlapping, offer a convenient schema for purposes of discussion: context, administrative organization, support services and facilities, and administrative expectations and predispositions.

One example that immediately comes to mind with respect to context is choosing between an internal and an external evaluator. The selection of an internal evaluator may impose indirect cost burdens on other units within the organization. Conversely, if an external evaluator is chosen, many more of the costs will be direct because of the necessity for a contract.

The administrative organization of the school district may impose different constraints upon the evaluation and thus affect its cost. If a complicated variety of approvals (for instruments, testing schedules, and so forth) and other administrative procedures are required, not only will the evaluator have to spend more time on the evaluation (a direct cost) but also other people in the organization will have to

spend more of their time processing requests and making approvals (an indirect cost).

Likewise, the nature of the support services and facilities available within a school district may have cost implications. For instance, the availability of computer services converts what would otherwise be a direct cost into an indirect cost. Similarly, if the evaluator has access to unused or underutilized facilities within the district, then the direct cost of renting office space will be saved.

Finally, the expectations of program administrators -- and especially their predisposition to evaluation use -- can affect both the direct and the indirect costs of an evaluation. Clearly, if the administrators who are the potential users of evaluation findings are hostile toward the evaluation and inclined to dismiss its findings, the evaluator will have to spend more time and energy trying to modify these predispositions, thus increasing both direct and indirect costs. (This issue is discussed further in the next section.)

Costs in Relation to Benefits

Most authorities agree that the real issue is neither the absolute dollar cost of an evaluation nor the cost of the evaluation relative to the cost of the program being evaluated. The real issue is the relationship between the costs incurred by and the benefits that accrue from the evaluation. For example, Scriven (1974) notes that evaluations should be "cost free,"

implying that they should at least pay for themselves in the benefits they provide. Clearly, Scriven is advocating that costs and benefits should be compared by subtracting costs from benefits (B-C). Similarly, in Standards for Evaluations of Educational Programs, Projects, and Materials (1981), the Joint Committee on Standards for Educational Evaluation says that "the evaluation should produce information of sufficient value to justify the resources expended" (p. 60; emphasis added). The Joint Committee's guidelines further indicate that one should thoroughly investigate the costs and benefits of an evaluation before deciding to undertake it, adding that the evaluator should "conduct evaluations as economically as possible" (p. 61).

If one accepts this point of view, a critical question becomes: How can the benefits deriving from an evaluation be determined? In other words, what constitutes an appropriate measure of benefits? Both Scriven and the Joint Committee would probably say that the dollar savings produced by the evaluation is the most appropriate measure. If the evaluation has resulted in recommendations as to how the program can be conducted in a less costly manner (with no reduction in the quantity or quality of educational outputs), and if the cost savings exceed the costs of the evaluation, then the cost of the evaluation is justified.

This simple notion has been successfully practiced by Steven Frankel, Director of the Montgomery County (Maryland) Department of Education Accountability. In essence, by selecting for evaluation only those projects and services which are readily amenable to the demonstration of cost savings (primarily business

service activities), his Department has dramatically increased its total budget for evaluation and has freed up resources for conducting evaluations in which cost savings are not involved.

The simple dollar-savings criterion is obviously inadequate in those instances where cost economies cannot be demonstrated but where the evaluation still confers benefits, though of a less monetary nature. For example, an evaluation may result in recommendations on ways to increase instructional effectiveness which, if implemented, will lead to an increase in student achievement. Such an evaluation may well be considered worth the dollar cost. Several issues arise here. The first is cost effectiveness: the dollar cost of the evaluation relative to the resulting increase in instructional effectiveness (and hence in student achievement). The second issue relates to the economist's concern with indirect costs in the second sense of the term: opportunity costs. Granted that the benefits of an evaluation -- in terms of increased student achievement -- exceed the costs of the evaluation, it is nonetheless possible that alternative uses of the evaluation funds might have led to even greater benefits (i.e., higher student achievement).

The third issue is more complicated: The recommendations of an evaluation are not always implemented, so one cannot always determine whether they produce benefits. Some authorities would take the position that, in such instances, one should consider the potential benefits that would have accrued, had the recommendations been implemented, and use those as a basis for judging the appropriateness of the costs. Other authorities

would maintain that, if the evaluation was not persuasive enough to convince decision-makers that its recommendations should be implemented, then the evaluation has failed and cannot be regarded as having produced benefits. The question comes down to this: Should one look at potential or at actual benefits? Is evaluation only a type of research, whose benefits are to be judged by the nature of the research findings? Or is evaluation a decision-oriented interactive process, whose benefits are to be judged not so much by the recommendations it makes but by the extent to which it beneficially informs the decision process? Obviously, the answers one gives to these questions will determine, in large part, just how one calculates the cost-benefit equation.

