

DOCUMENT RESUME

ED 211 094

IR 009 919

AUTHOR Rhodes, Lewis A.
TITLE The Communication of Experience. A Guidebook for the Management of Information.
INSTITUTION Applied Management Sciences, Inc., Silver Spring, Md.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
PUB DATE Oct 79
CONTRACT 300-78-0580
NOTE 184p.; Best available copy.

EDRS PRICE MF01/PC08 Plus Postage.
DESCRIPTORS *Administrators; *Art Education; Attitude Measures; Communication (Thought Transfer); *Documentation; *Evaluation, Methods; *Information Dissemination; Models; Problem Solving; *Program Evaluation; Questionnaires; Research Reports
IDENTIFIERS *Information Management

ABSTRACT

Designed to meet the information needs of arts educators who function as managers of programs, projects, or other administratively definable units in local or state education agencies, this guidebook synthesizes the experiences from four regional workshops held in 1978-79 to address project documentation, evaluation, and dissemination. Topics covered include the control and influence of information, and its use in project problem solving; reporting and dissemination, including packaging; evaluation for program planning, improvement, and justification, and the steps in evaluation; and the identification of a strategy or evaluation model. A discussion of the management of information summarizes the report, and appendices provide background material, bibliographies, and evaluation articles and reports. (REF)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED211094

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- * This document has been reproduced as received from the person or organization originating it
- Minor changes have been made to improve reproduction quality
- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy

THE COMMUNICATION OF EXPERIENCE

A GUIDEBOOK

FOR THE

MANAGEMENT OF INFORMATION

BY

U.S. OFFICE OF EDUCATION

ARTS EDUCATION PROJECT DIRECTORS

DEPT. OF EDUCATION
COPY

developed by

Applied Management Sciences
962 Wayne Avenue
Silver Spring, MD.

under

Contract No. 300-78-0580

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Lewis A. Rhodes

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

IR009919



The activity which is the subject of this report was supported in whole or in part by the Office of Education, U.S. Department of Health, Education, and Welfare. However, the opinions expressed herein do not necessarily reflect the position or policy of the Office of Education, and no official endorsement by the Office of Education should be inferred.

Acknowledgements

This Guidebook represents a synthesis of experiences. It is built upon the original design for four regional workshops developed for the 1978-79 USOE Arts Education Project Directors by Applied Management Sciences. However, it reflects more than that; for it has been enriched by the 80 local and State project directors who provided continuing feedback to improve the content as each workshop was conducted. Special thanks to each of those workshop participants and to Dr. Harold Arberg whose sensitivity, continuing understanding and support made this possible.

Finally, the project and workshop staff who planned and created the original design deserve particular mention:

Della Cooper, who coordinated all workshop logistics and facilitated sessions;

Dr. Sondra B. Gair, who provided specific content input from the arts;

Anna Beth Payne, who participated with conceptualization of the workshops and Guidebook; and

Nancy O. Gentile, who designed and delivered the evaluation components of both the workshops and this manual.

The task of synthesizing the experiences and producing this document was the responsibility of Lewis Rhodes and Nancy Gentile. Kenneth Freeman edited the material.

Lewis A. Rhodes
Project Director
Applied Management Sciences

October, 1979

TABLE OF CONTENTS

		<u>Page</u>
I.	INTRODUCTION	1
	Background	2
	Organization	3
	Definitions	5
II.	INFORMATION--who has it? who wants it? For what?	7
	Control and Influence	7
	"Soft" data/"Hard" data	7
	Information in Project Problem-Solving	12
III.	REPORTING AND DISSEMINATION	19
	Reporting	19
	Dissemination	21
	Why, what, how	22
	Packaging	23
	Multiple audiences, media, content	24
IV.	EVALUATION	27
	What Evaluation Is	27
	Conducting an Evaluation -	28
	o for program planning	
	o for program improvement	
	o for program justification	
	Steps in Evaluation -	30
	o Identification of audience needs	30
	o Identification of information sources: documentation	32
	The Growth Record Process	34
	Role of the Documentor	35
	The Project Growth Record	36
	o Identification of resources (staff and money)	44

TABLE OF CONTENTS (Continued)

	<u>Page</u>
o Identification of a Strategy--Evaluation Models	46
Discrepancy Evaluation	46
Goal-Free Evaluation	48
Illuminative Evaluation	49
Responsive Evaluation	50
CIPP Evaluation	52
Art-Criticism Evaluation	53
Jurisprudence Evaluation	54
V. SUMMARY--The Management of Information	56
<u>APPENDIX A - Background</u>	
Where we started ...	
<u>APPENDIX B - Bibliographies</u>	
Selected Tests in the Arts	
Questionnaire Design and Attitude Measurement	
<u>APPENDIX C - Evaluation Articles and Reports</u>	
A Blueprint for Program Evaluation	
Common Evaluation Hazards and How to Avoid Them	
The Camera as an Evaluation and Research Instrument	
Facilitation Evaluation	
How to Evaluate Educational Programs, Vol. 2, No. 1.	
Evaluating an AGE Program	
An Evaluation of the Twin City Institute for Talented Youth	

"Despite information theory and processing, no one yet has actually seen let alone used, an 'information system,' or a 'data base'....The main conclusion to which our experience with communications - largely an experience of failure and all the work on learning, memory, perception, and motivation point: communication requires shared experience....The effectiveness (of an information system) in other words, depends on the pre-establishment of communication."

Peter F. Drucker
Management

THE COMMUNICATION OF EXPERIENCE

A GUIDEBOOK FOR THE MANAGEMENT OF INFORMATION
BY U.S. OFFICE OF EDUCATION ARTS EDUCATION PROJECT DIRECTORS

I. INTRODUCTION

This guide is designed to address the information needs of the Arts Educator who functions as a manager of a program, project or other administratively definable unit in a local or State educational system. The material is derived from the experiences of four regional workshops sponsored by the USOE Arts and Humanities Staff for their 1978-79 grant Project Directors.

These guidelines provide more than a report on the workshops yet somewhat less than a detailed "how-to" handbook.

The two-day workshops, entitled The Communication of Experience, brought more than 80 local and State arts educators together in Boston, Oklahoma City, Phoenix and Indianapolis. The focus of the workshops, as of this guidebook, was project success and survival--and how to manage information in order to ensure it. Major emphases were put on the information arts educators need:

- for their own projects' effective management (documentation);
- to convince others (evaluation); and
- to share among their peers (dissemination).

The underlying rationale for the workshops was that this information, regardless of how it is used, comes initially from the experiences of the project and staff as they work toward their goals. A major task of each

project is the transformation of these experiences into information and its use to achieve success. Thus the title, The Communication of Experience.

While there were some presentations of "how others have done things" at the workshops, the major concepts were developed experientially--that is, each Project Director brought certain experiences into the sessions which helped shape what they took away from it. This guidebook, similarly, is not a transcript of workshop presentations or discussions. Rather, it is organized around a framework of questions which synthesize many of the concerns of the Arts Educator who serves as a manager. To the degree you share these concerns, the "answers" suggested on these pages may be of value.

BACKGROUND

This guide, and the workshops, resulted from repeated calls for assistance from local and state grantees of the USOE Arts Education Program, as well as from Arts Educators elsewhere. The contractor for this effort, Applied Management Sciences of Silver Spring, Maryland, conducted an extensive review of the literature in this area with particular focus on recent national and regional meetings directly concerned with evaluation in arts education. The results of this review are reported, with examples of many of the most relevant papers, in the appendices.

Additionally, an analysis of the experiences of Arts Educators in the 80 projects funded by USOE in 1978-79 produced the following inferences, upon which rest the particular design of the workshops and of this guidebook.

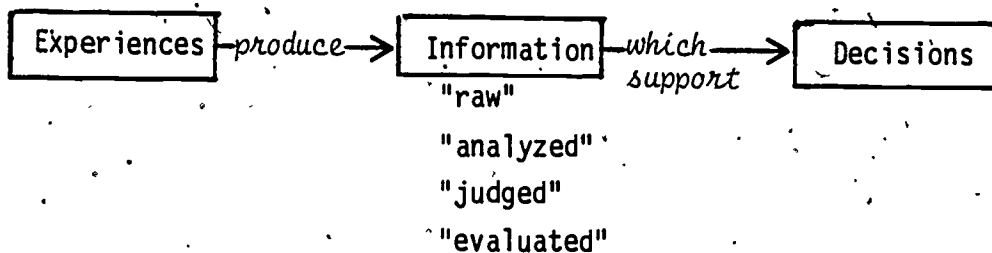
- First is that most people already know how to turn experiences into useful information for their own and others' decisions. But they do not usually associate these skills with the technical terms "evaluation," "documentation," and "dissemination."
- One of the factors constraining this association is a lifetime of professional experiences in which people have been burdened with collecting information for others; have found information they provide to others used primarily to make judgments about

them; and have found the information they needed inaccessible or unavailable at the time it was wanted. This has produced an "affective bias"--or, more simply stated, fear--of putting program information on paper.

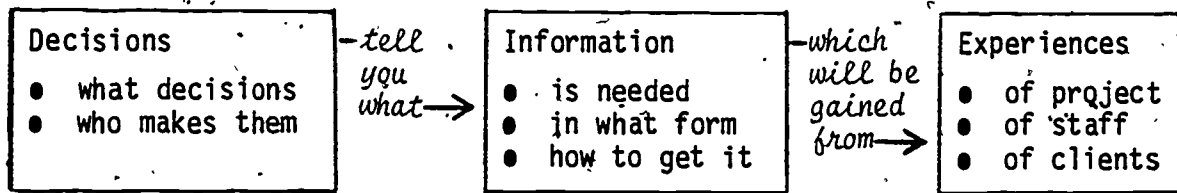
- A third inference serving as a foundation for this approach deals with the movement of information between people and programs. Good resource management usually requires that each decision-maker view his or her project or unit as an independent entity. But in reality, the success of any venture is based upon the ability to take advantage of the interdependence among elements. And, the use of information to support this interdependence has been constrained by the fears cited above, as well as the fact that seldom is any one person responsible for interdependence.
- Finally, we have an assumption of why a project or unit would want to collect and use information. The answer we propose is, survival, not research. Unless you get the information you need to assure success, you may end up not needing any information for measuring it. Too many good demonstration programs in education have ended up as cases of "The Operation was a Success...but the Doctor Died." Program survival often requires extending your concern beyond your grant period, on to the integration of your results with the ongoing processes of your or another's organization.

ORGANIZATION

This guidebook is organized "backwards." Although information processes usually end with decisions, we will start at that point. That is: the transformation of your project's experiences into information for use in the successful accomplishment of your purposes, can be displayed as:



However, the planning for this process has to start at the other end--with consideration of those decisions which are to be influenced by your project's experiences.



Therefore, following a listing of definitions, the first section of this guidebook addresses people, decisions and the information they need. This is followed by sections on reporting and disseminating information to affect decisions; and finally, documenting your project's experiences and evaluating them. The appendices include an array of articles and reports which were used as supplementary materials in the workshops.

DEFINITIONS

Following are a number of terms that will be used throughout the remainder of this guidebook. Note that these definitions are intended only to be what we mean by them.

- Assessment: Looking at the conditions existing in an environment--resources and people (both clients and providers).
- Decisions: Judgments based upon cumulative experience.
- Disseminating: Transforming your own experiences into useful information for others.
- Documenting: Collecting information about what you are doing and learning.
- Evaluating: Looking at what you did as it affects clients, staff, and environment--and making a judgment as to its worth for you and others.
- External Evaluator: A person who functions in the role of evaluator who is not connected with the day-to-day operations of the program.
- Formative Evaluation: The process of reviewing and evaluating ongoing programs for the purpose of changing strategies and making improvements before it is too late.
- Internal Evaluator: A person on the program or project staff who functions in the role of evaluator.
- Linkage: A relationship between two persons/organizations (with reference to a mutual concern) that serves the needs of both parties.
- Minimum Reporting Requirements: The convenience of compliance without the discomfort of thought.
- Model: The way someone else did it.
- Plan: Your best estimate of what should be done, until you get enough new data to guess again.
- Response Burden: The difference between a Federal or State decision-maker's desire to have good information and the practitioner's desire to provide it.
- Strategy: An approach to reaching a desired end that allows for unanticipated happenings.
- Summative Evaluation: Looking at the program upon completion of the effort.

II. INFORMATION: Who has it? Who wants it? For what?

Control and Influence

Project management is usually considered to include: budgeting, planning and scheduling, personnel management, development of support, progress monitoring, evaluation and dissemination. More simply stated, this means that there are just four elements that you manage--money, time, people, and information.

As manager, which of these do you feel you can control? or influence?

	<u>Control?</u>	<u>Influence?</u>
Funds	_____	_____
Time	_____	_____
People	_____	_____
Information	_____	_____

Most project directors are aware that only three of those elements can be "controlled." Two of them--money and time--are the focus of most management techniques. A third--information--can only be controlled if you have it. The fourth--people--can't be "controlled" at all, but they can be influenced. And it is those two elements--information and people--that are closely intertwined in the management of any demonstration or other type of project. Identifying, generating and using information to influence people can be a major factor in the successful integration of arts education into the school system of your community and state.

"Soft" data/"Hard" data

But what types of information are most influential? Ask yourself what forms of information most influenced the decisions you made in developing your current project? And, judging by the following article, you are probably not alone if you said that your decisions are most often influenced by direct or indirect experiences rather than reports and evaluations.

What Counts with Congress?

When Congress sets out to make major policy decisions, the most important variable is the personal values of six to 10 congressmen and their staffs, a top congressional aide told education researchers gathered in San Francisco this week.

In a candid discussion before a sparsely attended session on "Designing Educational Research to Reflect Federal Policy Needs," Minority Staff Director Robert Andringa of the House Education and Labor Committee made no bones about the gulf that separates policy makers and policy researchers. After a little impromptu research of his own, Andringa said he came up with this rank-ordered list of what counts most heavily in the congressional decision-making process on issues such as education:

- (1) personal values of six to 10 members of Congress and their staff,
- (2) strong views of respected and trusted friends,
- (3) assumptions about the budget,
- (4) public opinion and the popular media,
- (5) consensus among major interest groups, such as education associations,
- (6) program data, (descriptive data about who is served, etc.),
- (7) GAO studies.
- (8) the hearing process,
- (9) policy research
- (10) administration views and lobbying, and
- (11) program evaluation.

Educators are generally among the least effective witnesses on Capitol Hill, Andringa told the participants attending the panel session of the American Educational Research Association convention. "Evaluations tend to be untimely or too long or unreadable or in the wrong format," he said, urging his listeners to "use bar graphs and pie charts." Andringa said the researchers often seem more intent on impressing their colleagues than on helping lay people make wise decisions.

(From Education Daily, April 23, 1976)

The point is that decision-makers, unlike computers, are most influenced by their experiences and feelings (or views, assumptions, etc...), and the feelings of people around them, than they are by statistics. So your goal as manager and evaluator of your project is to gather and present information in a way which affects the way people think and feel, so that decisions they make will be in the direction which you think best.

The reason we call this process evaluation, and not propaganda, is because the foundation of your effort is the gathering of honest information--not the propagation of lies. But it is critical for you to realize that the effects of evaluation and propaganda are the same--they influence people's decisions.

On the next two pages are some questions that you and your staff may find useful to identify the various persons who influence, or are influenced by, your project.

"Statistics are no substitute for judgement."

--Henry Clay

- If your project were completed, what tangible services and/or products would exist or have been provided? Who would use or have used each product or service? Who would be ultimately affected by each product or service?

<u>Products/Services</u>	<u>User</u>	<u>Ultimate Target</u>
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____

- To what extent have the users or ultimate targets of your services been involved in your project?
 - ...in proposal development? _____
 - ...in a project advisory role? _____
 - ...in providing other feedback? _____
- How do you use the information you get from them? Do they know this?
- Who are the people who are totally committed to the success of your project?
 - .. on your staff? _____
 - .. in your organization? _____
 - .. outside your organization? _____
- Who might like you to be successful because it will facilitate some goal of theirs? What are their goals or purposes to which you relate?
- Do you know what you and each staff member "wants" out of the project?
 - .. How do you see your role as contributing to the success of the project?
 - .. Do others on the staff know this?
- How will you know you are being successful?
 - .. Who has the information you need? _____
- Who has information you need in order to be successful?
 - .. about the problem or need you are addressing? _____
 - .. about the relevance of your product or service? _____
 - .. about the effectiveness of your product or service? _____

● How often do you need the information from these people? _____

.. Why would they want to share that information with you? _____

.. How can you get it? _____

● What information mechanisms are you using?

.. Are staff meetings for giving information or getting it? _____

.. Are advisory meetings for giving information or getting it? _____

.. Do you ask people to fill out forms? How do you feel when you do this? _____

.. Do they know how the information is going to be used? _____

● Who is going to be judging your success? (Audience for your evaluation)

USOE? _____ You? _____ Others? _____

Your institution or agency? _____ Your staff? _____

Your community? _____ Your professional peers? _____

● What will each audience be looking at? your services or products? your project processes?

<u>Audience</u>	<u>Services/Products</u>	<u>Project Process</u>
_____	_____	_____
_____	_____	_____
_____	_____	_____

● Upon what information will they be basing their judgments?

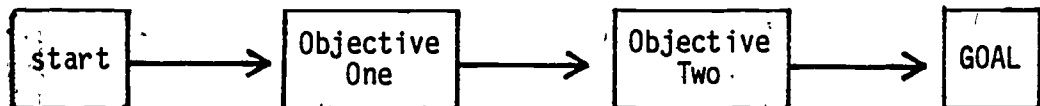
● When will they want the information? At the end of the project? Throughout the project? Quarterly? Other times?

● Why does each audience need the information?

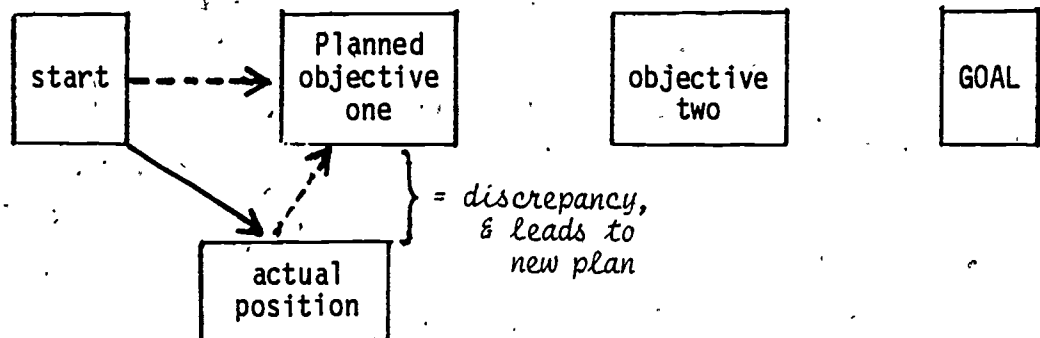
Information in Project Problem-Solving

Your task in managing the giving and receiving of information would be relatively simple if it were not for one factor. For some reason, when we approach the varying components of management--planning, operating, evaluating--many of us appear to forget much of what we do with information in other aspects of living. Almost mindlessly we assume that a different form of behavior is required. We begin to restrict information to symbols that can be recorded on paper; and tend to deal with it as an end in itself, apart from the impact it has on the experiences of those who receive it and give it.

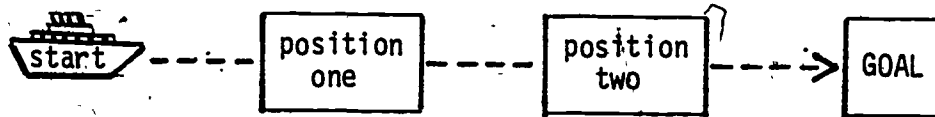
As an example of how this cultural blindness limits the information we can get for our management purpose, and which we can give to others for theirs, we might first look at the simple task of planning. We tend to deal with planning as a "straight line" process. That is, we have a starting point, a goal, and a series of intermediate objectives to accomplish it:



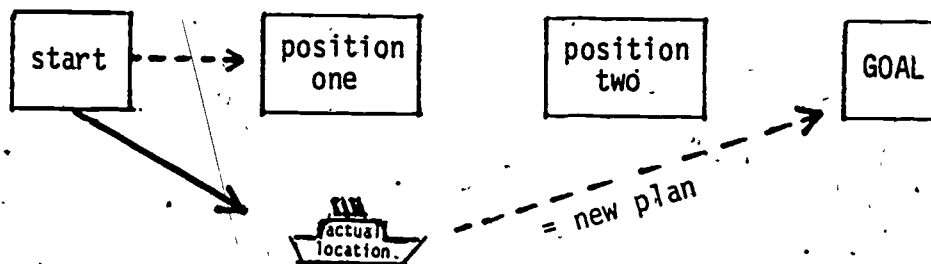
When, in the course of program operation, we reach the time that the first objective was to have been accomplished, we check to see if we have reached it; and if we have not, we label this a "discrepancy" and try to figure out how to achieve the intended objective:



This may seem logical, but think for a moment about how planning works in a task such as navigation. Here, too, the navigator begins with a "straight line" plan



...but when the first position, or check-point is reached, and he or she discovers that unanticipated wind, currents, wrecks, etc., have taken the ship off its planned course--a new course (or plan) is developed--but not to the old check point. Instead, the course is planned to the original goal from the point at which the person now is:



This operation continues as the navigator continues to take sightings every few hours to determine where the ship actually is, and to re-plan its course.



In looking at the actual zig-zag course that brings the navigator to the goal, it is possible to suggest that this planning process is based upon a different set of assumptions than those we use in most organizations. For example, it appears that:

- Knowing where you are is more important than knowing where you thought you'd be.

- There will always be unanticipated events or other things over which you will have no control.
- Discrepancy information is positive data that first tells you where you are relative to where you want to be, and then can help you identify forces that can be harnessed to achieve your goal.
- The more frequent the checking (evaluating your position) the less chance of deviating too far off course.

Interestingly, all of these assumptions can be equally true in project management, yet they are seldom used. One reason for this may be "fear"--fear of the discrepancy, of being judged wrong, of not achieving expectations. Unfortunately, the loss that results from this fear is one that affects more than the project itself. For the information inherent in the discrepancy--where we are, how did we get here, what happened that was unanticipated, etc.--is in effect the description of an experience. And experiences are not merely judged good or bad, and then passed over. One can learn from either one. When this information, however, is not gathered and shared because of "fear," then little or no learning takes place.