Cost Implications of High-Utilization Factors

The Evaluation Productivity Project has, over a period of years, been concerned with identifying those factors associated with instances of high evaluation utilization (see, for example, Alkin, Daillak, & White, 1979; Daillak, 1980; Stecher, Alkin, & Flesher, 1981). During the current fiscal year, we have completed a handbook for evaluation decision-makers that includes a factor framework (Alkin, Jacobson, Burry, Ruskus, White, & Kent, in press).

One conclusion to emerge from our years of research is that the way in which the evaluation is conducted affects the extent to which its findings are utilized. As mentioned in the previous section, the real issue in judging the appropriateness of an

evaluation's cost is the extent to which the benefits exceed the cost. If the findings of an evaluation are not utilized at any level, then it is questionable whether even the most minimal evaluation costs are justified. At the same time, one must recognize that the very aspects of the evaluation most likely to increase its use are also likely to involve high costs in terms of both dollars and also the time and energy of almost everyone involved.

The results of our studies have not been surprising. We have found that the dedicated evaluator -- the one who takes the time to understand the political complexities of the evaluation situation, to consider the needs and interests of potential users, and to involve them in the planning and conduct of the evaluation -- stands a better chance of seeing the findings of the evaluation utilized. But at the same time, this approach incurs some heavy costs, at least in terms of the evaluator's time and the time of program staff. Similarly, evaluation utilization increases when the potential user -- the program administrator who commissioned the evaluation -- is interested in the evaluation, is involved in the evaluation process, and is committed to integrating the evaluation findings into the decision process. Obviously, this degree of involvement on the part of potential users entails substantial indirect costs in the form of forgone opportunities to use their time in other, possibly more productive, ways.

Let us look more closely at some of the factors just mentioned, starting with the political sensitivity of the

evaluator. The politically sensitive evaluator must be particularly attentive to the people within the organization, their place in the administrative structure, their special interests, and so forth. Such attentiveness may require a considerable outlay of the evaluator's time in becoming familiar and dealing with the situation. Moreover, if sensitivity to multiple constituencies adds to what must be known, there may be additional costs associated with extended data collection and data processing. In addition to these direct costs, there are indirect costs: for instance, the time which program personnel spend interacting with the evaluator, the time which clients or pupils spend taking tests.

The evaluator's credibility is another factor that affects utilization. Credibility depends in part upon the evaluator's credentials (e.g., academic or professional degrees, prestige or reputation, institutional affiliations, experience). Obviously, a highly credentialed evaluator costs more than a relatively uncredentialed evaluator: in consultant fees for external evaluators and in salaries for evaluators internal to the organization. Our research shows, however, that credibility is a function not only of the evaluator's credentials at the outset of the evaluation but of the evaluator's actions during the course of the evaluation: "As evaluators engage in their activities, they may come to be viewed as credible on a wider range of topics or credible to new audiences." (Alkin et al., 1979, p. 247). In short, an evaluator can build credibility, but this process is time-consuming and involves both direct and indirect costs.

As indicated earlier, the potential user is also a key figure in evaluation utilization. The more the potential user is concerned about the evaluation, involved in its conduct, and interested in its results, the greater the likelihood that the findings of the evaluation will be utilized. Obviously, this kind of user involvement has cost implications, especially in terms of indirect costs. And the larger the number of potential users, the higher the costs.

Not only evaluator and user characteristics but also certain project characteristics have cost implications. One example is the particular requirements that the organization imposes upon the evaluation. To the extent that these contractual obligations or written requirements are not directly relevant to the central concerns of the evaluation or the interests of the potential users, they may entail high costs -- in terms of the time and research spent by the evaluator and the time spent by program personnel in interacting with the evaluator -- and yet have no accompanying payoff in high utilization.

The evaluation procedures, including design and methodology, obviously have cost implications. To the extent that these procedures are tailored to the particular needs of the program being evaluated and are applied rigorously, their costs may be high. These costs include outlays for materials, data processing, and so forth; the time and effort of the evaluation staff and program personnel; and client time spent in testing. Nonetheless, appropriate procedures are essential to high utilization, insofar as the way in which the evaluation is

conducted influences potential users' perceptions of the evaluator's credibility and of the quality of the evaluation.