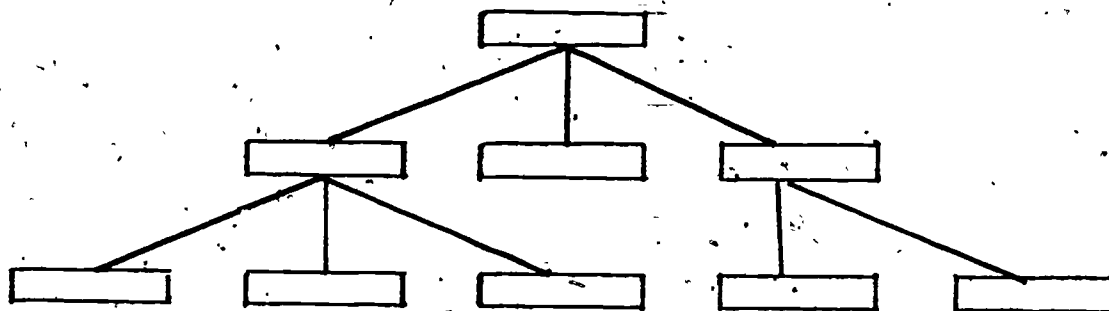
THE ROAD TO WISDOM

The road to wisdom? - Well, it's plain
and simple to express:

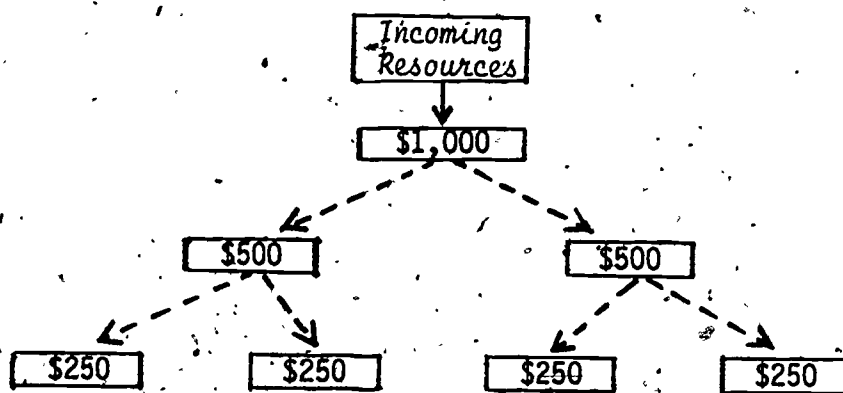
Err
and err
and err again
but less
and less
and less.

*Piet Hein
Grooks*

A second example of the limits we impose on ourselves by our perceptions of management is evident when we try to use an organization chart as a way to identify information needs and/or channels for communicating information.



These very familiar "maps," existing if not on our walls then at least in our minds, portray what we believe to be the relationships in an organization. And they do, but only one type of relationship--those that relate to decisions that control resources coming into the organization. The "higher" in the pyramid, the larger the amount of resources affected by a decision-maker.



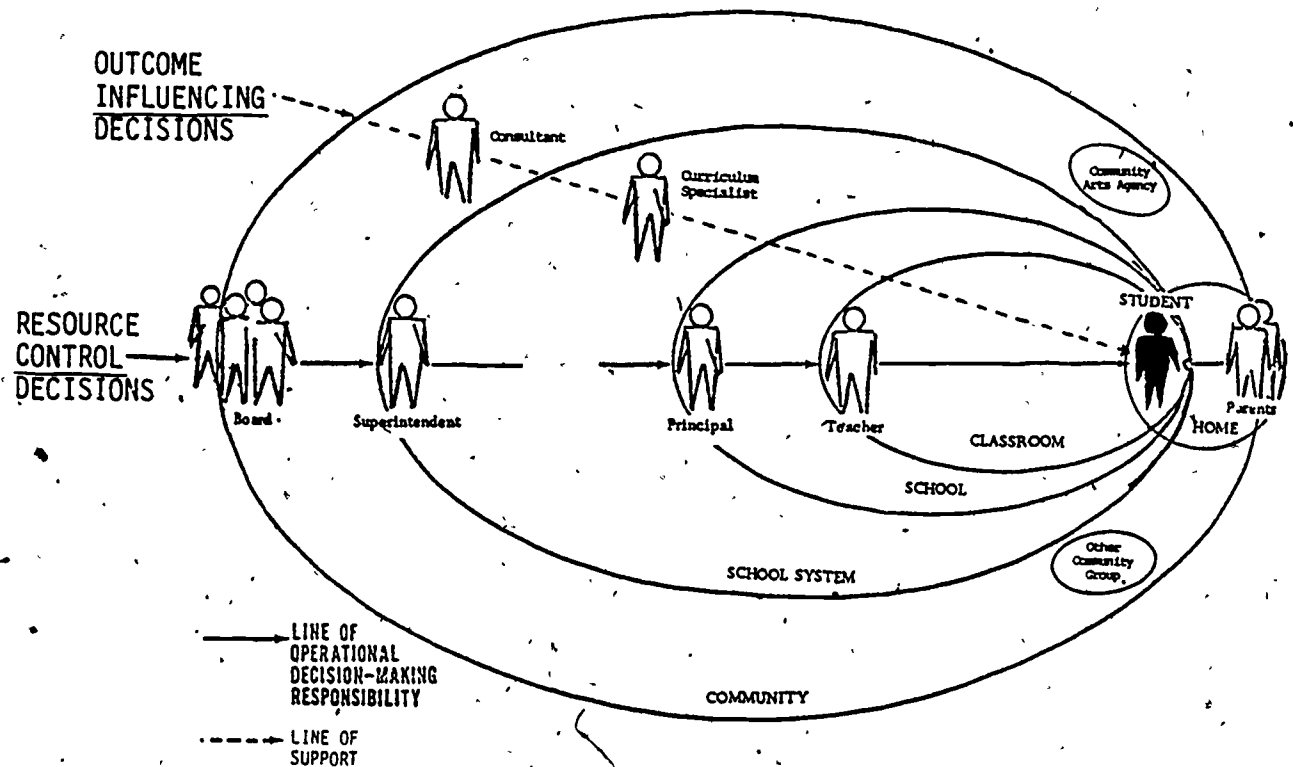
Organization charts, however, are difficult to use for portraying how decision-makers relate to each other in terms of the outcomes of the organization--which in most human service organizations involve some type of influence on another person. But if we think of an organization chart as a form of map, then we see that it can be re-drawn without necessarily changing reality--just as mercator and polar projections portray the same

reality while appearing to be different--the difference resulting from their being drawn to different reference points--the equator and the poles.

In the same way, if you re-map your organization in terms of a different reference point--that is, to portray how peoples' decisions relate to or affect the outcomes you desire--you may be surprised at the number of potential allies and supporters for your project that you identify, as well as the opportunities you perceive for building links, with information, to these others who share a mutual interest in your project's outcomes.

The following "map" is one way to portray these relationships in a typical school situation. It has these characteristics:

- each decision-maker in the total organization is displayed with reference to the student; yet the traditional "resource" relationships are not changed;
- the "space" that each decision-maker primarily influences with his or her resource decisions--e.g., classroom, building--is shown in a way that includes the "spaces" of others that are influenced by those decisions;
- each "space" also coincides to a great extent with that person's job-related sense-of-self. This is where each decision-maker strives each day to meet his or her needs for a sense of worth, achievement, power, etc. In other words, within this space is where the element of job satisfaction is achieved or is potentially available;
- those professionals who are not in direct contact with students or teachers, such as curriculum developers, are arranged on a line of influence that relates them to the outcomes they support but clearly shows that they do not share in the "power" of controlling the major resources of the organization;
- other groups or individuals outside the school which relate to the outcome can be shown. These can include the home, community organizations, other projects within the school system, etc.



Using a format such as this to plot the relationships of the persons within your project provides several advantages:

- You can identify those individuals who need information from you to accomplish their purposes, or from whom you need information to accomplish yours. In your project what information is this? When, or how often, is it needed?
- You can deal openly with the fact that each person in a program has dual motivation--effective service and self-interest. The knowledge that self-interest intrudes on almost every decision is increasingly being acknowledged in research, and in fact, satisfaction is now identified as a key element in the success of any program. What types of information would contribute to a sense of satisfaction for you? For your staff?
- If your program exists on the line of "influence" rather than "resource control" you may be able to see that attempts to achieve outcomes at the student level without involving those who make the everyday resource-decisions are bound for eventual failure. You may want to draw lines to those decision-makers for whose problem-solving decisions you could offer support. What types of problems do they have to deal with? How does what you offer relate to these problems?
- You can draw lines to those other agencies or individuals outside your organization that share interest in your outcomes. What types of information do they have that could support your objectives? How can they be involved so that you have access to them?

Within a goal-centered perspective such as this, you can begin to identify the various sources of the information you need to document or gather for achieving your project's objectives. You can also identify the information you have which could be useful to others in accomplishing their purposes. This latter information is the "content" of what is usually called evaluation or dissemination.

"It is not true, as a good many industrial psychologists assert, that human nature resists change. On the contrary, no being in heaven or earth is greedier for new things. But there are conditions....:

- the change must appear an improvement;
- it must not be so rapid or so great as to obliterate the psychological landmarks which make a man feel at home;
- nor obliterate his understanding of his work, his relations to his fellow-workers, his concepts of skill, prestige and social standing in certain jobs."

--Peter Drucker

III, REPORTING AND DISSEMINATION

Reporting

Before you start thinking about any required reports you have to produce as part of your project grant, are you sure you know what is expected from you? It was found at the workshop that many Project Directors believe that more is required by the funding agency than actually is. For example, consider the following excerpt from the Evaluation Monitoring and Reporting Requirements of the DHEW/USOE General Provisions:

§ 100a.276 Evaluation.

- Each project shall include procedures for effective evaluation of the extent to which project objectives are being met. (20 U.S.C. 1221c(b) (1).)

Subpart Q—Monitoring and Reporting of Program Performance

§ 100a.430 Scope of subpart.

This subpart sets forth the procedures for monitoring and reporting program performance. These procedures are designed to place greater reliance on recipients to manage the day-to-day operations of their federally-supported activities.

(OMB Circular No. A-102, Attachment I.)

§ 100a.431 Monitoring by recipients.

Recipients shall constantly monitor the performance under federally-supported activities to assure that adequate progress is being made towards achieving the goals of the project. This review shall be made for each function or activity of each project as set forth in the approved grant application or contract document.

§ 100a.434 Significant developments between scheduled reporting dates.

Between the scheduled performance reporting dates, events may occur which have significant impact upon the federally-supported activity. In such cases, the recipient shall inform the Commissioner as soon as the following types of conditions become known:

(a) Problems, delays, or adverse conditions which will materially impair the ability to attain the objectives of the grant or contract. This disclosure shall be accompanied by a statement of the action taken, or contemplated, and any Federal assistance needed to resolve the situation.

(b) Favorable developments or events which enable meeting time schedules and goals sooner or at less cost than anticipated or producing more beneficial results than originally projected.

In addition to these general requirements, the USOE Arts Education staff have indicated the types of information which are useful to them, as in the following excerpt entitled "Final Reports" taken from a USOE Staff Memo to the 1977 projects:

The staff at the national level responsible for administering the program will not be able to visit many of the 77 State and local projects, much as we would like to. We learn something of what you are doing by talking with you and your associates by telephone and at conferences, but we need to add to this information. The reports will help us answer the many requests we receive for details on the grants.

An informal narrative report will suffice. No specific form need be completed, nor are you required to supply specific statistical data. (However, if such data are available we would appreciate your including this information.)

We would like you to include in your report:

- a brief summary of specific activities the grant supported
- the project's strengths and weaknesses
- major problems and/or obstacles encountered during the year
- issues (those which emerged during the year as well as those present at the beginning of the project)
- project outcomes (progress, gains, compromises, implications for the future)
- specific plans for the future (for arts education in your school, district, or State; for the project or its goals)

We are interested in what you feel the project accomplished (whether or not these accomplishments relate to the original proposal). We hope you will consider the report an opportunity to record your own reflections, intuition, and feelings about the project, as well as the experiences of students, teachers, and administrators involved.

In addition, we would appreciate copies of the materials which the grant helped support; i.e., curriculum materials, in-service training materials, comprehensive plans, resource guides. Do not send but let us know if you developed slides, films, videotapes, or other audio-visual materials with grant funds.

It is obvious from the above that the types of information that persons operating at the Federal level most need include many of the feelings and learnings that do not fit easily on checklists and one page forms. This creates a problem for Federal programs because they are constrained from creating "undue response burdens" on local programs. Earlier we defined response burdens as the "difference between a Federal or State decision-maker's desire to have good information and the practitioner's desire to provide it." On the one hand, today, the awareness of response burden has led to shorter and less frequent report forms in many programs. But on the other hand, it has also slowly cut off the flow of learnings and other experiential information that can let Federal personnel play roles more helpful and supportive than just being compliance monitors.

One way to deal with this situation has been suggested in the proposed new regulations for the USOE Arts Education Program:

Subpart E--What Conditions Must Be Met by a Grantee?

§ 161c.4) Ongoing evaluation activities.

(a) Throughout a project period, a grantee is responsible for evaluation activities that provide information for effective management of the project, including--

(1) Documentation of all activities for self-monitoring purposes; and

(2) Collection of statistics and data they may feel is needed to assess progress and develop local support; for example--

students served, student and teacher achievements and attitudes, and relationships between other academic subjects and the arts.

(b) The Commissioner may require that some of the information in paragraph (a) of this section be included in project reports.

(c) Not more than 10 percent of total project costs may be used for outside evaluation of the project.

(20 U.S.C. 2962)

Here, in these new regulations, the program is articulating the principle that all reports are derived from information which, in the first instance, must be gathered by the local project for its own management purposes.

Dissemination

Dissemination of information about a project's accomplishments is a requirement of most Federal demonstration programs. It provides a way to maximize the investment that the government makes in the development of new ideas. In addition, most project directors want to share information about what they are doing, especially with peers and others operating similar programs.

Why then have the results of so many dissemination efforts been unsatisfying?

The idea of experience-sharing among individuals or projects in a Federal program is a deceptively enticing concept that would work ...

- If individuals could be aware of what is happening while it is happening.
- If experience and information were synonymous--that is, if the actions and thoughts of individuals, as they dealt with the major tasks of operating a successful project, could be translated into a useful and communicable form of information.
- If sharing weren't left to the "good intentions" of parties whose intentions soon get overridden by the daily problems of running a project or program.

- If the sharing could take place while the problems were current.
- If individuals felt it was as safe to share the critical information about what did not work as about what did.
- If the fear of evaluation and judgment could be removed from the process of providing information to others in a program, especially those in higher positions.

We have been and will be discussing ways to identify, and document this type of information. In this section, however, we will focus on dissemination as the movement of information among those who have the "questions" and those with the "answers."

Although this may be of more interest for State projects, or others with the responsibility for maintaining dissemination networks, the issues discussed also influence the way a specific local project can package and share information about its work.

There are three issues that need to be considered when planning dissemination--Why disseminate? What to disseminate? And how to do it?

Why Disseminate?--The Adoption/Adaptation Issue

Leaving aside for a moment the grant requirements to disseminate, it may be important for you to clarify for yourself how you feel about why dissemination is important.

True or False? The reason to disseminate is so that others do not have to re-invent the wheel!

Your answer will put you somewhere on a continuum between those who believe that good materials and ideas should be adopted without change, and those who believe that they must always be adapted. Or, said another way, between those who believe that the "wheels" are more important than the process of creating them, and those who believe that the satisfying process of creative adaptation is an important element in teaching and the professional growth of all educators.

What to Disseminate?--The What We Did/How We did It Issue

Another consideration related to dissemination relates to the type of information most helpful to others. Formal dissemination usually

encompasses what might be termed the "WHAT" information--what happened, to whom, with what effects? etc. On the other hand, the "HOW" information--how you did it, what worked, what didn't work--appears to be the information most needed by local programs. This latter information is the type that is exchanged, usually informally, through meetings and phone calls. Yet most projects do not have ways to systematically capture or document this "HOW" information while it is still fresh. (If you are interested in one way to address this problem, see the discussion of the Growth Record in the later section about documentation.)

How to Disseminate?--The Push/Pull Issue

A final consideration addresses the issue of how information gets to people. This might be envisioned as a "PUSH-PULL" continuum. On the one end are those who see the role of dissemination to be the "pushing out" of information to people who are perceived as needing it. On the other end are those who believe that dissemination systems should allow people to "pull" or get information only when they need it.

Actually there is no right position on the three issues dichotomized above. Sound dissemination practice requires that all six of the elements be taken into account in any dissemination system.

A process cannot be understood by stopping it.

Understanding must move with the flow of the process, must join it, and flow with it.

(First Law of the Mentat)
Frank Herbert
Dune

Packaging

There are also some considerations to be taken into account when packaging information about your project:

Multiple audiences

The more different audiences you envision for your material the less the chance you will communicate effectively with any one of them. If possible, develop separate materials--booklets, slide tapes, reports, etc.--tailored specifically to the audience you desire. Sometimes this only involves the re-ordering of information you already have developed.

Media

Using non-print media for your reports can have two effects. It can make it possible to communicate imaginative and experiential parts of your program. On the other hand, it can limit the access to the material. In choosing the media for your communication, consider whether or not your desired audience has access to the playback equipment that will be needed.

Content

The one thing you can be sure you have in common with anyone who will find your experiences useful is that you both are responding to some similar problem; e.g., integration of arts into the classroom. Therefore, a useful format many times can be one that focuses on problems and how your project responds to them. Ask yourself, if someone were visiting the project what would interest them the most? Use this in your report.

People have different abilities and different experiences which influence the way they receive and use final reports. These differences must be taken into account when interpreting findings and reporting information. For example, reporting achievement tests scores may be meaningful to some audiences, while others may be alienated by the use of such information.

The primary determination that should be made in the development of a report is to understand the criteria that will be used to judge a program or product.

Below is a summary of ideas about reporting that Project Directors offered during the course of the workshops. You may want to compare them to yours.

What Information Do Audiences Ask for About a Project?

- an abstract
- who was the target group?
- what was the objective?
- what strategy or technique was used?
- what features are different from other projects?
- cost: start-up; special equipment or staffing needs? continuation?
- evidence of effectiveness:
 - How was it tested?
 - On whom?
 - How many were tested?
 - How were they selected?
 - By what measures were they tested?
- responses to testing:
 - What type of responses occurred?
 - At what frequency?
 - Within what range of possible responses?
- crucial elements:
 - Which did you find essential?
 - When did you employ them?
- possibilities for modification:
 - What alternatives were tested?
 - What alternatives are suggested?
 - What problems might be encountered?
- prepared materials:
 - Clear, detailed instructions to primary user
 - Separate manuals/instructions for other users
 - Project brochure
 - Self-contained modules
 - Compatibility with existing programs or schedules
 - Public information booklet(s)

What are the various types of reports that can be considered for different audiences?

- written?
- oral?
- general?
- specific?
- technical?
- nontechnical?
- descriptive only?
- evaluative and judgmental?
- makes recommendations?

What Modes of Display Can You Consider?

- case studies _____?
- portrayals _____?
- graphs and charts _____?
- test score summaries _____?
- scenarios/anecdotes _____?
- questions/answers _____?
- product displays _____?
- dialogues/testimonies _____?
- photographs/newspaper clippings _____?

IV. EVALUATION

What Evaluation Is

Evaluation can be defined as looking at what you did as it affects clients, staff and environment--and making a judgment as to its worth for you and others. A breakdown of this definition provides further explanation:

"...looking at what you did..."

The first feature of any evaluation is documentation of the process in an on-going fashion (what has been called "formative evaluation.") When documentation is conducted in this way, a program manager can make assessments at any point in a project's life. The more detailed this documentation, the greater the understanding of the process for both the program staff and external observers.

"...as it affects clients, staff, and environment ..."

An important feature of any evaluation is the ability to show change in behaviors or attitudes as a result of an innovation. The degree of change or effect may be small, but noteworthy. Effects may be due to many things; the intent here is not to "prove" that what you've done has created a change, but rather provide evidence of differences which now exist in the environment since the program or innovation has been introduced--so-called "before and after" information.

"...making a judgment..."

Evaluation always implies judgment. Whether we like it or not people will judge our projects. As one evaluator, Michael Scriven (see his evaluation model on page 48) has put it,

"...if the goal of evaluation is to determine worth, then, to collect data, focus inquiry and inspect causal relationships without making a judgment of worth is not evaluation."

This judgment can be an informed one, based on documented information, or the judgment can be made without benefit of documentation, or of

information. When a program is evaluated, we seek to influence these judgments by providing evidence of success, change, and/or differences.

"...as to its worth for you and others..."

Evaluation is always subjective. Evaluation is a necessary part of the decision-making process, a process which utilizes information as well as the values and attitudes of the decision-maker. Another "evaluator," Laura Chapman, writes that

"...the purpose of evaluation is to discover the value or significance of something. Measurement is not an integral part of evaluation; it imposes on the concept of "worth" an ordering and ranking system which leads to a denial of intrinsic value. An evaluation of an arts program should be empirical, valid, broadly reliable and responsive to the qualitative character of art experience."

But, if it is important to "sell" a program to the "outside," one must acknowledge the criteria for success used by such external audiences and provide information (evidence) as to how the program has achieved that success.

Conducting an Evaluation

How do you go about conducting an evaluation? You can conduct an evaluation as others have done, using established evaluation "models." You can design your own evaluation technique which will satisfy the needs of your program, or you can use a combination of your own model and those of others. The important thing is to select an evaluation strategy that you are comfortable with and which satisfies your own decision-making needs.

Evaluation can assist in the decision-making process for

- program planning
- program improvement
- program justification.

When evaluation is used for program planning decisions it has several features. First, evaluation can help to identify needs within the program that are contingent upon the ultimate success of a program. Many times within the course of running a program, even if a formal needs assessment has been conducted, issues or needs will arise that must be addressed in order to accomplish program objectives. By establishing a formative feedback evaluation system for decision-making, these needs can be identified early in the program and steps can generally be taken to satisfy these needs.

Secondly, evaluation can help to identify the discrepancies between intended and actual program elements (see the discussion of the Discrepancy Evaluation Model, on page 46). When a project is conceptualized, we identify outcomes and the processes necessary to achieve these outcomes. In the course of running a program, we implement the processes in the anticipation of the desired results. Because programs are made up of people as well as processes, often what we suppose will happen, does not, in fact, occur. Sometimes this realization does not take place until late in the project cycle, or if the realization occurs earlier in this cycle, we feel locked-in to our process and unable to make necessary changes. Evaluating the process can assist here in two ways--it can provide early identification of discrepancies and it can also provide evidence for the need for change to both internal and external decision-making bodies.

Finally, evaluation is useful in the planning process, because it can serve as a checklist to the decision-maker to assure that the steps are being taken which will lead to project success (this, again, is what is meant by "formative evaluation").

Evaluation is also used in making decisions about program improvement. Management is a constant state of trying to improve and attain greater levels of satisfaction among staff and constituency. Evaluation can be used to develop change strategies and alter the course of a program to this end.

But evaluation, in the end, is most often used for program justification. Why should this program continue? When we need to justify a program we seek evaluation information that speaks to consumer utility and cost-efficiency.

Understanding why we conduct an evaluation for planning, improvement, or justification will help us in designing evaluation techniques and reporting findings.

Steps in Evaluation

There are four basic steps in conducting an evaluation:

- I. Identification of issues (what needs of the audience must be addressed?)
- II. Identification of information sources (documentation)
- III. Identification of a strategy (either an "evaluation model" or a design of your own)
- IV. Identification of resources (staff and money available for conducting the evaluation)

These four steps will be discussed more fully in the sections which follow.

Identification of Issues (Audience Needs)

Who is going to be judging your success?

- USOE _____?
- Your institution _____?
- Your community _____?
- You _____?
- Your staff _____?
- Your professional peers _____?
- Others _____?

In designing an evaluation, it is important to understand not only why you are collecting information, but also for whom you are collecting information. The audiences of your evaluation are those persons and organizations who will be making decisions about your program.

What will each audience be looking at?

- o Processes--what was it you did?
- o Cost--how much?
- o Consequences--how did your project affect your community, or the future of your students?
- o Change Factors--what appears to be the reason for changes which occurred?
- o Client Satisfaction--did the students like the program?
- o Cognitive or Affective Change--were there changes in what people have learned, or in ways in which they feel?

Whether or not you want certain groups or organizations to make judgments about your program will not prevent such judgments from being made. When you undertake the process of evaluation, you, as a program manager, have made a decision to "inform" these judgments. By providing your audience with information that is directed to their needs, you give them concrete evidence upon which to base a decision. Remember that just as in a court of law, evidence here can be considered as inadmissible and circumstantial to some, and as concrete proof to others (see the discussion of the Jurisprudence Evaluation Model, following on page 54). So another feature of an evaluation plan is the determination of what evidence will be acceptable to your audience.

How will you know and prove that your project has been successful?

- o identify the criteria by which different people make judgments
- o identify the "political atmosphere" in which the decision-makers operate.

While you and I have lips and voices which are
 kissing and to sing with
 Who cares if some one-eyed son-of-a-bitch invents
 an instrument to measure spring with.

--e.e. cummings

Identification of Information Sources:
(Documentation)

Following is a series of questions which deal with the information you will need to collect for your evaluation and where this information can be found. The sections on documentation which follow give a more complete description of how to collect and analyze information.

- What information mechanisms do you use?
 - staff meetings--for giving or getting information _____ ?
 - advisory meetings--for giving or getting information _____ ?
 - Are people filling out 'forms' for you? Do they know why? _____ ?
 - Who has the information you need to be successful _____ ?
 - How often do you need this information from these people _____ ?
 - Why would people want to share this information with you _____ ?
 - How can you get information:
 - discussions _____ ?
 - observations _____ ?
 - correspondence _____ ?
 - written materials _____ ?
 - meetings _____ ?

It is important for the success of Arts Education programs, and the evaluation efforts that parallel them, that documentation and evaluation processes not be confused. Regardless of evaluation type or strategy--internal or external, project or program, process or product--what is important is that the information upon which those judgments and decisions will be based comes, in the first place, from the experiences of the people at local and State levels. Documentation--the translating of those experiences into communicable information--is a prerequisite for both good management and good evaluation.

It cannot be assumed that information and experience are synonymous--that those who are in the middle of an experience are aware of what is meaningful in it (and if they were, that they could communicate it to others in a helpful way). Therefore the responsibility

for documenting what is happening and what decisions are being made should become a regular function of project operations, not something you "assume" will be done by someone else, or by yourself, only later

Documentation might be considered "process observation," but it is important to differentiate between the "process observer" role in traditional research, where an individual, working independently, gathers "objective" data on what is happening for later external analysis, and the process observer role in management. In the latter situation, the role involves the whole project team. Using various appropriate methods the Project Director, or another individual, increases the awareness of the staff to the experiences which the program is undergoing, why and how things are happening, and why various decisions are being made--and this person helps the staff to use the information they generate for project problem-solving and decision-making.

An additional and valuable use of this information, of course, is for "evaluation." While it is critical that Project Directors and staff have access to current information in order to run a project, it is equally important that they be able to turn it into timely reports for different audiences of internal and external decision-makers whose judgments can influence the project's success and ultimate existence.

The task of gathering information about the occurrence of events is relatively simple compared to that of getting the data about the relationships between and among events, and about the processes that precede them. It is not that the information does not exist, only that it is stored in the "experiences" of those who were the participants. They are frequently not aware of all of its meanings nor of its value to those who might be looking at the event from another perspective.

The nature of this experiential information is what makes it so valuable as a target of the documenting process--it deals with how and why things happened. Since the primary interest and emphasis of the Arts Education Project Directors' Workshops dealt with the "communication of experience," the Growth Record Process was presented as one way that Project Directors might find useful to translate regularly the experiences of the project participants into useful and meaningful information.

The Growth Record Process: Planning Ahead is a Looking Backward Process

As noted earlier, it has been our experience that there needs to be a distinction made between the role of documentation in research and in management. What makes this particularly important are the perceptions held by many managers that evaluation is an external process done by and for someone else; a process that makes judgments which are often based upon sketchy or incorrect information about what really happened.

As evaluators in one federal program noted,

"Our pervasive impression is that evaluation is not a very well-liked part of the program in many sites ... it still has a slightly unnatural feel in the base of project operations. It retains a flavor of an "add on"--of "something being done to us." It is extremely difficult to govern and conduct evaluation so that participants share in its ownership.

For these reasons, Applied Management Sciences advocates experimenting with an approach to documentation called the Growth Record Process. The fundamental concepts in this approach are:

- understanding of an experience--what happened, why and how--can best be derived after it is over.
- To do this requires a reference point to focus your awareness--something to look back at to create a structure for the review. "Planned" objectives serve this purpose well; for example, "What did we intend to do? How? What actually happened? Why? What were the influences that affected the outcomes?"
- The more points of view, or experiences of a situation that are considered in the review process, the "better" the picture of what happened. This also addresses the concern for a documentor's objectivity--the problem of self-selection of information. The more persons involved, the more likely that one person's values won't hopelessly bias the evaluation.
- This process of identifying what was important is an experience, analyzing it, making generalizations and documenting these learnings so that they can be used in future actions lets a staff or board experience the "ah-ha's" of discovery and learning that external evaluators usually reserve for themselves. This type of review should follow each major project activity or event, and be done at least once a week, e.g., at a staff or board meeting, to pick up other aspects of project development.

- Participation in this review process provides a way for various role groups, or agencies, involved in a project to collaborate on a meaningful task; that is, to contribute to the solution of problems that relate to their mutual concern. The "meeting," therefore, as a problem-solving forum, serves as a major source of process data for the documentation.
- The frequency (suggested as once a week) of the "process" reviews may seem too high to one who is accustomed to a culture which already seems to have too many meetings already built-in.

An analogy for this type meeting, however, can be drawn from the process of navigation, discussed earlier. Even though the original course (plan) is laid out as a straight line between two points, the navigator is accountable for continuous and frequent checking to determine where they actually are; what unanticipated forces caused them to be there; and to suggest a new course based upon where they are now which takes into account the previously unanticipated influences.

In this process, "knowing where you are" is more important for planning than knowing where you thought you would be. Checking frequently is essential because the longer the time between check points, the larger the possible "error," and the possibility of forgetting some of the dynamics of the events.

The strategies that support this process consist of a functional role for a documentor and self-monitoring "forms" which permit a project to document its own growth. A discussion of these two elements follows.

The Role of the Documentor

It is helpful to have one person accept the responsibility for the documenting process. This might be the Project Director, or it can also be a member of the project support staff. The role of the documentor can be perceived as that of an internal facilitator for the flow of information needed for project management.

In a previous application of this strategy in a national program, such persons were called "information facilitators." As described in that program,

his/her principal task is to assure that adequate information is generated and fed back into the decision-making process within the project, to "mirror" what is going on, and to facilitate the asking of "better" questions. To do this requires a recognition of this role in the project; a sensitivity to the needs and strengths of project personnel; and a commitment to the project's approach.

Other important elements of this person's role include serving as an interface between the project and external information requirements, i.e., primarily OE through its quarterly reports; and serving as a linkage agent to facilitate sharing of experiences among projects.

Final Report on the Development of the NDEP
Information System, OEG-0-70-2068, Sept. 1973.

The role of the documentor responds to one Federal program's recommendation "to make evaluation so much an integral part of the project that it becomes most indistinguishable ... so that information from participants can be gathered more easily and all project personnel can become accustomed to a regular feedback process and to the utilization of data for continuous program improvement." This documentor role helps everyone to be a process evaluator and therefore can remove much of the fear of evaluation as an external process.

The Project Growth Record: A Way to Translate Experience into Useful Information

The pieces of paper that we call the Project Growth Record are designed to provide a project and a documentor with a means for generating and collecting experiential information about how a project accomplishes its tasks; what it has to overcome in doing this; what it learns in the process; and how it applies the learning. It is both a planning and a self-reporting system that surfaces problems before they become unmanageable, and then challenges project staff to look for alternative ways to deal with the problems. It is not an "information system" but can interface with whatever formal or informal management information systems are used in a project.

The Growth Record Process is designed to document the outcomes of a project's ongoing decision-making processes, as opposed to a process which requires that project staff produce and send data to others for analysis before getting any value from it.

Copies of forms follow, with a flow chart that illustrates how they are used by project staff to look back at their month's experience, to learn from it, and to re-apply the learnings.

1

PROJECT: _____

TASK/OBJECTIVE: _____

EXPERIENCE REPORT

For the period of: _____

Describe the principal actions or events which took place during this period.

What were the observable indicators that each action had the effects you anticipated?... or that it had other effects?

37

43

44

2

PROJECT: _____

TASK/OBJECTIVE: _____

In Comparing your actual EXPERIENCES
this period with your plan, what
things happened that you might not
have foreseen?

LEARNINGS

For the Period of: _____

HINDSIGHT

What actions planned during this period did not turn out as expected? What caused the difference?

38

SERENDIPITY

What unexpected favorable developments occurred during this period?

GROWTH

What learnings can you derive from these experiences that can be applied to your PLAN for the next period? ... or can be shared with other WEEA grantees?

45

46

3

PROJECT: _____

TASK/OBJECTIVE: _____

_____ 7

POSSIBLE PROBLEMS & RESOURCES

For the period of: _____

What special OBSTACLES do you anticipate encountering during the next period?

What RESOURCES are there within your Project for dealing with them?

What RESOURCES are there outside your Project for dealing with them?

39

HOW do you plan to address the above OBSTACLES?

47

48

4

PROJECT: _____

TASK/OBJECTIVE: _____

ACTION PLAN

For the period of: _____

What are the major tasks that you plan to undertake this coming period?

At the end of the period what differences in Conditions, or in the Actions of people will satisfy you that this activity had the effects you desired?

40

49

50

4

ACTION PLAN

Awareness of an experience creates "tension." The **Action Plan** creates an "existence" which can be reviewed at the end of the period. Using the Action Plan as the focal point for review serves as a catalyst for awareness and openness to experience.

Success indicators, rather than objectives, are used for two reasons. First, they serve as the means to synthesize both task AND personal objectives, thus creating a reference point that is closer to the reality of the way these human beings function. Secondly, they avoid the backlash of the accountability movement in education that has resulted in individuals feeling that objectives are something that someone else used to judge them.



1

EXPERIENCE BEFORE

--	--

The process uses short planning cycles in the "trial and error" approach so that discrepancies can be identified before they grow into larger problems.

The project uses the original plan (4) as the reference point to describe its experiences during the month and the actual outcome achieved.

2

LEARNINGS

While comparing their actual progress to their original expectations, certain conditions can become visible which may be addressed in the next month's plan.

3

PROBLEM & RESOURCE ANALYSIS

This analysis serves as a way to look for alternative approaches

4

ACTION PLAN

Next month's expectations are documented to begin initial awareness of new experience.

The structure of the forms and the wording of the questions asked can vary from the prototypes shown on the previous pages, but the basic principles are usually maintained. The formats provide a way for a project to document its decision processes for activities related to each major objective. These decisions are perceived as the outcomes of problem-solving processes that require experiential information from more than one person.

The decisions (or plans) are the last step in the process. The focus is on looking back at a previous plan (expectation) and comparing it with actual accomplishments in order to learn from the positive or negative discrepancies. It should be noted that this differs from Management-by-Objective-related discrepancy evaluation where discrepancies are "judged", evaluated, and commonly perceived as "bad." Here, instead, a discrepancy is defined as an experience (what actually happened and why). There is no such thing as a good or bad experience since it is possible to learn from both; and so-called "bad" experiences frequently provide more useful information than the "good" ones. It is also possible to pick up information about "lucky," or "unexpected" occurrences, which would be lost in a Management-by-Objective reporting, because they were never planned for.

The project documentor, utilizing this type of strategy, serves to facilitate his or her projects' ongoing evaluation and planning. Moreover, time is a most important dimension of these reports. No single report is significant except in terms of its relationships to earlier or later information. It is this "picture" of change or growth over time--and the reasons for it--that provides a "process" picture frequently impossible to capture through traditional research or evaluation. A periodic (monthly or quarterly) review across a series of weekly reports can permit a project to perceive patterns or trends that

may be significant. Further, it allows them to identify areas of concern for additional probing and data collection by the documentor. This analysis of activity at quarterly review points can serve to generate a great deal of the process (how and why) data needed by Arts Education project directors.

One more thing needs to be said about this process. Since it is based upon both common sense and evaluation-theory, projects may have already implemented many of the elements. One of the attractions for project personnel is that they really do not have to learn something new. Many Project Directors with prior management experience already will have informal mechanisms that they have used for their own process evaluation. It will be possible to use those as a starting point or to modify these suggested methods and formats to relate to their own pre-existing procedures.

Richard Leakey, "There have been thousands of living organisms", he says, "of which a very high percent has become extinct. There is nothing, at the moment to suggest that we are not part of the same pattern: He notes that there is one point of difference: man is the only organism with power to reflect on its past and upon its future. That power to reflect, he says, "is what makes us able to plan our future in such a way as to avoid what seems inevitable."

(p. 76 *Time*, Nov. 7, 1977)

Identification of Resources

(Staff and Money)

Who will conduct the evaluation?

Evaluation can be done by the Project Director or staff member, or by someone who is external to the program. Whoever is selected to do the evaluation should have a clearly-defined role. When a member of the project team is acting as evaluator, that should be known to everyone. If one chooses to use an evaluator outside the project, as in the hiring of any professional, standard procedures should be used. An examination of:

- practical background (What do they know about the content and context of your program?)
- educational background (What type of credentials are important to the audience?)
- philosophic orientation (How do they define evaluation? Does this match your definition?)
- methodology (Do they suggest a technique which is realistic within the limits of your project? Do you understand all the implications of this technique?)

It is important in securing the services of external evaluators to examine their evaluations of other programs, to understand the technique and philosophic approach of the evaluator. Calls to past clients are a good idea.

Who will be responsible?

- for the design of the evaluation _____ ?
- for the collection of information _____ ?
- for the development of instruments _____ ?
- for the analysis _____ ?

Responsibilities should be clearly delineated and should be divided among the persons involved, using the above tasks as a guide.

What are the cost factors involved in evaluation, in terms of money, time, and resources?

Evaluation should be considered a cost-efficient measure, not a costly one. Things to be considered are:

- What resources do you have available?
- Do you have an evaluation budget? (A good rule of thumb here is not to exceed more than ten percent of the total project budget)
- Have you accounted for program time needed to respond to evaluation requests? (Even if you employ an external evaluator, staff time will be expended.)
- Have you measured the potential impact of the evaluation against the anticipated expenditure? (Sometimes expending money at the beginning of the project will result in greater savings later in the project life.)
- What are the "costly" considerations, such as invasion of privacy, defensive attitudes of staff and clients, and difficulties encountered in information collection.

FIVE COMMANDMENTS FROM W.H. AUDEN*

Thou shalt not answer questionnaires,
or quizzes upon world affairs, nor
with compliance take any test. Thou
shalt not sit with statisticians, nor
commit a social science.

*From "Under the Lyre -- A Revolutionary
Tract for the Times"

Identification of a Strategy (Models)

An evaluation strategy is simply the way in which one goes about analyzing evaluation information. All evaluation strategies, or models, can utilize almost any type of available data. They distinguish themselves in terms of the questions they pose, how information is arrayed, and how decisions will be formulated. What follows is an overview of seven of the major evaluation models, their features, uses, and procedures. It should be noted that these evaluation models are only presented as sample techniques; they need not be used exactly as described, nor are they "required" of Project Directors. Many successful evaluations have been done merging one or two of the techniques, or using features from several of them. The important point here is to select a strategy which you feel comfortable with, and which will be most satisfactory to your audience.

Discrepancy Evaluation: The Model of Malcolm Provos

<u>DISCREPANCY EVALUATION - PROVOS</u>	
PURPOSE:	TO RELATE OUTCOMES TO OBJECTIVES
METHOD:	<ul style="list-style-type: none">● PERFORMANCE DATA VS. OBJECTIVES● BEHAVIORAL OBJECTIVES● ACHIEVEMENT TESTS
COMMENTS:	<ul style="list-style-type: none">● DISCREPANCIES ARE NOT NECESSARILY NEGATIVE THINGS● REQUIRES EXTENSIVE GOAL CLARIFICATION● NOT MERELY A CHECKLIST
MAJOR QUESTIONS:	<ul style="list-style-type: none">● ARE OBJECTIVES BEING ACHIEVED?● ARE PEOPLE BEING PRODUCTIVE?

To use this model, you compare the differences between standards you've set for performance, and the actual levels of performance: the outcomes of your project are compared with your objectives within the context of program operation.

In order to do this, you have to be able to clarify your goals, not just in terms of end products, but the goals you set for the

design, operation, and production of interim and final products, and for costs.

At all of these stages, indicators of performance are compared with standards which serve as criteria of performance, and discrepancies between performance and standards are reported to program managers. A discrepancy is defined as a simple difference and need not be negative - i.e., it is a discrepancy to be operating under budget yet certainly not negative.

When you administer a program, you keep in mind your goals, or plans, and your actual attainments, or results, but when you use the discrepancy model for evaluating, you concentrate on the difference between goals and results which is the "process" of your program.

What's found by studying the differences between goals and results is a set of "learnings" which can influence the results of your program, if you use them.

Oftentimes, the Discrepancy Model is used like a checklist or report card, that is: this is what you said you were going to do, now let's see whether or not you did it. But the model's intent is far more than this. It can aid you in doing what you said you were going to do and further to help you determine whether or not this is still a desirable goal.

Provus advocated continuous communication between project staff and evaluation staff, so that discrepancy information can be used at the earliest possible occasion to modify the program whenever managers feel modification is needed to attain the desired outcomes. In the Provus model, it's important to remember that for purposes of evaluation, discrepancies are not necessarily negative--differences need not be deficits, as they help you get a sense of what you're doing. According to this model, the place where you end up is affected not only by where you are, but also by what's happened along the way, if you keep an eye on the differences between where you are and where you thought you'd be.

The key features of the Discrepancy Model are:

- the need for a close link between program planning and evaluation
- evaluation is a continuous process
- evaluation criteria need to be linked to stages of program development and implementation
- use of discrepancy information as the primary method of program evaluation
- the need to maintain rapport between evaluators and program managers.

Goal-Free or Consumer-Oriented Evaluation: the model of Michael Scriven

GOAL FREE EVALUATION - SCRIVEN	
PURPOSE:	TO EVALUATE THE WORTH OF A PROGRAM THROUGH THE SUITABILITY OF ITS GOALS
METHOD:	<ul style="list-style-type: none">• CHECKLIST• OBSERVATION• COMPARATIVE/NON-COMPARATIVE TECHNIQUES• ESTABLISH STANDARDS
COMMENTS:	<ul style="list-style-type: none">• VALUES ARE EMBODIED IN STANDARDS• DEFINED ROLE OF EVALUATOR• ADVOCATE PROGRAM PROFESSIONAL AS EVALUATOR
MAJOR QUESTIONS:	<ul style="list-style-type: none">• WHAT ARE ALL OF THE EFFECTS?• WHAT IS THE CONSUMER UTILITY (BEST BUY)?

Goal-Free Evaluation, though the title of this model conjures up a variety of images, does not imply that the evaluation, project or evaluator is without goals. Rather this model is concerned with measuring all of the effects of a project and Scriven contends that this can only be done if the project is viewed holistically apart from its original goals or intents.

The model is based on the assumption that consumers don't usually care about whether a program achieves its organizational goals--they care about what a program really does, whether or not it's a "good" program, and they expect an evaluation to deliver this information to them. Scriven feels that these consumer "values" are embodied in standards which support the design, implementation and evaluation of a program. If evaluators are in fact saddled with this responsibility, they need to put a lot of emphasis into understanding the standards they use for evaluation: are they subjective, biased, appropriate? Scriven feels that the program professional (e.g., the arts educator) is best suited to conduct the evaluation as these persons are most familiar with the standards to be used.

One of the features of Scriven's model is that it can use comparison, either between programs, or in the same program "before and after," or in the field at large. A non-comparative technique also can be used. What's important is to determine whether or not the audiences of the evaluation will be making comparisons and to gather information accordingly.

In developing his evaluation model, Scriven distinguishes between "formative evaluation" (used to improve a program while its being implemented) and "summative evaluation" (used to judge the completed program, process, or product). Formative evaluation should be done by someone "inhouse," while summative evaluation is done by an independent, outside evaluator.

Scriven also distinguishes between "intrinsic evaluation" which examines the "means", and "pay-off" evaluation which examines the "ends."

Scriven developed a taxonomy of criteria for education evaluation studies which can be used as a checklist in designing an evaluation. The steps which Scriven developed for designing an evaluation methodology are:

- (1) Defining the type of evaluation needed by intended use--formative, summative or both
- (2) Defining the role of the evaluator(s)
- (3) Determining whether the goals of the program are adequate and worthwhile
- (4) Determining the instruments to be evaluated (intrinsic evaluation)
- (5) Determining the effects of the instrument(s) (pay-off evaluation)
- (6) Determining whether the evaluation is to be comparative or non-comparative.

Scriven's methodology for evaluation is more appropriately thought of as a set of conceptual distinctions (and thus related to evaluation design) than a procedural model for conducting an evaluation.