Finally, evaluation reporting, which is strongly related to utilization, has cost implications. To prepare a report that can be easily understood by all potential audiences and to disseminate it in a timely manner requires considerable effort on the part of the evaluator, and thus considerable cost. Similarly, follow-up procedures -- designed to assure that users understand the report and its recommendations -- call for extra time and effort and therefore carry extra costs. Our research shows that simply preparing and distributing a report of the evaluation is not enough, if high utilization is desired. Evaluators should regard evaluation as a process, not a product. They should view evaluation reporting as an almost-continuous set of activities designed to sensitize potential users to the information being developed, to prepare them for the findings that will emerge from the evaluation, and to encourage them to implement the recommendations. Without question, such activities ail high costs, both direct and indirect. Nonetheless, they ssential to attaining high levels of evaluation use.

Summary

These reflections on the costs of evaluation represent a brief summary of the work we have completed in this area and of the implications that can be drawn from it. In synthesizing our reflections, we have drawn heavily on The Costs of Evaluation (Alkin & Solmon, 1983). Many authors contributed chapters to

this volume, and some of their ideas are incorporated into this piece. Our perspectives on the costs associated with various utilization factors grew out of our previous effort in developing the utilization framework. Thus, the paper provides a synopsis of our work relative to the costs of evaluation, particularly evaluation which has a high potential for utilization.

REFERENCES

Alkin M.C., Daillak R.H., & White P. (1979). Using evaluations: Does evaluation make a difference? Beverly Hills, CA: Sage Publications.

Alkin M.C., Jacobson P., Burry J., Ruskus J., White P., & Kent L. (1985, in press). Handbook for evaluation decision-makers. Beverly Hills, CA: Sage Publications.

Alkin M.C., & Solmon L.C. (Eds.). (1983). The costs of evaluation. Beverly Hills, CA: Sage Publications.

Alkin M.C., & Stecher B. (1983). A study of evaluation costs. In M.C. Alkin & L.C. Solmon (Eds.), The costs of evaluation (pp. 119-132). Beverly Hills, CA: Sage Publications.

Catterall J.S. (1983). Fundamental issues in the costing of testing programs. In M.C. Alkin & L.C. Solmon (Eds.), The costs of evaluation (pp. 71-80). Beverly Hills, CA: Sage Publications.

Daillak R.H. (1980). A field study of evaluators at work (CSE Report No. 154). Los Angeles: UCLA Center for the Study of Evaluation.

Drezek S., Monkowski P.G., & Higgins P.S. (1982). Current vs. perceived-ideal procedures for determining educational program-evaluation budgets: A survey of school evaluators. Educational Evaluation and Policy Analysis, 4 (1), 97-108.

Haggart S.A. (1983). Determining the resource requirements and cost of evaluation. In M.C. Alkin & L.C. Solmon (Eds.), The costs of evaluation (pp. 59-70). Beverly Hills, CA: Sage Publications.

Joint Committee on Standards for Educational Evaluation. (1981). Standards for evaluations of educational programs, projects, and materials. New York: McGraw-Hill.

Levin H.M. (1983). Cost-effectiveness: A primer. Beverly Hills, CA: Sage Publications.

Levin H.M. (1975). Cost-effectiveness in evaluation research. In M. Guttentag & E. Struening (Eds.), Handbook of evaluation research (Vol. 2). Beverly Hills, CA: Sage Publications.

Morell J.A., & Weirich T.W. (1983). Determining the costs of evaluation: Principles from mental health. In M.C. Alkin & L.C. Solmon (Eds.), The costs of evaluation (pp. 81-100). Beverly Hills, CA: Sage Publications.

Rusnell D. (1979). Cost-effective evaluation: Is there a $500 solution for a $1,000 problem? New Directions for Continuing Education, 3, 97-107.

Sanders J.R. (1983). Cost implications of the standards. In M.C. Alkin & L.C. Solmon (Eds.), The costs of evaluation (pp. 101-117). Beverly Hills, CA: Sage Publications.

Scriven M. (1974). Evaluation perspectives and procedures. In W.J. Popham (Ed.), Evaluation in education: Current applications. Berkeley, CA: McCutchan.

Solmon L.C. (1983). Economic issues in considering the costs of evaluation. In M.C. Alkin & L.C. Solmon (Eds.), The costs of evaluation (pp. 15-26). Beverly Hills, CA: Sage Publications.

Stecher B.M., Alkin M.C., & Flesher F. (1981). Patterns of information use in school level decision making (CSE Report No. 160). Los Angeles: UCLA Center for the Study of Evaluation.