Illuminative Evaluation: The model of Parlett and Hamilton

ILLUMINATIVE EVALUATION - PARLETT & HAMILTON	
PURPOSE:	TO DESCRIBE AND INTERPRET RATHER THAN MEASURE
METHOD:	<ul style="list-style-type: none">• OBSERVATION• FURTHER INQUIRY• EXPLANATION
COMMENTS:	<ul style="list-style-type: none">• ASSUMES CONSENSUS ON ACTIVITIES• OUTCOME OF UNDERSTANDING DIVERSITY
MAJOR QUESTIONS:	<ul style="list-style-type: none">• WHAT IS THE PROGRAM'S RELATIONSHIP TO ITS ENVIRONMENT?• WHAT DISCOVERIES HAVE BEEN MADE?

The focus of this evaluation strategy is on describing and interpreting, rather than on measuring and predicting. Your goal as evaluator is to obtain the best picture of what's going on with the project--you use descriptive data obtained from observations, interviews, questionnaires, and archival records. After you've described the project, you share this information with your staff, clients, and others who are

interested and/or involved in the project, the purpose being to assure that there is consensus on what the activities of the program were.

The next step in this process is to interpret the program according to the interests and needs of the staff, clients and other interested groups. This step is made in order to gain some understanding of the project's diversity as perceived by external audiences. As evaluator, using this model, you want to know what is, and why: describe and interpret! The entire program should be studied in depth--its rationale, evolution, operations, achievements, and problems--in other words, the program in relationship to the environment in which it operates.

This approach is intended to be unstructured. The assumption behind it is that pre-defined hypotheses, which you want to test, end up constraining your data gathering, and unduly influence the nature of the findings.

The developers of this model specified three stages for the evaluation:

- observation
- further inquiry
- explanation

Each of these stages is conducted independently--the second and third stage building on the first stage. Observation is directed towards program activities. The second stage or further inquiry is used to provide information which will "illuminate" the program activities in terms of goals and outcomes. The final stage is used to explain or analyze the program in terms of goals, activities and outcomes.

Responsive Evaluation: the model of Robert Stake

<u>RESPONSIVE EVALUATION - STAKE</u>	
PURPOSE:	TO DESCRIBE, UNDERSTAND AND ASSESS THE WORTH OF A PROGRAM IN ORDER TO DEVELOP CHANGE STRATEGIES
METHOD:	<ul style="list-style-type: none"> ● ANTECEDENTS ● TRANSACTIONS ● OUTCOMES ● CONGRUENCE
COMMENTS:	<ul style="list-style-type: none"> ● UNDERSTANDING DIVERSITY ● NEGOTIATIONS ● NECESSITY OF BOTH DESCRIPTION AND JUDGEMENT
MAJOR QUESTIONS:	<ul style="list-style-type: none"> ● WHAT DECISIONS NEED TO BE MADE ABOUT THE PROGRAM? ● HOW CAN THESE DECISIONS BE INFORMED?

This model focuses on both description and judgment as being necessary parts of evaluation. In making a judgment, you are assessing the worth of a particular aspect of your program, e.g., an educational strategy, as it relates to improving an aspect of learning. What you're trying to do is make a judgment about your program, and share it with an external audience.

An important aspect of Stake's Model is the distinction of data sources by their relationship in time to the practice being evaluated:

- o Antecedents--those conditions existing prior to the teaching and learning experiences which may relate to outcomes, such as the students' aptitudes or previous experiences, previous curriculum, etc.
- o Transactions--interactions of persons in the learning environment, your program.
- o Outcomes--the results of antecedents and transactions.

In this model you look at things from two perspectives--the antecedents, transactions and outcomes which you intended, and those which you observe. In evaluation, you are looking for congruence between what you intended and what you observe. This also combines the use of observations to obtain descriptive information. But it is more structured than other models in that the program intents, or goals, serve as the framework for gathering information. And the emphasis in this model on the relationships (or contingencies, as Stake refers to them) between the antecedents, transactions, and outcomes makes it more environmentally-responsive than the Provus Discrepancy Model.

In essence, using this model you are making judgments about your program all along the way--are things good or bad--and you're continuously checking out your assumptions.

In this model, the descriptive data that you obtain are based on logic, while the judgments are based on empirical comparisons between results (outcomes) and standards. The model looks at logical consistencies in order to understand diversity between program intents and program operations.

The basic procedure of this model is to make a determination as to what decisions need to be made and then to gather information to inform these decisions and/or recommend change strategies.

CIPP (Context, Input, Process, Product) Evaluation, or
Decision-Oriented Evaluation: the model of Daniel Stufflebeam

CIPP (CONTEXT, INPUT, PROCESS, PRODUCT) - STUFFLEBEAM	
PURPOSE:	TO REMEDY PROBLEMS THROUGH EVALUATION (CHANGE)
METHOD:	<ul style="list-style-type: none">• LOOK AT THE DECISIONS TO BE MADE• CRITERIA VS. EFFECTIVENESS
COMMENTS:	<ul style="list-style-type: none">• QUALITY CONTROL• MANAGERIAL• SUGGEST CHANGE
MAJOR QUESTION:	HOW SHOULD A GIVEN PROGRAM BE PLANNED, EXECUTED AND DISSEMINATED TO FOSTER HUMAN GROWTH?

The CIPP model, a combination of other evaluation models, was designed to remedy problems in previous models. It assumes that there will be a large mutual influence between evaluation and decision-making.

The three major steps in the evaluation process are delineating, obtaining, and providing information for judging decision alternatives:

- delineating information refers to identifying required information
- obtaining information includes collecting and analyzing the data
- providing information refers to the organization of information for reporting purposes.

Seen from the perspective of this model, there are three environments, or contexts, in which changes can occur, within which decisions by managers must be made:

- those guided by technical standards, with a routine, cyclical data collection system
- those in which there is ongoing developmental activity aimed at continuous improvement
- those in which innovative activities are being used for inventing, testing, and diffusing new solutions to problems.

In this model four types of decisions are delineated, and determine the type of evaluation to be conducted:

- planning decisions: these are aimed at major program changes, based on discrepancy information. The evaluator performs a "context evaluation" in which the relevant environment is defined; intended and actual outcomes and discrepancies between them are described; problems that prevent a high degree of congruence are diagnosed.
- structuring decisions: these specify the means to ends that were established through planning decisions. The kinds of variables studied include program organization, personnel schedules, etc. "Input evaluation" is used to provide information on methods of resource utilization for meeting program goals.

- o Implementing decisions: these involve carrying out the planned actions. "Process evaluation" is directed towards these decisions needs in several ways--detecting or predicting problems in the procedural design or its implementation; providing information for programming decisions; and maintaining a record of implementation processes as they occur.
- o Recycling decisions: those which determine the course of any activity at any given point in the program. "Product evaluation" serves these decision needs by measuring and interpreting attainments, both during and after a project.

Art-Criticism Evaluation: the model of Elliott Eisner

ART CRITICISM MODEL - EISNER	
PURPOSE:	TO JUDGE A PROGRAM'S WORTH IN RELATIONSHIP TO ITS FIELD
METHOD:	<ul style="list-style-type: none"> o BLUE-RIBBON PANEL o RECOGNIZED STANDARDS
COMMENTS:	<ul style="list-style-type: none"> o SEEKS TO IMPROVE STANDARDS o AFFECTIVE INFORMATION o STAMP OF APPROVAL
MAJOR QUESTIONS:	<ul style="list-style-type: none"> o WOULD A CRITIC APPROVE THIS PROGRAM? o WHAT DISTINGUISHES A PROGRAM FROM OTHERS OF THE SAME TYPE?

This evaluation model is patterned after the process of art criticism. It's also called "connoisseur-based evaluation."

Programs are evaluated either by expert, so-called "blue-ribbon," panels; or they are compared to recognized standards of excellence. These standards of excellence need not be only those from the field at large, but more importantly, these standards should be recognized by the audiences of the

specific project. Standards can be as measurement-oriented as increases in test scores or qualitatively-oriented such as artistic excellence. A blue-ribbon panel can be composed of recognized experts in the field or professionals in the field can assume the role of the panel by using agreed-upon standards for judgment, e.g., a group of music teachers within a school arts program can be brought together and establish criteria for judging their own program.

The goal is to critically describe, appraise, or illuminate a program element, and to determine what merits or demerits distinguish a program from others that are at the same level of development.

Basically, the goal is to place a "stamp of approval" on a program by determining "would a critic approve this program?"

As a model, it is particularly dependent on the evaluators' expertise and sensitivities, and on their ability to convey meaning and feelings. This model assumes that judgment is based on affective information or "feelings" and so evaluations should address the affective domain.

Another feature of this model is that by utilizing critical standards or criteria as a basis for judgment, you not only determine the project's status in relationship to these standards, but you can seek to improve or "raise" the standards of the field.

Jurisprudence or Adversary Evaluation: the model of Robert Wolf

JURISPRUDENCE MODEL - WOLFE	
PURPOSE:	TO ESTABLISH AN ARGUMENT AS TO A PROGRAM'S WORTH
METHOD:	<ul style="list-style-type: none">• GATHERING EVIDENCE• UTILIZING ADVERSARY• ESTABLISHING PROCEDURES• JUDGEMENT AND RESOLUTION
COMMENTS:	<ul style="list-style-type: none">• SEEKS TO MAXIMIZE ACHIEVEMENT OF DESIRED RESULTS• UTILIZE OPPONENTS SO AS TO CONVINCE
MAJOR QUESTION:	WHAT ARE THE ARGUMENTS FOR AND AGAINST THIS PROGRAM?

The jurisprudence/adversarial evaluation model is characterized by: information management by courtroom procedures; use of citizens for information review; and bipolar exploration of key issues. Two evaluation teams are organized. One team serves as the advocate team and seeks to gather information and develop logical arguments in support of the program being evaluated. A second team serves as the adversary team and attempts to present information highlighting

the inadequacies and weak-points of the program. Neither of these evaluation teams is responsible for the ultimate evaluation of the program; rather, their respective responsibilities are to collect relevant information and to develop evidence and arguments in support of their position for or against the program.

A jury of "citizen evaluators" hears the evidence and arguments of each evaluation team. The evidence is presented in court room style with various witnesses giving testimony (a review of test scores, explanation of decisions, descriptions of the effects of certain decisions, etc.).

Each team is allowed the opportunity to "cross-examine" witnesses. This permits the development and exploration of alternative hypotheses. The jury of citizen evaluators may direct questions (funneled through the jury foreman) to any witness for the purposes of clarification or to examine other hypotheses.

After both evaluation teams have completed their presentation of evidence and arguments and after each team has presented summary statements, the jury of citizen evaluators retires to deliberate on the evidence that has been presented. It is then the responsibility of this citizen jury to make judgments concerning the program being evaluated point-by-point for each of the prespecified key issues. The jury may report favorably or unfavorably, in part or whole, and may also identify areas where the evidence is too contradictory, too minimal, or too one-sided and thus withhold judgment on certain issues.

The jurisprudence model has three advantages not offered by any other evaluation strategy: 1) it insures the full exploration of key issues through the presentation of evidence and arguments reflecting both sides of the issue; 2) it enhances the opportunity to develop alternative inferences from data; and 3) it removes from the evaluator, the responsibility for summary judgments and places this responsibility with a panel of citizen jurists.

The most obvious disadvantages are increased cost (by funding two evaluation teams rather than one) and the administrative complexity of agreeing on evidentiary rules, specifying the scope of key issues to be addressed, and selecting instrumentation common to both evaluation teams. Some of these "disadvantages" can be countered by utilizing advisory panel members as the evaluation team.

V. SUMMARY

The documentation, evaluation and dissemination strategies in this guidebook have been presented within a larger context which views the management of information as a process as important to the success of a program as that of managing resources and personnel.

Moreover, these ideas have been based upon our particular view of the needs and uses for information in Federal demonstration programs such as the USOE Arts Education Program. Within this perspective national, state, local, and individual information needs are considered interdependent. Thus, effective information management starts with the experiences of the local program staff and clients and, to the extent it provides useful information at each level of program support, it eventually produces the quality of information needed for national management and dissemination needs.

Finally, we view information as more than experience-derived data. It also can be used as part of experiences, and especially to influence others. An Arts Education project does have a complex task: as a "temporary system" it must have a positive influence on individuals and permanent social institutions over which it has little or no direct control. Each of these has its own goals, objectives and primary responsibilities. It is for this reason that documentation, evaluation and dissemination can be major management strategies. Many of the purposes of an Arts Education program are determined and shared by the individuals and institutions with which it works. To the extent that these individuals and institutions are provided with information to perceive the achievement of their purposes through the project, to that extent the Arts Education project may successfully influence changes in the permanent institutions with which it interacts. In other words, when individuals perceive their purposes and growth enhanced by an Arts Education project's activities, the probability of achieving project objectives increases and persons both within the project and representing cooperating institutions feel ownership in the project and its strategies.

Effective management of information in Arts Education project management, to us, therefore, provides:

- a unifying process which links the needs of the individuals and institutions;
- a learning process which enables a project to stay on top of the complex and changing conditions it faces by accurately perceiving its environment, drawing information from situations while still in process, learning from these situations, and applying the learning to its advantage;
- an involving process which recognizes that while management decisions may be made by one person, no single individual has the experience or information to support the decision. When all project members share information, share in decision-making, and understand clearly the rules, regulations, and operating limits, the project runs well and persons gain psychic rewards from participation;
- a controllable process which acknowledges that, while the functioning of many of the project's components are only indirectly controlled by the project, management of information is one major function which is totally under the project's control. Too often a project will relinquish what is actually its "power" to affect its own outcomes because it feels that documentation, evaluation or dissemination are functions that have to be done for Washington, or that have to be separate components distinct from other objectives. In these cases, management of information often becomes a purpose in itself for which only a few people are responsible, and in whose successful accomplishment there is little investment by others.
- and finally, information management is a natural process--an ongoing problem-solving process that responds to the basic human need to have some influence on surrounding conditions. The individuals to be involved in the project--administrators, parents, support personnel--already have their personal "management" system in place permitting them to identify needs, to select appropriate actions and to self-correct before acting again. The major management task for the Arts Education project director is one of providing these individuals and groups with the information and opportunities to apply their natural problem-solving abilities and interest to the accomplishment of project objectives.

"...communication requires shared experience..."

Peter F. Drucker
Management

Two final thoughts . . .

about information that goes in, and comes out,
of our programs . . .

The Government are very keen on amassing statistics. They collect them, add them, raise them to the n th power, take the cube root and prepare wonderful diagrams. But you must never forget that every one of these figures comes in the first instance from the village watchman, who just puts down what he damn pleases.

Sir Isiah Stamp

INLAND REVENUE DEPARTMENT (England) 1896-1919

THE GOVERNMENT IS VERY KEEN ON PRODUCING EXEMPLARY PRODUCTS AND PROGRAM MODELS. THEY RESEARCH THE NEED FOR THEM, DEVELOP, IMPLEMENT AND EVALUATE THEM; DISSEMINATE AND DIFFUSE THEM, AND PRODUCE WONDERFUL REPORTS ABOUT THEM. BUT YOU MUST NEVER FORGET THAT EVERY ONE OF THESE PROGRAMS OR PRODUCTS WILL, IN THE FINAL INSTANCE, BE USED BY A HUMAN BEING WHO WILL DO WHATEVER HE DAMN PLEASURES TO MEET HIS NEEDS FOR SATISFACTION IN HIS WORK.

LEWIS A. RHODES
(USA) 1976

APPENDIX A

BACKGROUND MATERIAL

WHERE WE STARTED ...

This current effort of the USOE Arts and Humanities staff to provide assistance in evaluation, documentation and dissemination is limited to a 2-day workshop. However, it is built upon an understanding that has been developing over the past few years concerning the unique needs of the art educator who often functions as a program manager in a local or state educational system.

There have been several state and national meetings and conferences to explore the particular evaluation concerns of the art educator. We have attempted to utilize the findings of these meetings as indicators of the directions this workshop should go. For example, in March 1977 the Department of Public Instruction of North Carolina held a symposium on "Measurement in the Arts: The Search for X." This symposium was held to give participants a view of the many efforts being made to assess arts programs and their effects on students. There were several conclusions that were felt to be of major importance.

1. We need to give much more thought to the reasons we collect data.
2. We need to give much more thought to the use we plan to make of data.
3. We need to make far greater use of the immense amount of data now at hand.
4. We need to look more carefully at the consequences of data collection as those consequences relate to our best defined desires for children...since the simple act of collecting data lets loose implications over which the collectors have no control."

In March 1978, the American Theatre Association and USOE conducted a seminar entitled "The Analysis and Dissemination of Arts Education Evaluation Techniques." The concept of the seminar centered on the fact that

... the demand for evaluation of arts projects has grown in the past ten years. Since arts evaluation is still in its infancy, few state and regional arts practitioners and

THE COMMUNICATION OF EXPERIENCE
A WORKSHOP
ON THE USES OF INFORMATION

USOE
Arts & Humanities
Staff



leaders know very much about its function and methods. What is appropriate, useful, manageable and cost-effective for arts-in-education projects is critically important to the growth and development of new programs. Therefore, there is a need for: 1) knowledge of methods available for evaluation of arts projects, and 2) dissemination of the information to project administrators and directors in the field.

One of the seminar participants, Dr. Jerome Hausman, presented several comments and recommendations at the conclusion of the seminar. He felt that

...within each of the arts fields, there are problems and ambiguities pertaining to evaluation. This is "built-in" by virtue of the dynamics of the arts themselves--changing conceptions for form, content, and context. All of this is further compounded by thinking that links the arts and environment. No wonder we are loathe to develop simplistic approaches to evaluating "state-wide" arts in education efforts. No wonder there is a turning toward more generalized data sources and more open-ended and flexible evaluation strategies. We have found that overall evaluation strategies need to be responsive to multiple and perhaps diverse criteria. But, we need to be clearer and more structured in the strategies we develop.

My own tendency toward self-study and anecdotal review procedures is based upon the assumption that critical issues must be reviewed in context. We also need to develop means to distill critical elements for analysis and assessment. Bob Stake's use of terms like portrayal and depiction are very useful and exciting to me. The real challenge rests in telling the story with attention to critical details and then assessing the meaning and significance of data. Of course, measurement approaches and techniques may be applicable along with assessment of performance or outcomes.

Hausman goes on to point out that

...evaluation strategies need to address themselves to processes and outcomes accountable to the arts! Once we have developed our general picture, specific criteria pertaining to behaviors, skills, and outcomes should be distilled from this general picture. Responsibilities for persuasion should only be assumed after the data are reconciled with values and purposes consistent with the program's objectives.

In addition to these meetings, the writings of both arts education and evaluation theorists reflect the concerns for evaluation. Robert Stake finds that

...accreditation reports rarely show the real worth of an educational program.

Michael Scriven says that

...if the goal of evaluation is to determine worth, then, to collect data, focus inquiry and inspect causal relationships without making a judgment of worth is not evaluation.

Stake elaborates on this idea by stating that controls, tests, regression equations and behavioral language may all be essential for test construction but may not facilitate the conduct of an evaluation project at all. These theorists have responded to people who are concerned with documentation, evaluation and dissemination in the field of arts education. Laura Chapman writes that

...the purpose of evaluation is to discover the value or significance of something. Measurement is not an integral part of evaluation; it imposes on the concept of "worth" an ordering and ranking system which leads to a denial of intrinsic value. An evaluation of an arts program should be empirical, valid, broadly reliable and responsive to the qualitative character of art experience.

Stake takes this theme even further when he says

...We need a reporting procedure for facilitating vicarious experience. We need to portray complexity, holistic impression, mood and even the mystery of experience.

Finally, we have also based this workshop on the comments which were expressed by current Arts and Humanities grant project directors:

- "Evaluation and research are essentially the same thing."
- "Documentation and dissemination are devices to show others what we did that was good."
- "Sophistication in evaluation means a strong preordinate design usually done by someone else."

- "If we do deal with the relationship between what is actually done and what is evaluated, might not Arts Education lose out."
- "Evaluation means numbers to show administrators."
- "Evaluation is never simple, if it is, it is not good enough. Our teacher-made tests are just good for us but they are not what a real evaluator would do."
- "Good evaluations are expensive and take a lot of money and time to do."

APPENDIX B

BIBLIOGRAPHIES

SELECTED TESTS IN THE ARTS

CREATIVITY

ART

DANCE

DRAMA

MUSIC

Compiled by:

Division of Cultural Arts
Department of Public Instruction
Raleigh, No. Carolina
March, 1977



SELECTED TESTS IN THE ARTS

Creativity

Torrance, E.P.
ASK AND GUESS TEST

Neutral forms, age, originality, fluency, flexibility, elaboration, production (written)
Personnel Press, Inc.

Item: Consist of three activities:

- I. Asking - subject required to list as many questions as he can about a drawing presented to him.
- II. Guessing Causes - subject required to list as many as possible causes of the action in the given drawing.
- III. Guessing Consequences - subject required to list as many possibilities as he can of what might happen as a result of what is taking place in the picture.

Scoring: Based on Fluency (no. of relevant responses), Flexibility (1 point given for utilization of each of several categories given in the scoring manual), and originality of responses.

Validity: The author presents considerable detailed information on content validity, construct validity, concurrent validity, scoring reliability, and test-retest reliability in the technical manual of the test series.

Eisner, E.W.
BUILDING TEST

Art, unspecified, originality, production (building, drawing)
Typology of creativity in the Visual Arts. Eisner, E.W.
Studies in Art Ed., 1962, 4(1), 11-22

Item: Subject must construct an object out of 1/4 pound of clay and colored tooth picks.

Scoring: Based on opinions of 3 trained judges. Evaluation proceeded on a nine point scale and objects scored on several criteria: Boundary Pushing (subject and form), Boundary Breaking (subject and form) aesthetic organizing.

Torrance, E.P.
CIRCLES TEST

Neutral forms, age, originality, fluency, flexibility, elaboration, production (building, drawing)
Personnel Press, Inc., Princeton, N.J.

Item: Subject presented with a series of 36 circles printed on 2 pages and is instructed to draw as many different objects as he can incorporate in the circles.

Item: Four Items-Forms: Subject lists different things each of 5 ambiguous silhouettes could possibly be.

Functions: Subject lists different uses for each of 4 common objects.

Concepts: Given an attribute of a class of objects, e.g. "things round", subject must list appropriate objects.

Scoring: Scoring for each test based on frequency (no. of responses) variety (no. of different classes of responses), and uniqueness (in frequency of occurrence).

Validity: The author presents detailed information regarding validity and reliability in the article.

Torrance, E.P.

PICTURE COMPLETION TEST

Neutral forms, age, originality, fluency, flexibility, elaboration, production (building, drawing)

Personnel Press, Inc. Princeton, N.J.

Item: Subject given a series of 10 drawings of incomplete figures and is instructed to complete a drawing using the lines given.

Scoring: Based on fluency (No. of pictures completed), flexibility (No. of different categories of responses), originality (Imagination), and elaboration (Amount of detail).

Validity: The author presents considerable detailed information on content validity, construct validity, concurrent validity, scoring, and test-retest reliability in the technical manual of the test series.

Torrance, E.P.

PICTURE CONSTRUCTION (CURVED SHAPE) TEST

Neutral forms, age, originality, elaboration, reproduction (building, drawing)

Personnel Press, Inc., Princeton, N.J.

Item: Subject given a piece of colored paper in the form of a curved shape and must draw a picture utilizing it.

Scoring: Based on originality (on a scale from zero to five), and elaboration (amount of detail).

Validity: The author presents considerable detailed information on content validity, construct validity, concurrent validity, scoring, and test-retest reliability in the technical manual of the series.

Likert, R. Quasha, W. H.
MINNESOTA PAPER FORM BOARD TEST

Form, age perceptual grouping, successive categories
Psychological Corporation, 304 E. 45th Street
New York, New York 10017

Item: 64 multiple choice items each consisting of a geometric figure divided into 2 or more parts. The subject must choose one of five figures presented that would represent the original should the parts be fitted together.

Scoring: The authors present considerable data on norms, validity, and reliability in the manual accompanying the test.

Child, Irwin L.
PREFERENCE AND AESTHETIC SENSITIVITY TEST

Painting, age, preference, rank order
Personal preferences as an expression of aesthetic sensitivity
by Irwin L. Child, Journal of Personality, 1962 vol. 30 pp 496-512

Test consists of 12 sets of pictures each set containing 60 pictures. Subject must rate reproductions of paintings according to his preferences by dividing each set of pictures into 10 piles of 6 pictures each and rating each group of 6 individually.

Scoring: Norms derived by the scaling of reproductions by 13 art graduate students according to aesthetic quality.

Reliability: Determined by test-retest method yielding a correlation of approx. .90 for 44 undergraduates.

Benner, G. Seashore, H. Wesman, A.G.
SPACE RELATIONS - DIFFERENTIAL APTITUDE TEST BATTERY

Form, age, perceptual grouping, successive categories
Psychological Corporation 304 E. 45th Street
New York, N.Y.
Grades 8 thru 12

Item: Test consists of 40 patterns which can be folded into figures. For each pattern, five figures are shown. The subjects must decide which of the figures can be made from the pattern.

Scoring: Extensive information on scoring and norms is presented in the manual of the test series.

Lewerenz, Alfred S.
TESTS IN FUNDAMENTAL ABILITIES OF VISUAL ARTS

California Test Bureau Delmonte Research Park
Monterey, Calif.

Eisner, Elliot W.

EISNER ART INFORMATION INVENTORY

Questionnaire, age, background knowledge, understanding, standard objective

Arts Curricula for the gifted.

Eisner, Elliot W. Teachers College Record, 1966, 67(7)
492-501

Consists of 60 multiple choice items equally divided into subtests dealing with art terms, art media and processes, artists and their works, and art history.

Scoring: No norms given.

Reliability: Application of the Kuder-Richardson Formula 20 to the tests results of 1488 subjects yielded a coefficient of .934.

Graves, Maitland

GRAVES DESIGN JUDGEMENT TEST

Figure, age, sculpture, preference, pair comparisons
Psychological Corporation, 304 East 45th Street
New York 17, N.Y.

Item: Ninety items each consisting of 2 or 3 plates from which the subject must select the one he considers best.

Scoring: Scoring key furnished. Norms given in percentiles for high school and college art and non-art students. Norms derived from test results of 826 art students and 480 non-art students.

Reliability: Coefficients range from .81 to .93 with a median of .86 based on retests of 14 groups of students from 2 high schools and 3 post high school institutions.

Horn, Charles, C.

HORN ART APTITUDE INVENTORY

Task, age, drawing, production
C.H. Stoelting Co.
424 North Homan Avenue
Chicago 24, Ill.

Item: Part One (scribbling and doodling): Subject must sketch twenty familiar objects.

Part Two (imagery), subject must manipulate twelve rectangles containing "key" lines to arrive at a picture.

Scoring: Items rated as either excellent, average, or poor based on examples of the 3 levels of work illustrated in the manual. No method given for obtaining quantitative scores.

Validity: Determined by correlations between two scores and teacher ratings of 36 high school seniors, the product-moment correlation coefficient equaled .659.

SELECTED TESTS IN THE ARTS

Dance

McCulloch, Margaret L. TEST OF RHYTHMIC RESPONSE

Aural (music), age, rhythmic performance, motor performance.
The development of a test of rhythmic response through
movement of first grade children. McCulloch, Margaret L.
Thesis (Ed.D.) Univ. of Oregon.

Item: Test consists of 14 items requiring the subject to
step, clap, walk, and bend in time to given rhythmic patterns.
Test purports to measure response to pulse beat, response to
accent, response to rhythmic patterns, and response to musical
phrasing.

Scoring: Based on a rating scale from 1 to 3 indicating
whether the subject was always, sometimes, or never in time
with the music.

Validity: Test items selected by a dance and a music "specialist".

Reliability: Test-retest reliability coefficient reported to
be .90 based on test results of 27 first grade children.

Coppock, D.E. TEST OF RHYTHMIC MOTOR RESPONSE

Aural (music), rhythmic performance, motor performance.
Development of an objective measure of rhythmic motor response.
Coppock, D.E. Doctoral Dissertation, Univ. of Iowa, 1964.

Item: Subject is presented with 23 rhythmic patterns (drum)
of increasing complexity and is required to step in time to
each pattern.

Scoring: Based on a 5 point rating scale (from 0 to 4).
Subjects rated on meter, tempo, and total for each item.

Validity: Determined by correlations between test results and
teacher ratings of 92 subjects. Coefficient found to be .267
(p.01).

Reliability: Based on test results of 92 subjects. Spearman-
Brown Prophecy Formula yielded a coefficient of .91.

Withers, Maida Rust CREATIVITY OF MODERN DANCERS

Action, dance training, dance performance, motor performance.
"Measuring the Creativity of Modern Dancers," by Maida Rust
Withers. Masters Thesis University of Utah, 1960

SELECTED TESTS IN THE ARTS

Music

Aliferis, James Stecklein, John E.
ALIFERIS-STECKLEIN MUSIC ACHIEVEMENT TESTS

Rhythm, melody, harmony, age, aural identification, visual note reading, standard objective
University of Minnesota Press Minneapolis 14, Minn.

Item: Consists of 79 items at a piano music (played) to be compared with printed notations of melody, rhythm, or harmony. Multiple choice of 4 alternatives per played item.

Scoring: Norms given are based on the testing of 2500 music students. Separate norms for geographical location, type of instrument studied, and type of school.

Beach, Frank A. Schrammel, H.E.
BEACH MUSIC TEST

Music, pitch, musical training, age, aural identification, composer identification, visual note reading, pair comparisons, standard objective
Bureau of Educational Measurements Kansas State Teachers College Emporia, Kansas

Item: 70 items divided into eleven sub-tests consisting of knowledge of musical symbols, recognition of measure, tone direction and similarity, pitch discrimination, application of syllables, time value, musical terms and symbols, correction of notation, syllables and pitch names, representation of pitches, composers and artists.

Scoring: Norms given in percentiles for grades 7 through college.

Validity: Based on test scores and teacher rankings of several thousand students. Correlations with musicianship = .65, with knowledge of musical fundamentals = .74.

Drake, Raleigh M.
DRAKE MUSICAL APTITUDE TESTS

Rhythm, melody, verbalization, age, percussion playing, visual note reading, pair comparisons,
Science Research Associates, Inc.
259 East Erie Street, Chicago 11, Ill.

Item: 2 Parts, musical memory, rhythm. Musical Memory: Subject listens to a number of pairs of 2 bar melodies played on the piano and compares them as to sameness, or change in notation, key, or tempo.

Music

McAleavey, Grace Anne

GROUP TEST OF MUSICAL ACHIEVEMENT FOR THE THIRD GRADE

Pitch, rhythm, age, aural identification, visual note reading, standard objective

"Formulation of a Group Test of Musical Achievement for the Third Grade" by Grace A. McAleavey, Masters Thesis Northwestern Univ., Evanston, Ill.

Item: Group test of 4 parts. Tester reads questions aloud and S's raise hand to answer. I. Knowledge of musical terms - 28 multiple choice items, II. Knowledge of syllable names - 20 fill-in blank items, III. Tonal recognition - 6 forced choice items, IV. Rhythmic recognition 0 16 fill-in blank items.

Scoring: Tentative norms established by percentile rankings of 300 children.

Validity: Determined by judgements of experts, and analyzing 40 courses of music study.

Reliability: Determined by correlating two forms of test yielding correlation of .76 for 300 pupils.

Drake, Raleigh M.

INTERVAL DISCRIMINATION TEST

Timbre, age, aural discrimination, pair comparisons
Drake, R.M., "Four New Tests of Musical Talent", J. Applied Psych. XVII, 1933, pp. 136-147

Item: Test requires comparison of two musical intervals played on the piano to determine if the last interval is longer or shorter than the first. 80 pairs of intervals to be compared.

Validity: Validity coefficient measured by teachers' estimates .77 with N=46.

Reliability: Self-reliability coefficient - .83 with N=46.

Taylor, Corwin H.

MEASURES OF MUSICAL BACKGROUND FORM A

Rhythm, melody, harmony, age, aural identification, background knowledge, understanding, successive categories, production
"The Construction and Validation of certain Experimental Measures of Musical Potentiality" by Corwin H. Taylor.
Doctoral Thesis Univ. of Cincinnati, 1941

Item: Consists of 4 tests: Test I. Recognition of Tonality - a) 32 melodies are presented and must judge whether each is minor or major key. b) 52 chords are presented and S must choose major or minor. Test II. Rhythmic Discrimination - a) 32 selections are presented and S must indicate meter of each.

Music

Reliability: Kuder-Richardson formula 20 yielded a coefficient of .78 based on the test results of 2314 subjects from various size schools in Colorado.

Hooper, Sharon
MUSIC READINESS TEST

Questionnaire, age, background knowledge, understanding, visual note reading, standard objective

Item: Test consists of 112 items divided into five sub-tests dealing with interpretation of terminology used in notation, ability to determine titles of familiar songs from a few measures of notation, concepts and skills concerning rhythm.

Scoring: Norms given in means and standard deviations based on test results of 200 students in grades 3 and 4.

Reliability: Coefficients derived from test results of 200 subjects in grades 3 and 4. Split-half = .99, Kuder-Richardson Formula 20 = .937.

Seashore, Carl E. Lewis, D. Saeveit, J.G.
SEASHORE MEASURES OF MUSICAL TALENTS, REVISED EDITION

Pitch, rhythm, volume, melody, musical training, age, aural discrimination, pair comparisons
Psychological Corporation, 304 East 45th Street, New York 17, N.Y.

Item: Test consists of six sub-tests requiring the subjects discrimination of pitch, intensity, time, consonance, tonal memory, and rhythm.

Scoring: Scoring key furnished. Norms given in percentile.

Validity: The authors present considerable information regarding validity and reliability in the manual accompanying the test.

Watkins, John G. Farnum, Stephen E.
WATKINS-FARNUM PERFORMANCE SCALE: A STANDARDIZED ACHIEVEMENT TEST FOR ALL BAND INSTRUMENTS

Music score, musical training, brass playing, string playing, woodwind playing, percussion playing, instrument playing
Hal Leonard Music, Inc. 64 E. 2nd St., Winona, Mich.

Item: Consists of a series of musical exercises of increasing difficulty presented to the student for sight reading and individual playing.

Scoring: Any and each error in pitch, tempo, rhythm, etc. incurred while the subject plays each piece counts as minus 1 towards the total score, Norms given based on the number of years the instrument has been studied.

A BIBLIOGRAPHY FROM A. N. OPPENHEIM'S
QUESTIONNAIRE DESIGN AND ATTITUDE MEASUREMENT

Problems of Survey Design

Sampling Methods

1. A. Stuart, Basic Ideas of Scientific Sampling (London: Griffin, 1962).
An excellent, nonmathematical account.
2. Isidor Chein, "An Introduction to Sampling," Appendix B in Claire Selciz, Marie Jahoda, Morton Deutsch, and Stuart W. Cook, eds., Research Methods in Social Relations (New York: Holt, 1959).
3. F. Yates, Sampling Methods for Censuses and Surveys (London: Griffin, 1949).
4. Morris H. Hansen, William N. Hurwitz, and William G. Mordon, Sample Survey Methods and Theory (New York: Wiley, 1953), Vol. I.

Practical sampling methods applied to surveys.

Experimental Design

5. Ronald A. Fisher, The Design of Experiments (Edinburgh: Oliver and Boyd, 1947).
A classical textbook on experimental design.
6. W. G. Cochran and G. M. Cox, Experimental Designs (New York: Wiley, 1950).
7. A. E. Maxwell, Experimental Design in Psychology and the Medical Sciences (London: Methuen, 1958).
8. Allen L. Edwards, "Experiments: Their Planning and Execution," in Gardner Lindzey, ed., Handbook of Social Psychology (Cambridge, Mass.: Addison-Wesley, 1954).
9. D. R. Cox, Planning of Experiments (New York: Wiley, 1961).

Survey Design

10. Claus A. Moser, Survey Methods in Social Investigation (London: Heinemann, 1958).

A good general text on survey methods. See especially chapters on sampling and on interviewing.

11. Herbert H. Hyman, Survey Design and Analysis: Principles, Cases, and Procedures (Glencoe, Ill.: Free Press, 1955).

For the advanced student.

12. Robert L. Kahn and Charles F. Cannell, The Dynamics of Interviewing (New York: Wiley, 1957).

A useful textbook on surveys, questionnaire-design, and interviewing.

Matching Problems

13. William A. Belson, "Matching and Prediction on the Principle of Biological Classification," Applied Statistics, VIII (1959), 65-75.

On the uses of the stable-correlates method of matching.

14. W. Z. Billewicz, "Matched Samples in Medical Investigations," British Journal of Preventive and Social Medicine, XVIII (1964), 167-173.

Interviewer Errors

15. Herbert H. Hyman et al., Interviewing in Social Research (Chicago: University of Chicago Press, 1954).

Experimental studies of interviewer bias and the effects of interviewers on respondents.

16. Robert L. Kahn and Charles F. Cannell, The Dynamics of Interviewing (New York: Wiley, 1957).

See especially the section dealing with interviewer bias.

Epidemiology and Social Medicine

17. J. N. Morris, Uses of Epidemiology (Edinburgh: Livingstone, 1957).

A short basic textbook.

Panel Studies

18. Morris Rosenberg, Wagner Thieleus, and Paul Lazarsfeld, "The Panel Study," in Marie Jahoda, Morton Deutsch, and Stuart W. Cook, eds., Research Methods in Social Relations (New York: Dryden Press, 1951), Part II.

19. James W. B. Douglas and J. M. Blomfield, "The Reliability of Longitudinal Surveys," The Millbank Memorial Fund Quarterly, XXXIV (1956), 227-252.

Problems of Questionnaire Design

Interviewing

1. Robert L. Kahn and Charles F. Cannell, The Dynamics of Interviewing (New York: Wiley, 1957).

Chapter 6 is a useful introduction to questionnaire design; chapters 7, 8 and 9 deal fully with interviewing skills, sources of bias, and so on.

2. Herbert H. Hyman et al., Interviewing in Social Research (Chicago: University of Chicago Press, 1954).

Experimental studies of interviewer bias and the effects of interviewers on respondents.

3. Robert K. Merton, Marjorie Fiske, and Patricia L. Kendall, The Focused Interview (Glencoe, Ill.: The Free Press, 1956).

A careful and clear analysis with many examples of interviewing problems and procedures.

4. Paul B. Sheatsley, "The Art of Interviewing and a Guide to Interviewer Selection and Training," in Marie Jahoda, Morton Deutsch, and Stuart Cook, eds., Research Methods in Social Relations (New York: Dryden Press, 1951), Part II.

An introduction to the conduct of interviews and the problems of interviewer training.

5. Eleanor E. Maccoby and Nathan Maccoby, "The Interview: A Tool of Social Science," in Gardner Lindzey, ed., Handbook of Social Psychology (Cambridge, Mass.: Addison-Wesley, 1954).

Basic textbook reference on problems of research in interviewing.

Mail Questionnaires

6. Christopher Scott, "Research on Mail Surveys," Journal of Royal Statistical Society, XXIV, Series A (1961), 143-195.

An excellent contribution to research into mail questionnaires, together with a full review of the literature.

Question Sequence

7. George Gallup, Qualitative Measurement of Public Opinion: The Quintamensional Plan of Question Design (Princeton, N. J.: American Institute of Public Opinion, 1947).

7. William A. Belson, Studies in Readership (London: Business Publications Ltd., 1962).

A carefully conducted follow-up inquiry into the causes of unreliability in a national readership survey.

Occupational Grading

8. Hilda T. Himmelweit, A. N. Oppenheim, and P. Vince, Television and the Child (London: Oxford University Press, 1958).

See pp. 82-84 for problems of obtaining father's job information from school children.

9. Fortune Surveys, The People of the United States--a Self-Portrait (Fortune, February, 1950, Vol. 21).
10. Richard Centers, The Psychology of Social Classes (Princeton: Princeton University Press, 1949).

Embarrassing Questions

11. Alfred C. Kinsey, W. B. Pomeroy, and C. E. Martin, Sexual Behavior in the Human Male (Philadelphia and London: W. B. Saunders, 1948).

See especially pp. 35-82 on interviewing techniques.

12. Allen J. Barton, "Asking the Embarrassing Question," Public Opinion Quarterly, XXII (1958), 67-68.

A humorous set of suggestions for wording embarrassing questions.

Reliability and Validity

13. Lee J. Cronbach, Essentials of Psychological Testing (New York: Harper, 1960).

A useful textbook on tests and their uses. See especially chapter 5 on validity, and pp. 126-142 on reliability.

14. T. W. Adorno, Else Frenkel-Brunswik, D. J. Levinson, and R. N. Sanford, The Authoritarian Personality (New York: Harper, 1950).

Checklists, Rating Scales, and Inventories

Psychological Measurement

1. Harold Gulliksen, Theory of Mental Tests (New York: Wiley, 1950).

A basic textbook on test-construction theory.

- 7 -
13. Anne Anastasi, Psychological Testing (New York: Macmillan, 1963).

Attitude Statements

Attitude Theory

1. David Krech, Richard S. Crutchfield, and Egerton L. Ballachey, Individual in Society (New York: McGraw-Hill, 1962).

See Chapter 5, "The Nature and Measurement of Attitudes," for the characteristics of attitudes and attitude systems, an exposition of scaling methods and types of survey questions.

2. D. Katz and E. Stotland, "A Preliminary Statement to a Theory of Attitude Structure and Change," in S. Koch, ed., Psychology: A Study of a Science, Vol. 3: Formulations of the Person and the Social Context (New York: McGraw-Hill, 1959).

A wide-ranging essay on attitude theory.

3. Hans J. Eysenck, The Psychology of Politics (London: Routledge & Kegan Paul, 1954).

See especially Chapter 4 on the organization of social attitudes.

4. Edward A. Suchman, "The Intensity Component in Attitude and Opinion Research," in Samuel A. Stouffer, ed., Measurement and Prediction (Princeton, N. J.: Princeton University Press, 1950).
5. T. W. Adorno, Else Frankel-Brunswik, D. J. Levinson, and R. N. Sanford, The Authoritarian Personality (New York: Harper, 1950).

A classic study of prejudice, showing the relationships between different kinds of prejudice and the development of the concepts of ethnocentrism and authoritarianism.

6. E. Wellin, "Water Boiling in a Peruvian Town," in B. D. Paul, ed., Health, Culture and Community (New York: Russell Sage Foundation, 1955).
7. Bernard M. Kramer, "Dimensions of Prejudice," Journal of Psychology, XXVII (1949), 389-451.

Writing Attitude Statements

8. Allen L. Edwards, Techniques of Attitude Scale Construction (New York: Appleton-Century-Crofts, 1957).

See in particular p. 13 showing how to edit attitude statements.

4. S. S. Stevens, "Mathematics, Measurement, and Psychophysics," in S. S. Stevens, ed., Handbook of Experimental Psychology (New York: Wiley, 1951).

On general measurement theory.

5. A. L. Edwards, Techniques of Attitude-Scale Construction (New York: Appleton-Century-Crofts, 1957).
6. Matilde W. Riley, J. W. Riley Jr., and J. Toby, Sociological Studies in Scale Analysis (New Brunswick, N.J.: Rutgers University Press, 1954).
7. W. S. Torgerson, Theory and Methods of Scaling (New York: Wiley, 1958).

For the advanced student.

8. Leonard W. Ferguson, Personality Measurement (New York: McGraw-Hill, 1952).

See chapter 4 on attitude measurement.

9. J. P. Guilford, Psychometric Methods (New York: McGraw-Hill, 1954).

See especially chapter 16 on factor-analysis, chapter 4-7 on psychophysical methods, and chapter 15 on item-analysis.

10. Lee J. Cronbach, Essentials of Psychological Testing (New York: Harper, 1960).

A useful textbook on tests and their uses. See especially chapter 5 on validity and chapter 9 on factor-analysis. Percentiles are discussed on pp. 75-78; standard scores on pp. 78-87; calculating correlations on pp. 110-125; problems of reliability on pp. 126-142.

Social-Distance Scales

11. E. S. Bogardus, "Measuring Social Distance," Journal of Applied Sociology, IX (1925), 299-308.
12. E. S. Bogardus, "A Social-Distance Scale," Sociological and Social Research, XVII (1933), 265-271.
13. Harry C. Triandis and Leigh M. Triandis, "Race, Social Class, Religion, and Nationality as Determinants of Social Distance," Journal of Abnormal and Social Psychology, LXI (1960), 110-118.

Use of a revised version of social-distance scale, developed with the aid of the Thurstone successive-interval procedure and checked for linearity, unidimensionality, reliability, and reproducibility.

Factorial Scales

- 22. Raymond B. Cattell, Factor Analysis (New York: Harper, 1952).
- 23. Hans J. Eysenck, The Psychology of Politics (London: Routledge & Kegan Paul, 1954).

See especially chapter 3 on scaling procedures and validation and chapters 4 and 5 on factor-analysis applied to attitudes and their clustering.

- 24. Hans J. Eysenck, "Primary Social Attitudes: A Comparison of Attitude Patterns in England, Germany, and Sweden," Journal of Abnormal and Social Psychology, XLVIII (1953), 563-568.

A cross-national comparison of attitude structures with the aid of factor-analysis.

- 25. Earl S. Schaefer, "Converging Conceptual Models for Maternal Behavior and for Child Behavior," in John C. Glidewell, ed., Parental Attitudes and Child Behavior (Springfield, Ill.: Charles C. Thomas, 1961).

A presentation of the two main components of maternal attitudes, in terms of factor-analysis.

Scalogram-Analysis

- 26. Louis Guttman, "The Basis for Scalogram Analysis," in Samuel A. Stouffer, ed., Measurement and Prediction (Princeton, N.J.: Princeton University Press, 1950).

The main source book for the Guttman scaling technique and its theoretical background.

- 27. L. Podell and J. C. Perkins, "A Guttman Scale for Sexual Experience -- a Methodological Note," Journal of Abnormal Psychology, LIV (1957), 420-422.
- 28. Paul Wallin, "A Guttman Scale for Measuring Woman's Neighborliness," The American Journal of Sociology, LIX (1953), 243-246.
- 29. Joseph Trenaman, "Guttman's Scalogram Analysis of Attitudes," in Attitude Scaling (London: The Market Research Society, 1960).

A useful and critical exposition of the scalogram-board technique.

39. Kurt Lewin, Principles of Topological Psychology (New York: McGraw-Hill, 1936).

Lewin's definition of behavior as $B = f(P, E)$ and its implications.

40. George G. Stern, "B = F (P, E)," Journal of Projective Techniques and Personality Assessment, XXVIII (1964), 161-168.

An attempt at studying environmental influences on perception, attitudes, and behavior.

41. Gardner Murphy, Lois B. Murphy, and Theodore M. Newcomb, Experimental Social Psychology (New York: Harper, 1937).

See especially their comments on the relationship between attitudes and behavior on pp. 898-900.

Projective Techniques in Attitude Study

General References to Projective Techniques

1. Donald T. Campbell, "The Indirect Assessment of Social Attitudes," Psychological Bulletin, XLVII (1950), 15-38.

A very useful review of projective methods in attitude research; particularly thorough on "information" tests, guessing tasks, tests of critical thinking, and pictorial devices.

2. H. H. Anderson and Gladys L. Anderson, eds., An Introduction to Projective Techniques (New York: Prentice-Hall, 1951).

A basic textbook.

3. J. E. Bell, Projective Techniques (New York: Longmans, Green, 1948).

General textbook.

4. Glaira Selltitz, Marie Jahoda, Morton Deutsch and Stuart W. Cook, Research Methods in Social Relations (New York: Holt, 1959).

See chapter 8 on projective and other indirect methods.

5. H. Henry, Motivation Research (London: Crosby Lockwood & Son, 1958).

Chapter 4 deals with the use of projective techniques in surveys.

6. G. H. Smith, Motivation Research in Advertising and Marketing (New York: McGraw-Hill, 1954).

Includes a review of projective techniques used in market research.

15. John W. Atkinson, ed., Motives in Fantasy, Action, and Society (New York: Van-Nostrand, 1958).

A series of studies using TAT measures of the needs for achievement, affiliation, and power, complete with scoring systems.

16. W. Caudill, The Psychiatric Hospital as a Small Society (Cambridge, Mass.: Harvard University Press, 1958).

See especially the use of drawings of hospital scenes as projective devices (Figures 7-3 to 7-6).

17. M. Haire, "Projective Techniques in Marketing Research," Journal of Marketing, XIV (1950), 649-656.

Shows the use of shopping lists as a projective device, to obtain the "image" of the housewife who buys instant coffee.

18. H. Proshansky, "A Projective Method for the Study of Attitudes," Journal of Abnormal and Social Psychology, XXXVIII (1943), 383-395.

A pictorial method for studying attitude to labor.

19. M. Shepherd, A. A. Cooper, A. C. Brown, and G. W. Kalton, Psychiatric Illness in General Practice (New York: Oxford University Press, 1966).

Pseudofactual Questions

20. E. L. Hartley, Problems in Prejudice (New York: Kings Crown Press, 1946).

See especially the use of "none-such" groups in attitude research, e.g., Wallonians, Danireans, and Pireneans.

21. Theodore M. Newcomb, Personality and Social Change (New York: Dryden Press, 1943).

An early use of the "Guess-Who" technique.

Play Techniques

22. Hugh Hartshorne and Mark A. May, Studies in Deceit (New York: Macmillan, 1928).

Experimental studies of children's attitudes.

23. L. E. Ucko and Terence Moore, "Parental Roles as Seen by Young Children in Doll Play," Vita Humana, VI (1963), 213-242.

- 17 -
4. Gardner Lindzey and Edgar F. Borgatta, "Sociometric Measurement," in Gardner Lindzey, ed., Handbook of Social Psychology (Cambridge, Mass.: Addison-Wesley, 1954).

An excellent textbook reference on the whole field of sociometry.

5. Robert J. Wherry and Douglas H. Fryer, "Buddy Ratings: Popularity Contest or Leadership Criteria?" in J. L. Moreno, ed., The Sociometry Reader (Glencoe, Ill.: The Free Press, 1960).

Semantic Differential

6. Charles E. Osgood, George J. Suci, and Percy H. Tannenbaum, The Measurement of Meaning (Urbana, Ill.: University of Illinois Press, 1957).

Introduces the technique of the semantic differential. See especially chapter 3 "The Semantic Differential as a Measuring Instrument," and chapter 6 "Semantic Measurement in Personality and Psychotherapy Research."

7. J. J. Jenkins, W. A. Russell, and J. Suci, "An Atlas of Semantic Profiles for 360 Words," American Journal of Psychology, LXXI (1958), 688-699.

Repertory Grid

8. George A. Kelly, The Psychology of Personal Constructs (New York: W. W. Norton, 1955).

Volume I contains the basic theory, and the development of the repertory test (see especially chapters 2, 3, and 5).

9. D. Bannister, "Personal Construct Theory: A Summary and Experimental Paradigm," Acta Psychologica, XX (1962), 104-120.

A clear, expository paper.

10. D. Bannister, "The Nature and Measurement of Schizophrenic Thought Disorder," Journal of Mental Science, C VIII (1962), 825-842.
11. L. M. Levy and R. D. Dugan, "A Factorial Study of Personal Constructs," Journal of Consulting Psychology, XX (1956), 53-57.

Stereotypes

12. W. Buchanan and H. Cantril, How Nations See Each Other (Urbana, Illinois: University of Illinois Press, 1953).

Studies of the role of stereotypes in international attitudes.

7. J. P. Guilford, Psychometric Methods (New York: McGraw-Hill, 1954).
8. Palmer O. Johnson, Statistical Methods in Educational Research (New York: Prentice-Hall, 1949).
9. E. F. Lindquist, Statistical Analysis of Educational Research (New York: Houghton Mifflin, 1945).
10. Celeste McCollough and Loche Van Atta, Statistical Concepts (New York: McGraw-Hill, 1963).

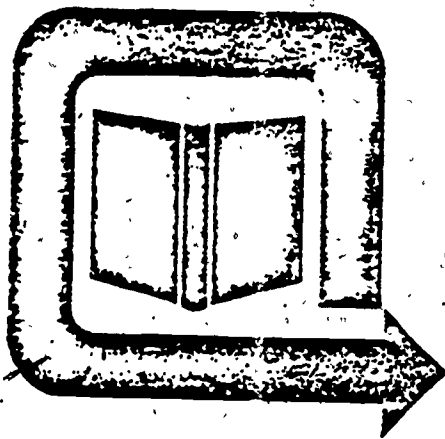
A useful programmed learning text.

APPENDIX C

EVALUATION ARTICLES AND REPORTS

A BLUEPRINT FOR PROGRAM EVALUATION

Kathleen Brophy
Arden Grotelueschen
Dennis Gooler



Occasional Paper Number One • February 1974

A BLUEPRINT FOR PROGRAM EVALUATION

by Kathleen Brophy, Arden Grotelueschen, and Dennis Gooler

Why an evaluation plan?

A plan for evaluation may be viewed as a blueprint which provides information about the intended activities (and implicit philosophy) of an evaluation endeavor. A plan allows the evaluator to explain his conception of the evaluation task to the client before implementation. The client then has an opportunity to see if he approves of the evaluator's intended actions and may restate his concerns if he feels this is necessary.

Certain elements should be taken into account in designing any evaluation plan, but each plan must necessarily reflect the uniqueness of the particular program being evaluated. The commentary and checklist featured in the following pages suggest some design considerations.

Kathleen Brophy is a graduate assistant in the College of Education at the University of Illinois at Urbana-Champaign. Arden Grotelueschen is associate dean for research and service in the College of Education. Dennis Gooler is an alumnus of the College of Education and is currently chairperson of the Area of Instructional Technology at Syracuse University.

Why evaluate?

In general, evaluation is conducted in response to a need expressed by people within a program, or individuals or agencies interested in a program but external to it. Evaluation may focus on what has been done, what is being done, and/or what might be done.

Evaluation can assist in:

Planning procedures, programs, and/or products;

Improving existing procedures, programs, and/or products;

Justifying (or not justifying) existing or planned procedures, programs, and/or products;

Evaluation may yield:

Descriptions of procedures, programs, and/or products;

Judgments of the value of procedures, programs, and/or products.

Evaluating for program planning

What program preferences and needs are held by various people and institutions?

What discrepancies exist between desired and actual status of various program elements?

What means are feasible for attaining desired goals?

What will happen if proposed goals are attained?

Who are the advocates and adversaries of the proposed program?

Evaluating for program improvement

Are strategies working as planned?

What unanticipated results are affecting program operation?

Are program issues being dealt with satisfactorily?

Is there sufficient program flexibility to meet new preferences or concerns?

Is the program content appropriate and effective?

Are internal/external relations hindering or enhancing program strategies?

Is the program as effective as similar programs?

Evaluating for program justification

Are the goals of the program justifiable as viewed by various people?

What was not attempted because of the existence of this program?

What exactly was done and what were the results?

Should the program be altered, expanded, or disbanded?

What do advocates and adversaries want to know about the program?

Who is the evaluation for?

The people who are typically most interested in an evaluation are the individuals or groups commissioning the evaluation, the persons involved in conducting the program being evaluated, the program participants, and the supporters or detractors of the program. Special attention should be given to all these audiences, especially the audience that actually commissions the evaluation. The evaluator must be responsive to each audience and recognize that different kinds of information may be desired by or useful to different groups.

Potential audiences or consumers

- Students/participants
- Teachers/resource persons/facilitators
- Program administrators, implementers, and sponsors
- Policy makers and public officials
- Community interest/action groups
- Academic discipline specialists
- Technical and methodological specialists
- General public
- Other: _____

(Specify)

What focus should the evaluation have?

Both evaluator and client should be able to define those variables which they believe have an important influence on the specific situation being evaluated. The delineation of these variables will depend greatly upon the concerns of the client, the orientation of the evaluator, and the sources used by both to obtain this initial information. Further, the relative importance of variables may shift as the program and/or the evaluation proceeds. New variables may be delineated. The focus of the evaluation will reflect the aggregated viewpoints of the various audiences (evaluator, client, and others) who might be affected by the processes and results of an evaluation. In most instances, there will be many more potentially interesting variables than can be reasonably examined. An evaluation plan represents a statement of variable priority. Although priorities will differ across programs, all evaluative efforts should seek to illuminate the wider social, economic, and/or political context within which the program operates.

Information desired by interested audiences

- Program dynamics and processes
- Program outcomes
- Program context and setting
- Program issues

Information needed to make decisions

- Content
- Instructional/administrative strategies
- Student participant capabilities
- Personnel and institutional resources

General categories of information

- Program intents
- Behaviors
- Interactions
- Interrelationships
- Causes/effects (including side or unintended effects)
- Costs
- Successes/failures

What should you know about the evaluator?

Any process used to determine the merit or worth of a program necessarily involves evaluator judgments. Judgments are reflected in the selection of variables requiring attention, the sources from which data are obtained, the techniques used to gather information, and the messages finally conveyed to an audience. Because so many judgments may be made, the client may wish to inquire into the philosophical and methodological orientation of the evaluator, his motivation for wishing to conduct an evaluation, his knowledge of the problem under study, his experience, his capacity to work with people, and his ability to report information. Such information about the evaluator (similar questions might be raised about the client) provides insight into how a particular evaluator may make judgments. The client is then better prepared to select the appropriate evaluator for the specific program, procedure, or product to be evaluated.

Practical background

- Experience in evaluation
- Professional experience
- General work experience

Educational qualifications

- Academic background
- Experience in education
- Professional interests

Orientation

- Perspective on purpose of evaluation
- Service/practitioner
- Research/theoretician
- Outcomes/processes
- Prespecified/responsive

Methodology

- Academic discipline related (e.g., historical)
- Informal techniques
- Behavior analysis
- Judgmental
- Descriptive
- Interaction analysis

Past standards

- Knowledge of other evaluations
- Reputation
- Delivery on what was asked, when needed
- Trustworthiness

Style

- Responsiveness to client
- Personable
- Ability to work with others

What are the moral, ethical, and legal responsibilities of the evaluator?

An evaluation may examine many critical and sensitive aspects of a program. Consequently, a number of conflicts can arise among clients, evaluators, and audiences. An evaluation plan should reflect an awareness of potential problems, and should specify to which existing information (e.g., student records) the evaluator may have access, and what he may do with such information. Agreement should be reached as to anonymity of sources of information — when such anonymity is deemed appropriate. The plan should stipulate as clearly as possible contingencies on delivering the promised information, by a predetermined date. There should be agreement as to who will have access to the evaluation report, and who will not. All of these elements affect the way in which an evaluation may be conducted and reported, and the obligations incurred by both evaluator and client.

Establish guidelines

Penalties for not meeting expectations (especially deadlines)

Extent of report — reporting only what is asked for, or giving information beyond original request

Reporting detrimental information

Access to evaluation information — who will have access

Authority of the evaluator or client to determine what to report

Generalizability of results and implications

Specify the client's parameters of power to terminate the evaluation process

Conduct of the evaluation

Use confidential information appropriately

Avoid invasion of privacy

Obtain permission for use of certain information

Maintain open exchange of ideas between client and evaluator concerning each other's interests

Tell participants in the program what the evaluation will involve and provide

How do you conduct an evaluation?

In an attempt to focus the evaluation, the client and evaluator typically define critical audiences for the evaluation, and the issues or variables to be examined. A next step in evaluation is an identification of the information relevant to each issue or variable, and the likely sources of that information (e.g., learners, citizen committees, class activities). Next, the evaluator must determine how much information should be collected, and from what proportion of the data sources (which sampling procedures to use). Once these estimates have been made, appropriate instruments and procedures should be selected or constructed and data gathered. Finally, the comparison of data (including judgments) with appropriate standards is also essential.

Questions about timing of information collection, explanations as to why certain information is being collected, and means of coding and storing information must also be considered. The conduct of an evaluation will interact with or intervene in the conduct of a program. Therefore, it is important that careful consideration be given to the process of phasing evaluation into program operations.

Identifying primary audiences

- Program funders
- Program managers or administrators
- Program participants
- Program or product consumers
- Others who may exert an influence on the program (e.g., content-specialists, legislators)

Identifying critical issues

- Outcomes
- Processes
- Costs
- Consequences
- Justifications

Data sources

- People
- Existing documents or records
- Related research in evaluation

Techniques for collecting data

- Questionnaires and opinionnaires
- Interviews
- Observation schedules
- Active participation in program
- Historical inquiry
- Standardized tests
- Rating scales
- Behavior analyses
- Checklists
- Attitude scales
- Interaction analyses
- Utilization of biographic data
- Anecdotal records

Techniques for gathering standards

- Statements by program personnel
- Statements by content experts
- Reports/recommendations by boards, commissions, etc.
- Statements by regulatory agencies

How do you report the evaluation?

People have different abilities and different experiences which influence the way they receive and use evaluation reports. Both evaluator and client should take these differences into account when interpreting and reporting evaluation information. Reporting standard scores from an achievement test, for example, may be meaningful information to some audiences. For other audiences, information may need to be presented in a different form. Similarly, the inferences drawn from certain information may be useful and defensible when presented to one group, but may be inadequate (or even have a negative effect) when presented to another group. It is important to understand what criteria various people use to judge programs, what standards are relevant and meaningful, what indicators people accept as legitimate, and in what form and in what language all of these things are to be meaningfully discussed. Finally, the evaluation report must be timely — there is often too little information too late.

Types of reports

Written

Oral

Progress reports

Final reports

General

Specific

Technical

Nontechnical

Descriptive only

Evaluative and judgmental

Makes recommendations

Does not make recommendations

Modes of display

Case studies

Portrayals

Graphs and charts

Test score summarizations

Scenarios

Questions/answers

Product displays

Multimedia presentations

Simulations

Dialogues/testimonies

How do you encourage people to use evaluation?

The process of evaluation, as well as the utilization of evaluation results, requires a clear understanding of the social, personal, and political dynamics surrounding any educational activity. An evaluation plan should show how the results of an evaluation will "fit" into these dynamics, if any satisfactory use is to be made of the evaluation information.

People will be more likely to use evaluation information if they perceive the process and/or resulting information as relevant to their needs, if they see the potential benefits deriving from the evaluation, and if they find the evaluation results to be available when needed. People who have been involved in the original planning and implementation of an evaluation are more likely to use evaluation results than those who were not.

Ways to encourage use

- To develop commitment, involve people from inside and outside the program in the evaluation process
- Report results when desired by project staff
- Indicate alternative courses of action
- Indicate implications of the findings
- Make presentation clear and attractive.
- Provide assistance for additional analyses of results as needed
- Provide technical assistance for implementation of evaluation suggestions
- Provide time for the study and use of evaluation findings
- Results should be valid
- Reporting of results should take into consideration the background and orientation of audience
- Reporting must show sensitivity to political/social factions

What does it cost to do an evaluation?

Conducting an evaluation necessarily requires expenditure of resources: money, time, and personnel. An evaluation plan should include estimates of cost in terms of dollar outlay, time expenditures of staff and others (such as those from whom data are sought), particular expertise needed, etc. Evaluation may be more critical to some kinds of programs than to others, thus making greater allocations of resources legitimate in some cases. However, some form of evaluation can fit into any budget. Reasonable cost can be determined by an estimate of the significance of the issues and the likely impact of the evaluation. The value of conducting an evaluation may also be influenced by the availability of information about similar programs, procedures, and products, and the availability of existing information on the program to be evaluated.

Money

Evaluator fees

Travel and accommodations

Supplies and communication

Dissemination and implementation of findings

Costs to program personnel

Time of personnel in gathering information about various aspects of the program

Time and effort in facilitating evaluator requests

Explaining and defending evaluation to program constituents

Considering, rejecting, or modifying evaluation suggestions

Potential costs of acquiring information

Invasion of privacy

Interpersonal relationships affected negatively

Psychological influences on specific individuals

Other difficulties

Formation of defensive attitudes in people involved in the program

Negative reactions of audiences not considered

Controversy generated because certain issues not considered

IX. COMMON EVALUATION HAZARDS AND HOW TO AVOID THEM

Each of the following subsections provides a brief description of a common evaluation hazard and what should be done to avoid it. The hazards are presented in the approximate order of their judged frequency of occurrence and severity. Not all hazards, however, are relevant to all evaluation designs. An attempt has therefore been made to describe as early as possible in each presentation the situations where each hazard is likely to arise. Unfortunately, because many of the hazards are relevant to two or more designs, it was not feasible to organize this chapter according to evaluation design.

Hazard 1: Claiming Much, Providing Evidence of Little

The goals of an innovation may be multiple and grand; the evidence may be modest and limited, as in one instance where preventing school failure through increased parent involvement, home/school cooperation, building positive attitudes, and academic gains were claimed while the only data were pre-post measures of spelling.

In another submission, it was claimed that participants would get in less trouble with the law. Evidence for this claim was not included in the submission; in a discussion, the claim turned out to be based on the fact that "The children are being watched from 8:30 to 3:30 every day. This would help keep them out of trouble." In the same submission, the claim that regular progress during the fall-spring year resulted from the summer program turned out to mean "If our children gained three months in six weeks, they'd be gaining the equivalent of one year in eight months."

The panel can only reach judgments based on the evidence; where evidence matches the claims, a favorable decision is far more likely than where evidence falls far short of goals, objectives, and claims.

Hazard 2: Selecting Measures Not Logically Related to the Intervention

Valuable as standardized achievement tests may be for assessing many interventions, in many instances there are other, more sensitive, more appropriate sources of data. Measures should be selected because they will assess the central consequence of the intervention, be it improved career decision making, adequate knowledge of human reproduction or contraception, improved consumer skills, minimum competencies for self care in dressing and toileting, increased pride and self confidence as a member of one's ethnic or racial group, improved teacher classroom control, or a better school planning and budgeting system.

For almost all educational outcomes, reliable sensitive measurement is possible, using techniques as varied as analysis of basic school records, interviews, questionnaires, case studies, observations, and tests.

Hazard 3: The Use of Grade-Equivalent Scores

Grade-equivalent scores present a severe methodological problem when used in norm-referenced evaluations. No norm-referenced evaluation based on grade-equivalent gains can be accepted. When used with control-group or regression-evaluation models, the methodological problems are far less severe although problems of interpolation remain.

Grade-equivalent scores provide an insensitive, and, in some instances, a systematically distorted, assessment of cognitive growth. The concept of a grade-equivalent score is misleading--for example, a grade-equivalent score of five attained by a third grader on a math test does not mean that he knows fifth-grade math. Possibly he can do third-grade math as well as the average fifth grader, but it is likely that no fifth-grade students have ever taken the third-grade level of the test.

The use of grade equivalents for evaluation purposes creates a second problem in that they do not form an equal-interval scale and should never be averaged. Finally, grade equivalents are constructed based on the assumption that growth occurs at the same rate throughout the school year.

Research has shown, however, that learning typically does not follow this regular pattern and, consequently, gains measured in grade equivalents will be artificially inflated or reduced.

To avoid this hazard, it is recommended that normalized standard scores (often provided by test publishers) be used for computations and that gains be interpreted through reference to pre- and posttest percentile standings with respect to a norm group.

Hazard 4: The Use of a Single Set of Test Scores for Both Selecting and Pretesting Participants

When students are selected for participation in a special group because they obtained relatively high or relatively low scores on some test, use of these scores as pretest measures invalidates any kind of norm-referenced evaluation. It can also invalidate control-group evaluations when the treatment and control groups are drawn from different populations (see Hazard 8). This problem stems from what is known as "statistical regression," "regression toward the mean," or simply, the regression effect.

If students are selected from a group because of their low test scores and are retested on the same or a comparable test, they will tend to score higher on the second testing, while an initially high-scoring group will score lower. The result is that low-scoring groups appear to learn more from a special program than they actually do, while gains in special programs for high-scoring students may be obscured.

To avoid this hazard, students should be selected for participation in a special treatment based on one set of test scores and then be pretested using an alternate form of the same test or a different test. Another alternative is to base student selection on some form of needs assessment other than test scores.

Hazard 5: The Use of Comparisons with Inappropriate Test Dates for Obtaining Information

In norm-referenced evaluations, tests should be administered at nearly the same time as the test publisher tested

the norm group. When control groups are available, few evaluators would consider testing the treatment and control groups more than a few days apart. When norms are used as a substitute control group, this same consideration needs to be given to test dates.

To avoid this hazard, treatment group students should be tested within two weeks of the midpoint of the interval during which the normative data were collected. Particularly, where the evaluation spans a summer, both pre- and posttest should deviate from the norming date in the same direction and by the same amount of time. Testing within six weeks of empirical normative data points is permissible if linear interpolations or extrapolations of the normative data are made. Tests that provide normative data for only one point in the year should not be used in fall-to-spring norm-referenced evaluations. The principles of comparability in time of data collection also apply to other outcomes, such as job placement rates that will show systematic seasonal variation.

Hazard 6: The Use of Inappropriate Levels of Tests

If most of the pupils achieve very high or very low test scores, the level of the test may be inappropriate for assessing their performance. If pupils encounter the test floor at pretest time or the ceiling at posttest time, treatment effects will be underestimated. Conversely, if the ceiling is encountered on the pretest or the floor on the posttest, gains will be overestimated. Ideally, students should score in the middle of the range of possible raw scores. Average performance for groups of students should generally fall between 30 and 75 percent of the total number of items answered correctly.

To avoid this hazard, test levels should be selected on the basis of the achievement levels of the students, not on the basis of their grade in school. In most cases, the nominally recommended test level or one level above or below will be suitable for testing most groups of students.

Using a test level other than that nominally recommended for a particular grade is likely to mean that norms tables for the tested students are not included in the test manual for that level. However, it is not meaningful to assess

either status or gains by comparisons with students at a different grade level. The status of a sixth grader should be assessed using sixth-grade norms even if he is tested with a fourth-grade test. Most major test publishers, fortunately, have interlocked their test levels by providing an expanded standard score scale which enables the determination of score equivalencies between adjacent test levels. These scores make it possible to predict from a pupil's score on one test level how he would have scored on the next higher or lower level, thus providing access to the in-level norms. Another possibility (although rare) is that the manual may provide normative data for pupils tested out of level.

Hazard 7: Missing Data

Regardless of the evaluation model used, data analysis should be based only on those students with both pre-treatment and post-treatment scores. Interpretation of these data, however, should take into account the characteristics of the students who dropped out, entered late, or graduated from the project. For example, if all of the lowest scoring students on a pretest dropped out before posttest time, the average posttest score would increase with respect to the pretest scores simply because of the missing students. This increase could be misinterpreted as a gain. Likewise, if the high-scoring students graduated from the group, the mean posttest score would be artificially deflated.

To avoid this hazard, every effort must be made to obtain pre-treatment and post-treatment data for each project participant, and to base comparisons on those students for whom both data elements are available. Data from students having only pretest or only posttest scores must be carefully examined to see if they differ in some systematic way from the data of students having both pre- and posttest scores. A description of any of these differences should be included in the evaluation report.

Hazard 8: The Use of Noncomparable Treatment and Control Groups

In "true" experimental designs, treatment and control groups should be similar in all educationally relevant

respects before the treatment begins. Groups that differ in terms of pretest scores present an obvious source of bias. Other more subtle factors such as differences in age, sex, race, or socioeconomic status can also exert strong biasing influences and should be avoided. Nonvolunteers should never be used as controls for pupils who volunteered (or were volunteered by their parents) for a particular instructional treatment.

Whenever possible, students should be assigned to treatment and control groups on a random basis. For example, with a semester-long reading program, pupils could be randomly assigned to first- or second-semester groups. For the first half of the year, one group would serve as the control group for the other, but both groups would ultimately receive equal amounts of the treatment.

In some cases, pre-existing groups will be enough alike so that they can appropriately be considered equivalent to random samples from a single population. In other cases, a control group will be known to differ systematically from the treatment group. Where the difference is small, the control group model may still provide the best method of evaluating the project, and statistical adjustments can be made to compensate for between-group differences. Where the differences are large, however, there is no way in which a noncomparable control group can provide an accurate estimate of how well the treatment group would have done without the treatment.

Hazard 9: The Use of Inappropriate Statistical Adjustments with Nonequivalent Control Groups

There are several statistical procedures that are widely used in an attempt to compensate for initial differences between treatment and control groups. Some are legitimate while others are not. Making between-group comparisons using either "raw" gain scores or "residual" gain scores falls into the latter category. Both procedures should be scrupulously avoided.

A raw gain score is simply the difference between a pre- and a posttest score and reflects the gain made between testings. It is argued that, although two groups may have been

somewhat different in terms of initial achievement levels, their expected gains would be roughly comparable after the same educational treatment. This would be true, however, only when each group's posttest standard deviation is the same as its pretest standard deviation. Where the posttest standard deviations are larger than those of the pretest scores, a raw gain score analysis will systematically underestimate treatment effects. Conversely, the procedure will systematically overestimate treatment effects where the standard deviations of pretest scores exceed those of posttest scores.

A residual gain score is the difference between an actual posttest score and a posttest score estimate derived from the combined treatment and control group regression line. Presumably the mean residual gain score for a group that received an effective treatment would be positive while that for the control group would be negative. Also, the sum of the absolute values of the two differences would provide an index of the size of the treatment effect. Unfortunately, it can be shown algebraically that a residual gain score analysis always underestimates the size of the treatment effect except where the groups' pretest scores are equal. Furthermore, the amount of underestimation is directly proportional to the size of the initial difference between groups.

There are other factors, such as how the treatment and control groups were formed, which determine the appropriate adjustment procedure to compensate for their initial differences. The issues involved, however, are very complex. Expert advice should be sought.

Hazard 10: Constructing a Matched Control Group after the Treatment Group Has Been Selected

Finding "matches" for treatment participants in some other group is a fundamentally unsound practice. Unless they and the treatment pupils are equally representative of the groups from which they are drawn, statistical regression will act differentially on the two groups and artificially inflate the apparent gains of one group with respect to the other.

In the most common situation, the group(s) from which the matching control pupils are drawn will be higher achieving than those from which the treatment group pupils are selected. Consequently, the control group pupils will be farther below the mean of the group(s) to which they belong than the treatment group children. On retesting they will thus show greater statistical regression and their posttest scores will be too high to serve as a no-treatment expectation for the treatment group participants.

The correct procedure for establishing matched control groups is to do the matching first and then assign members of each pair randomly to the treatment or the control group. That is, a large group of students, all eligible to be in the project, must be available. The first step is to divide the group into matched pairs based on test scores, ethnic background, sex, etc., so that the two members of each pair are as similar as possible. Then, after the matching process is complete, some random procedure such as flipping a coin should be used to decide which member of each pair goes into the treatment and which into the control group.

Since this procedure will rarely, if ever, be possible in real-world situations, the only way to avoid this hazard is to avoid matching. A better practice is to work with unmatched and therefore slightly different groups and to correct for between-group differences with appropriate statistical adjustments, such as some form of covariance analysis. Again, practitioners should seek the help of experts.

Hazard 11: The Careless Collection of Data

Testing and other types of data collection must be accomplished with scrupulous attention to detail. For most evaluation models, the primary requirement is that treatment and control or comparison groups be treated in exactly the same way. Minor variations from the procedures described by the test publisher are permissible in control group models and in certain quasi-experimental designs. It is essential, however, that treatment and comparison groups be tested in exactly the same manner. In norm-referenced evaluations, treatment group students should be tested in the same way as the students in the norm group. This requirement means that procedures outlined by the test publisher must be followed precisely.

Problems arise if tests are administered or scored in an inconsistent and careless manner. If there are differences in the ways in which the test takers and the norm group students are tested or if there are differences in the procedures, conditions, and scoring at pretest and posttest times, then it is impossible for the resulting data to be accurately interpreted. There are no statistical manipulations that can compensate for mistakes made in administering or scoring a test.

To avoid this hazard, the following steps should be taken:

1. Test procedures must be orderly and accurate if scores are to be meaningful.
2. Test administration and scoring procedures must be exactly the same for the treatment group as for the control, comparison, or norm group used to generate the no-treatment expectation. Testing treatment group pupils in exactly the same way as pupils in the norming sample means following the test publisher's directions in every detail.
3. The procedures, conditions, and scoring methods used during posttesting must be exactly the same as those used during pretesting.

Hazard 12: The Use of Different Instruments for Pretesting and Posttesting

In the norm-referenced design, it is not advisable to change tests between pre- and posttesting because there is no adequate way to compare pretest scores on one test with posttest scores on a completely different test. Since each test publisher follows slightly different norming practices, it is likely that one test's norms will be slightly "easier" than another's. These differences do not matter if the same test is used both pre and post but could magnify or obscure real gains if changes were made. While it is not essential to use the same form and level of an achievement test pre and post, this practice is also recommended.

46

Some tests have been developed so that the lower levels are intended for use at the end of one grade and the beginning of the next. In these instances, to use the same form and level of test for fall pretesting and spring posttesting, it will be necessary either to pretest or posttest out of level. In some grades where spring-to-spring or fall-to-fall evaluations are conducted, it may be necessary to change test levels in order to avoid ceiling or floor effects; unfortunately, this practice will introduce an unknown amount of error into the measure of gain.

Hazard 13: The Use of Inappropriate Formulas to Generate No-Treatment Expectations

Many projects use an unrealistic theoretical model or formula to calculate "expected" posttest scores from IQ or other pretest scores. If students do better than the calculated expectation, the project is considered a success.

Many methods have been devised for calculating performance-level expectations that rest on untenable assumptions. Neither IQ scores nor grade-equivalent scores should be used to generate no-treatment expectations. For example, a student who has gained .7 year per year, on the average, since beginning school, is presumed to continue at the same rate unless a special program increases his rate. Unfortunately, grade-equivalent gains measured from fall to spring will usually exceed this rate--even for typical Title I children--and treatments will appear to be more effective than they really are.

In norm-referenced models, this hazard may be avoided by generating the no-treatment expectations solely from empirical percentile norms tables. When control groups are used, the actual posttest scores of these groups provide the proper basis for evaluating treatment effects.

Hazard 14: Mistaken Attribution of Causality

Observed gains may have resulted from the Title I treatment, but there are always plausible alternative explanations. The plausibility of these alternative explanations should be carefully examined before evaluation results are attributed

THE CAMERA AS AN EVALUATION AND RESEARCH INSTRUMENT;

SNAPSHOTS OF A SCIENCE CURRICULUM

by

Edward L. Wachtman

Contract Research Corporation
Cambridge, Massachusetts

Prepared for the Annual Meeting of the
American Educational Research Association

Toronto, Ontario

CANADA

1978

48

116

Printed in U.S.A.

Introduction

This paper is in two parts. The first section will deal with the use of photography in an actual evaluation study. It will be a loosely stated case-study of how one evaluator was able to utilize the camera in several related but distinct ways. This section is written after the fact - with the evaluator reflecting upon his use of the camera during a series of on-site evaluation visits. Many of the thoughts in this reflection have been influenced by Susan Sontag's book On Photography and the author owes a great debt to Sontag for articulating some of the questions and issues that are related to this paper and which should have an impact on the use of the camera as an evaluation or research tool.

The second section of this paper will discuss some of the broader issues that are raised with regard to the use of the camera as a research tool. These issues will focus on some of the possible limitations and advantages that should be considered when using the camera as an evaluation tool and a reporting medium.

Overview

In 1976, a team of evaluators from Syracuse University's Center for Instructional Development attempted to meaningfully describe and portray the workings and characteristics of a National Science Foundation funded curriculum known as ESSENCE or Environmental Studies. This curriculum presented a unique challenge to the evaluation effort because it claimed to be free of specified behavioral objectives and even went so far as to state that "Evaluation raises hell with trust." This challenge was heightened because ESSENCE also claimed to be content-free. Its goal, if we can say that it had one, was to introduce children to the process of inquiry; content was viewed as a vehicle for this process.

After reviewing the ESSENCE materials, it was felt that the best way to approach ESSENCE would be portrayal - the telling of the story of ESSENCE. We

hoped this would provide the client - the Director of the Environmental Studies Institute at Syracuse University - with a meaningful understanding of the curriculum as it was actually used in the schools. Therefore, the aim of our inquiry was primarily description. We attempted to make no claims of merit per se. Rather we hoped to capture the various characteristics and activities of ESSENCE as we discovered them so that those who were not present when these actions were taking place might still be able to share our appreciation of some of their meanings. This would enable the client to come to his own conclusions and make his own judgments.

We decided early in our planning that we would use a camera as one of several instruments for the site-visits in order to capture and portray what ESSENCE was in real classrooms, with real teachers and real students. The study consisted of three major components: site-visits to schools where ESSENCE was or had been used; a separate review of the literature on affective education that we hoped would link our observations with the research literature; an extensive questionnaire sent to participating teachers and administrators to be responded to in writing or on an enclosed cassette cartridge. The camera was used in the site-visits.

The notion of using a camera as one of our research tools was reinforced by some of the ESSENCE activities which we were probing. One of the more highly-touted activities of the ESSENCE curriculum was the use of a Polaroid camera by the students as a tool of inquiry. The students were given Polaroid cameras with the assignment of "take a picture of something invisible," or "take a picture that is some positive evidence that something natural happened," or "go outside and take a picture of power," or "take a double exposure that portrays some change." ESSENCE purported to be a process of inquiry, and the camera was

50

viewed as instrumental in this process. It focused student attention on a specific problem or issue. It framed specific actions to the exclusion of others. It recorded how things change and/or remain the same.

Because the evaluation team was uncertain as to what to expect while conducting our study of ESSENCE we adopted a certain ESSENCE point of view. This point of view stressed process rather than specified outcome, openness to whatever or wherever the study lead, and a willingness to deal with a certain amount of uncertainty as our inquiry unfolded. Therefore, we decided to use the camera as an instrument for focusing on specific issues, and as a means of visually recollecting some of those myriad events and activities that were witnessed during a hectic three weeks of site-visits. And, most importantly, we hoped that what was captured on film would be useful to the client in better understanding some of the very complex and unique characteristics of ESSENCE and the people who utilized it.

Three Uses of Photography in the Portrayal of a Science Curriculum

Because our knowledge of what ESSENCE comprised was very vague, the camera initially functioned as a means of framing and isolating certain actions and events. The use of the camera as a means of focusing upon one specific event or activity during the site-visits was a value-laden act in itself. Implied in the photographs of the site-visits is the assumption that what was photographed was somehow worth considering as a small but isolated portion of our portrayal study.* The camera forced the site-visitor to, at least momentarily, focus his attention on a specific activity. In what could loosely be described as a goal-free evaluation the camera became an instrument of goal-seeking, helping the site-visitor search

* Stephen Kemmis speaks to this issue briefly in "Telling it Like it is: The Problem of Making a Portrayal of an Educational Program, January 1974, University of Illinois, Champaign-Urbana, (mimeograph).

out patterns and look for order in the seemingly unordered activities going on continually around him. From the very beginning, the site-visitor was compiling a series of individual images - both mentally and photographically which would, much later in the study, be pieced together to tell the story of ESSENCE. Just as the Polaroid cameras were used by the students to focus their questions, their inquiry, the site-visitor utilized his camera (Olympus 35PC) to bring order and clarity of value to his study. This was especially important in the early stages of the study when we were never certain what to expect from ESSENCE or even what questions to ask about it.

This ordering was vital to our study because we understood from the beginning that ESSENCE, because it concerned itself more with process than product, could mean many things to different people. Our objective was to portray, in the most meaningful way possible, "A Story of ESSENCE." The emphasis in this title must be put on the article "A" - because we were convinced that what we were portraying was but one of a number of stories that could be written about this National Science Foundation curriculum project. The camera was one way events and activities were singled out and isolated and then pieced together to re-create one of many possible stories that would accurately and adequately portray ESSENCE.

The camera was also used as a primary mnemonic device, along with a portable tape-recorder, to aid in the site-visitor's later recollections of what he observed during his visits. This took some of the burden off the evaluator in terms of trying to keep track of the many happenings that he experienced during his visits. The camera and the tape-recorder almost completely replaced note-taking in the site-visits, thus leaving the site-visitor free to more readily "grasp the insider's view of the program" (Kemmis, 1974): asking probing questions of

teachers and learners; watching for subtle interactions between teachers and children as well as between children and children; and trying to understand the total curriculum without fear of missing some of its more specific components.

There is a sense of professional well-being in knowing that one is collecting as much usable evidence as possible in any evaluation study. This was particularly true in the site-visitor's observations of ESSENCE in action. Not knowing beforehand what to expect, it was comforting to know that data (in the form of taped conversations and photographs) were being accumulated, and that these data would be helpful, in fact necessary, in piecing together a story, an adequate portrayal of this curriculum. Like the atomic physicist trying to capture images of sub-atomic particles, we could only speculate about what we would find - but were confident that if there were something to capture on film, we would discover it. And like the physicist, we were very aware that we might not fully understand what it was we were observing until well after the study's on-site observations were completed. The photographs taken during the visits to the schools were a form of insurance that guaranteed that the site-visitor would have something to say about what he "saw" during his travels. The camera was his guarantee that he would not forget the obvious - that those small details that give meaning to the total picture would not be omitted.

Although the camera was utilized both as a means of focusing on specific action and as a memory back-up device, we felt from the beginning of the study that good photographs of the curriculum in practice would enhance our study of ESSENCE. We hoped to make our report as descriptive as possible and we did not want to rely solely on words. To paraphrase the photographer Lewis Hine if we felt we could tell the story in words alone, we wouldn't have needed to lug a camera around (Sontag, 1977). The ESSENCE study is a story of words. It is a story about people.

young and old, together engaging in the activity of teaching and learning.

"The Story of ESSENCE" relies heavily on the real words of real people - but where it was possible, photographs were used to illustrate the objects of conversation and to heighten the worded story. Thus, the camera was a most helpful tool in helping us report our findings to the client in a way that was both factually informative and meaningful. Although the photographs by themselves are only isolated images, mere bits and pieces of evidence, when they were paired with certain conversations or used to illustrate particular circumstances, we believe they added depth and human feeling to the report. There were several instances when it was felt that the photographs led to a better understanding of a conversation that was taking place or a unique activity that was happening.

When the taped conversations and photographs are used together they take on a reality themselves, a reality that may be greater than the sum of each taken separately. This reality is basically independent of the recorder even though it is his point of view that informs what questions are to be asked and what segments of visual reality are to be captured on film. This marriage of photography and taped conversation gives the client a sense of autonomy in his judgments and decisions. He becomes a vicarious participant in the events that have taken place and is freed from his reliance on the verbal interpretation of the event through a mediating voice. The use of photographs especially in portrayal-type evaluations allows the client to come to more informed conclusions, to make his own decisions about the worth of a program, curriculum, or other activity. Without dealing with the question of the subjectivity or objectivity of the photography in evaluation research, the camera can help the client in better verifying his own opinions about the findings and recommendations of an evaluation report. Photographs are additional types of evaluation evidence, and if they are not used surreptitiously and deemed admissible within the particular context of the evaluation study they

can add a sense of visual validity to more traditional modes of evaluation.

Issues and Problems Surrounding the Use of Photography as an Evaluation Instrument

In the previous section the usefulness of the camera as an evaluation tool was briefly, but positively discussed. This raises some questions as to how different the camera is from other evaluation tools - for example statistical analysis. In fact, statistical analysis and the use of photographs in evaluation research may be analogous - both focus upon one event or activity to the exclusion of others; and both have a quasi-mystical fascination for some people who see both as some symbolic representation of "Truth" or reality.

The fact that photography is but a tool in the evaluation process can not be underplayed. It is instrumental in achieving other ends (description, determination of worth, reasonable decisions) and is not an end in itself. One uses a shovel to dig a hole; and unless one is a craftsman or tool designer, the type and design of the shovel are relevant only in so far as they fit the needs of the job to be done. The camera will sometimes fit the needs of the job to be done, sometimes it won't. We have been warned by Robert Stake that evaluators "should choose methods to fit the issues, even if forced then to employ weaker and less respectable methods" (Stake, 1977). The evaluation team who compiled the data for "A Story of ESSENCE" and the client for whom it was done, felt that in this case the use of the camera was a "respectable method." However, in another situation, at another time, another method, a different tool might be more appropriate.

The fact that the camera excludes some activities to the inclusion of others is not a positive or negative aspect in itself. Rather, it is how we deal with this limitation that makes it relevant to evaluation research. The site-visitor of the ESSENCE study believed that the act of focusing on some events and excluding others was a valuable endeavor in bringing an initial sense of order to a seemingly disordered array of events. The benefit of focused specificity in this case was

seen as a necessary beginning to the formulation of an adequate overview of the curriculum. The problem at issue here is how much specific focus is enough to pick out the important components of a program and how much is too much, so that the total overview is lost in obscure detail. Again it is Robert Stake who best articulates this question "What is more important: to tell of some very special things about the program or to provide the most veridical portrayal of the program?" (Stake, 1972). Stake advocates the latter. Is it possible that the prudent use of photography as an evaluation tool might enable evaluators to both - tell some special things about a program without detracting from the overall portrayal of that program? It is not so much that "special things about a program" necessarily detract from its "most veridical portrayal" - rather the perseveration of the evaluator with these "special things" that cloud and distort meaningful program portrayal. This can be said about any evaluation method or tool that is seen as an end in itself rather than an instrument to another more informative, value-based end.

It is this infatuation with the tools of the evaluation trade rather than with process of valuing itself, that must be avoided. The camera does not seem to offer any great panacea to the problems of evaluation research. It does offer a novel point of view to the field. But just as the limitations of a complex statistical analysis or research design must be kept up-front in evaluation research - so too the limitations of the camera must be fully explored and dealt with.

Because the number of photographs that can be taken is virtually unlimited, photography reinforces a segmented view of reality as consisting of small, separate units of an apparently infinite number. This makes the world more manageable because the photographer or the viewer deals with only one aspect of reality - one separate photograph - at a time (Sontag, 1977). It was this manageable nature

of photography which aided the ESSENCE site-visitor - bringing order to the initial stages of the study.

The individual photographs only showed how separate and distinct parts of the curriculum looked. They did not show how it functioned (Sontag, 1977). How something functions implies action and time is a necessary condition of action. Individual photographs of ESSENCE did not, by themselves, have the power to meaningfully convey the workings of the curriculum over time, even when the photographs were viewed as a group. It was the narrative quality of the taped conversations which placed the activities of ESSENCE (and the photographs) within the context of time. Narrative deals with time and action over time. It is only within the context of time that we can talk meaningfully of process. An evaluation study that stresses process must deal with action over time. Narrative, therefore, becomes a necessary condition of any process-oriented evaluation.

Photography and the discriminate use of the camera have a great potential as useful and informative evaluation tools. The evaluation profession must make certain that these tools are not elevated to ends in themselves - but rather that they be seen, and remain to be seen, as only tools in the growing repertoire of the evaluator. Photography can aid the evaluator in focusing inquiry and collecting data. But the ends of evaluation, determination of merit or worth, the making of rational decisions, and the describing of events and activities must not be preempted by the tools which serve those ends.

REFERENCES

Kelly, E.F., Wachtman, E.L., Reeves, T., Cashell, J., McKee, B., and Hutcheson, S. A Story of ESSENCE: The Portrayal of a Science Curriculum Center for Instructional Development, Syracuse University, Syracuse, New York, May 1977.

Kemmis, Stephen. Telling It Like It Is: The Problem of Making a Portrayal of an Educational Program. University of Illinois at Champaign-Urbana, 1974 (Mimeograph) p. 15.

Sontag, Susan. On Photography. New York: Farrar, Straus and Giroux, 1977. p. 185, p. 22-23, p.23.

Stake, R.E. An Approach to the Evaluation of Instructional Programs. Paper presented at the annual meeting of the American Educational Research Association (Chicago, Illinois, April 1972), p.1.

Stake, R.E. Some Alternative Presumptions (editorial). In CEDR Quarterly, Center on Evaluation, Development and Research at Phi Delta Kappa International Headquarters. Vol. 10, No.3, Fall, 1977.

FACILITATIVE EVALUATION

Pamela A. Miller & Barbara B. McDonald

**Vermont State Department
of Education,
Montpelier, Vermont**

**Paper Presented at
Annual Meeting of American
Educational Research Society
Toronto, Ontario**

March 29, 1978

(Printed in U.S.A.)

59

Facilitative Evaluation

Pamela A. Miller & Barbara B. McDonald
Vermont Teacher Corps, Act II
University of Vermont

This article will describe an evaluative process derived from the authors' experiences in meeting the mandated evaluation/documentation requirements of a federal project designed to foster educational change and improvement. This article conceptualizes evaluation as obtaining information utilizable at two levels. Evaluation is seen as a means for empowering project staff to be able to monitor progress towards meeting project goals, yielding information useful in the process of designing and redesigning techniques for advancing to these goals. In addition, evaluation is seen as yielding enabling information for the drawing of generalizable conclusions about the change process.

The federal mandate for documentation led to a number of issues that needed resolution. In order to acquire wide-ranging information about issues bearing on project outcomes, such as differing perceptions of the political forcefield, strategies used by project staff to cope with the forcefield, and the interim and ultimate goals of the project staff, the active cooperation of all staff members in the keeping of records was necessary. This meant that staff members needed expertise in record keeping and motivation to keep these records. Realistically, many people are not trained to keep accurate records and need support in developing their own evaluative skills. The problem of motivating harrassed staff to keep records is even more difficult. It is partially solved by concerned support from the evaluator and by making information kept for documentation purposes a vital component of individual and overall project management.

Evaluation is seen as a reciprocal process between the evaluator and the people he deals with. The model assumes that staff must be empowered with a proactive role in which they have an impact both on the instruments that they use to collect data and the techniques used to analyze that data. The evaluator is obligated to portray to the people he collects information from the emergent themes he finds in the data. They, in turn, have the right both for instruction in interpreting these themes and an opportunity for affirmation or denial of them based on their personal perceptual framework of the validity of the evaluator's findings.

Beyond the issues of accurate, comprehensive documentation lies the problem of theoretical context, an issue not customarily dealt with in evaluation. The authors feel that the generalizable information available from a multi-faceted, multi-data source, case study has been underrated. There are some theoretical models which provide insights into the delineation of the change process. The current federal mandate for documentation/evaluation can provide information of a scope that can test a theoretical, explanatory framework so that generalizable information about the change process can be obtained.

This paper will describe four aspects of the facilitative evaluation process. The first section will discuss how the evaluator goes about producing the evaluation design while establishing a good working relationship with practitioners. Next, a description of the steps taken by the evaluator and practitioner for implementing the evaluation design are described. Then, some examples of theoretical contexts in which to consider the change process will be discussed. Finally, some suggestions will be sketched as to techniques for analyzing the data.

Designing the Evaluation

The following seven events describe how the evaluator initiates and continues his work:

- Event 1: Interview between Evaluator and Practitioner to:
- (a) establish trust level
 - (b) discuss practitioner's role
 - (c) assess level of evaluative sophistication of practitioner
 - (d) determine practitioner's perceived needs for information
 - (e) determine practitioner's ideas for collecting information
- Event 2: Compare results of total interviews with the overall project design, analyzing discrepancies and exploring causality.
- Event 3: Evaluator derives overall project design, given considerations of:
- (a) project goals and objectives
 - (b) practitioner's roles
 - (c) practitioner's ideas for collecting information and
 - (d) level of evaluative sophistication of practitioners
- Event 4: Handcraft an individualized instrument(s) for each practitioner based on overall evaluation design and practitioner's level of sophistication.
- Event 5: Check instrument(s) with practitioner to determine if it fits his needs and is usable. Instruct the practitioner in utilizing the instrument. Redesign instrument(s) if necessary.
- Event 6: Implement work process using individualized instruments to collect needed information.
- Event 7: Conduct follow-up interviews to:
- (a) monitor progress
 - (b) update instruments
 - (c) describe emergent themes and check with practitioner's perceptions.

The purpose for using the seven event model is to make it possible to collect data that is "accurate information" which will be utilized cooperatively by staff and evaluator in planning project development.

The Facilitative Evaluation model is intended to encourage growth both in the competencies of individual staff members and in the achievements of the project as a whole. In this model, the evaluator's role involves gaining the trust of the staff, involving staff in the design of instruments to obtain information and keep records, educating staff in evaluation techniques as they participate in the evaluation/planning process, and participating with staff members in utilizing instruments. These events empower staff with evaluative techniques so that they can use data for their own needs as well as the general project needs.

This model uses an interview technique, Event One, to gather information and to facilitate staff cooperation. Primarily, one-to-one interviews with staff and other appropriate people are scheduled. The interview process, when done well, closely parallels a helping-counseling process. There are many factors that contribute to successful use of this technique. The evaluator must assume that the person being interviewed is doing important work and is worthy of consideration as a professional. Both parties should be seen as equals, discussing a situation of mutual interest and concern. This attitude will convey itself in attentive listening, positive non-verbal encouragement and appropriate responses, with the evaluator putting himself into a relaxed state with focused attention upon the person. A casual environment such as a informal lunch or discussion over a cup of coffee provides the low-key unaggressive, non-judgmental atmosphere that is most productive. If the process is going well, a synergistic phenomenon should take place with two minds becoming more

productive and energizing than one. Work is important and people are usually eager to discuss and reflect on it. This synergism is so rewarding and productive that experience has shown that staff members often strive to duplicate these conditions with other colleagues, increasing the creative power of the total project.

It is important in the interview process to determine the practitioner's current capability to evaluate information and his ideas on how to collect it. Of necessity, when documentation is important, staff must keep records. However, it is futile to expect them to keep records they are unable or unwilling to maintain.

Event Two describes the process of comparing what people actually do based on interviews and observations with the overall written project design. It is well known that roles specified in proposals or planning documents often radically change both because of the people filling them and because of the evolving needs of a project. The evaluator needs to make sure that staff realize that this is a common occurrence that does not imply blame. What is important is that some sort of record be kept concerning the context in which the change happened so that knowledge can be derived from this phenomenon. Event Three, the creation of the overall evaluation design, is self-explanatory, albeit the difficult crux of the matter. Event Four describes the selection or building of individualized instruments and documentation systems based on a realistic assessment of the practitioner's interests and capabilities in conjunction with the overall evaluation design.

Event Five describes a critical step in the evaluation/documentation process. It is important to check back with the practitioner to see if he can utilize the instruments designed for his use. The practitioner's capabilities as well as pragmatic considerations of available time and resources may have been over or underestimated. It is often necessary to work through an instrument with the practitioner using a real life situation as an example so that the practitioner may get realistic practice in using it. Finally, instruments may have to be redesigned. It is better to redesign an instrument so the practitioner will record data than to lose the information by using an instrument that is too complex to be accurately filled out by the evaluator.

Event Six sees the practitioner using the newly designed instruments to collect information. His role in this capacity is that of a participant observer. Event Seven describes the conduction of follow-up interviews to monitor progress and update the individually designed instruments. It is important that the evaluator show sustained interest in the data the practitioner gathers. It is not always convenient or agreeable to record information. The evaluator's personal concern with the data and the practitioner as well as the practical utilization of the data in the project will help to alleviate this natural phenomenon. In addition, the evaluator needs to monitor changing project goals and practitioner roles and update his instruments accordingly. Then, too, it is hoped that the practitioner will become more sophisticated in utilizing evaluation as he gains practice in the evaluation process and adjustments in instrumentation should be made accordingly. These encounters also offer

7

the evaluator the opportunity to portray to the practitioner the insights and repetitive themes that he has discovered from his overall perusal of the evaluative data and to discover whether or not these themes are validated by the perceptions of the practitioner.

In summary, these seven events describe what the evaluator does to procure practitioner participation in the collection of information. The events are conceptualized as a means to train practitioners so that they can serve as participant-observers in the field. They will then have the capacity to validate the emergent themes presented by the evaluator and to derive their own emergent themes for presentation to the evaluator. From overall project goals and, where available, emergent themes, the practitioner develops work goals. The planning-replanning cycle to implement these work goals will be described in the next section.

Implementation and Utilization of Data by Practitioner

In this section, facilitative evaluation will be discussed primarily in terms of steps that practitioners take to bring project goals into reality. These steps are designed to roughly replicate, in a pragmatic manner, the steps of goal setting, implementation and reassessment that are so useful in an innovative project.

The steps in implementation and utilization of data by practitioners are:

- I. Select a work goal
The practitioner defines clearly a goal and discusses the setting where it will occur. The evaluator should offer clarification and help in this process.
- II. Discuss the present reality and identify the problem
The practitioner takes an objective look at the situation. The evaluator should encourage this process by responsive listening and questioning.

- III. Conceptualize the most effective solution
The practitioner conceptualizes the most effective solution(s) to the problem. The evaluator can be especially helpful by offering encouragement conducive to inventive, yet feasible conceptualizing of the future.
- IV. Plot steps to the solution
The practitioner plots the steps needed to bring the goal from where it is presently to where the practitioner would like it to be in the allotted time.
- V. Establish checkpoints
The practitioner establishes completion times of steps as tentative checkpoints. These checkpoints serve to motivate the practitioner to move forward.
- VI. Design evaluative indicators for each work step
These indicators are milestones which indicate readiness to move to the next step. (Or a failure which indicates need for re-appraisal of (a) the steps to achieve the goal or (b) the appropriateness of the goal).
- VII. Implement initial steps of the plan
As each step is implemented and the evaluative indicator has or hasn't become reality, the practitioner should log warning signs, that indicate the steps to compile the goal may need revising due to new information. Conversely, data may indicate success.
- VIII. Assess progress at checkpoints
This is done to determine whether the next steps are feasible considering the cues that the practitioner has collected.
- IX. Continue progress
Continue to the next step or replot steps. If useful, devise a new, more effective solution to better accommodate the new reality.
- X. Complete all steps or discontinue goal
Data indicates whether practitioner should continue on present course or discontinue the work goal.

In actuality, steps one to three, the selection of a work goal, delineation of present reality and conceptualization of an effective solution to a problem, are not necessarily done in linear sequence. Instead, suc-

cessive approximations of one step reverberate on the findings of another step which in turn initiate new thinking of an earlier step.

The selection of a work goal, step one, is dependent on a number of factors. Among them are overall project goals, identified emergent themes, staff interests and competencies and the time span of the project with the ramifications for what is feasible. Step two, an assessment of the present reality has a heavy impact on the selection of a work goal. The factors to be considered in such an assessment are numerous. Critical among these factors is an unblinking, realistic assessment of political realities both in the actual system being worked on and the center environment. Aside from the political realities, the socio-cultural-physical climate of a system needs careful consideration. In the end, the number of factors considered in an assessment of current reality is determined by the perspicacity of the staff as well as the reflective probing of the evaluator.

Step three involves the conceptualization of the most effective solution(s) to the problem. Given the context of an overall work goal and the present reality as perceived by the people in the environment, practitioners decide what the most effective solution(s) to the problem are. Practitioners, aided by informed questioning by the evaluator, need to imagine what a better reality would look like. A discussion of the context within which the problem will occur and feasible alternative solutions will help the practitioner think through the direction to take. If the problem is particularly complex, a brainstorming of helping and hindering factors may be of assistance. During the process of the practitioner working through these three steps, the evaluator can also be sorting out the project's perceived underlying problems, its underlying philosophical assumptions,

checking their veracity against the perceptions of practitioners. . . These sessions offer the evaluator the chance for reflective portrayal to practitioners of emergent themes that he has seen from his work.

Once a consensus on the future objectives is reached, the actions necessary to bridge the gap between present reality and a future objective need to be derived. This is a careful meticulous process which relies heavily on step two, i.e., a forthright unblinking assessment of present reality, both political, and social. Step four requires the practitioner to estimate a realistic time line for the completion of each action of the work goal. Having established the actions for the work goal and affixing each action with a completion date, step five continues the process by establishing indicators to help the practitioner know that he has achieved or failed to achieve the necessary actions.

The sixth step is to implement the plan, recording whether or not the evaluative indicators have been met. Different indicators signifying that an action is completed may indicate a need to readjust subsequent actions. Steps seven, eight and nine are to assess progress at checkpoints, continue to next step or replot steps and finally, complete all steps or discontinue goals as data indicates. The steps are conceptualized as a circular process of goal setting, implementation, checking and re-appraisal. The choice of abandonment of an untenable position is built into the process. It is assumed that a position may become untenable because of an inability to achieve the intermediary steps necessary for attainment of the goal or because the goal itself begins to be seen as a mistake. It is important to document the events leading up to failures so that appropriate analysis may be made of the causes of such failures. This whole series of steps is designed to yield written records that are of systematic use in decision making

by practitioners as well as providing participant observation field records for the evaluator.

Theoretical Framework

As the authors stated in the introduction, this evaluation model serves two purposes. The first purpose, discussed in the previous sections, is to provide data for better project management. The second purpose is to gather generalizable data about the change process. In this section a discussion of the theoretical frameworks dealing with the change process will reveal possible themes that may assist the evaluator in a better understanding of the proceedings that are being assessed.

"Industrial managers are fond of noting that change is the only thing that remains constant in their lives. Yet despite the common occurrence of organizational change, its dynamics and underlying processes are understood in only rough, ill-defined ways. Managers and social scientists who create and study change situations find that organizational changes involve multiple sets of complex variables whose identify, interaction and impact vary from situation to situation."
(Barnes, 1967.)

Because of the varied and rich data at his disposal, the project evaluator is in an excellent position to identify emerging themes and to portray these themes in a larger theoretical framework. The authors intend to briefly describe three sociological theories that provide conceptual contexts by which to consider the change process. The three theories are: conflict theory, exchange theory and interaction theory. Although there are certainly other theories that would be appropriate to

consider in gaining a larger theoretical framework in which to consider the process of change, it is not the authors' intention to present an exhaustive study of theoretical models related to change but merely to spark interest in the notion of incorporating theoretical thinking into the evaluation process.

Conflict is a theme that recurs frequently in the change process. Coser (1956), in his study of conflict, cites a number of scholars such as Simmel, Sorel and Marx who maintain that conflict is a binding as well as a disruptive one. He goes on to state that interest groups will form and reform around issues that are conflict laden. "...when a social structure is no longer considered legitimate, individuals with similar objective positions will come, through conflict, to constitute themselves into self-conscious groups with common interests." (Coser, 1956.). Most theorists in talking about conflict tend to agree that conflict cannot be repressed. How it is expressed, however, is reflected in the characteristics of the individuals or groups dealing with it. Coser states that there is a difference in group behavior when conflict exists within a group and when it is directed at a group from an external source. Groups who have very close internal relationships tend to repress internal conflict longer, making the conflict much more intense when it finally occurs. Groups that are being attacked from without tend to demand more commitment from their members with fewer conflict issues allowed to surface from within.

A final element in conflict theory that the authors will consider is Coser's definition of realistic and non-realistic conflict. He states that

while realistic conflict involves resolving conflict focusing on the actual issues, non-realistic conflicts are based on real issues, but deal with them in an indirect manner, resulting in a conflict resolution dealing with false issues. Therefore, there is not a true resolution of the actual issues. Noting that conflict theory concerns itself with the behaviors of individuals and groups engaged in conflict situations, the authors feel that a look at exchange theory would expand the reader's theoretical perspective of human behavior around issues of change by examining and analyzing other factors that affect what people do. In this theory, Thibault & Kelley (1967) are concerned with interpersonal relations and group functioning or put another way, how people's behaviors are influenced by the costs or rewards that they perceive they will experience as a result of their actions. During the process of deciding what action to take, a person is continually weighing the consequences of his behavior, often with opposing notions about what his actions should be. Thibault & Kelley (1967) define a variety of explanatory terms that help decide what action a person will take. Norms, or those agreements about what behaviors are and aren't acceptable that carry with them a social process to enforce compliance, afford a certain degree of predictability to the actions people may take. Status is also important when deciding who will do what. The positive status of a person influences others to more frequently behave as he does, giving him more power. Power and dependence are interrelated phenomena. An individual obtains power by others being dependent on him. His power is greatest when he can operate as he wishes without being concerned with others actions interfering.

Exchange theory takes into account all the possible behaviors a person may exhibit in any one instance and analyzes the consequences each of the behaviors will have for him. While exchange theory deals largely with people's behaviors, interaction theory deals with the perceptual events in people's minds. As change in an organization involves constant human interaction, it is useful to look at some conceptualizations of how and why humans interact with each other as they do. Interaction theory looks at human conduct and how people function with other individuals or groups.

George Herbert Mead and Herbert Blumer (1969) identify two kinds of interaction in human society. Blumer calls them "non-symbolic" and "symbolic" interaction. Non-symbolic interaction takes place when a person responds indirectly to the actions of another without interpreting them. Symbolic interaction assumes that a person takes meaning from others' actions and responds by behaving in accordance to the meaning he has affixed to the others' actions. (Blumer, 1969). Blumer also uses the term "joint action." This is characterized by actions ranging from the interaction between two people to the complex social workings of an organization. These "joint actions" make up society, according to interactionist thought. "Each participant necessarily occupies a different position, acts from that position, and engages in a separate and distinctive act. It is the fitting together of these acts that constitutes joint action." (Blumer, 1969). People perform tasks by assessing the job to be done, looking at their own actions and the actions of others and fitting these actions together to accomplish the task. Interactionists believe that people's actions are determined by the way that they perceive reality. A person's perception of reality and the role that

he plays in that picture of reality, therefore, effect the alliances he develops and the actions that he takes to maintain or change these alliances.

In summary, the authors have briefly described parts of three sociological theories that have implications for the change process. Our description of conflict theory looked at how conflict affects people in their dealings with others. The distinction between realistic and non-realistic conflict was also described.

Exchange theory deals with people's behavior and how these behaviors are influenced by the costs and rewards that people perceive they will experience as a result of their actions. Norms, status, power and dependence were all discussed in relation to this theory.

Finally interaction theory was discussed. Interaction theory states that people's actions are determined by their perception of reality and the role that they play in that reality. Symbolic and non-symbolic interaction and joint action were defined.

Each theory gives the evaluator a different but not unrelated perspective of the change process. A conceptualization of the evaluator's role in portraying this knowledge for the practitioner might be seen as Miller et al. have portrayed the change process in their theory of social reform (1977). First, an identification of the actors in the process is made and then an identification of the variables which describe the process in which the actors are involved is determined. An analysis of the interrelationships that exist between the actors and the process that they are involved in characterize the change process. With this conceptual model as a base, the usefulness of a theoretical framework becomes evident.

Conclusion

Evaluation has been customarily conceived of as formative and summative in terms of success and failure. Given the time and money limitations of much evaluation, this is indeed, all that can be hoped for in many cases. However, given the scope of information that can be collected using techniques of evaluation/documentation, the authors would like to briefly sketch some techniques of synthesizing this information into a comprehensive whole.

As the evaluator follows the techniques outlined in previous sections, he is faced with reams of data collected both from his vantage point and from the vantage point of the participant-observers he has trained. He is confronted with the task of placing differing world views into a portrayal of the whole. The authors concur with the techniques suggested by Carini, Engel & Hein (1978) for the application of qualitative methods to program evaluation. As information is collected, the evaluator should immerse himself in the contemplation of it. Themes should eventually emerge from the contemplation of this collection of data. The evaluator should then begin combining this data into a loose weave of meaning. Once this is done, he can commence cross-matrixing these emergent themes preliminary to an analysis of the data to determine the relative influence of various actors and their actions on the process. The data can then be cross-checked with theoretical premises. The evaluator, because he has access to all the rich array of information is in an excellent position to portray data in a broader, theoretical framework.

The authors feel that the interaction of practitioners with the evaluator is crucial throughout this process. People's perceptions of

their experiences trigger theoretical understandings which when revealed to the practitioner empower him with new insight. Besides utilizing the perceptions of a number of participant-observers with differing world views in his immersion in the data, the evaluator is obligated to continuously check and cross-check with participants both the emergent themes that he perceives as well as the ones that they perceive. The evaluator needs to make fragmented data whole and mirror it back to practitioners making it a comprehensible entity for them. Emerging themes are portrayed to managers so that they can change their procedures, if necessary. Data is thought of as representing a phenomenologically rich world in the process of becoming.

The authors feel that, while human beings are amazingly complex organisms, there is still much that is universal from human situation to human situation. They suggest consideration of this model to gain information that will be useful in the "practical art of getting things done" (Stake, 1978, paraphrase) while allowing for the chance of relating case study data to a larger theoretical framework.

SELECTED BIBLIOGRAPHY

- Barnes, Lewis B. "Organizational Change and Field Experiment Methods." Victor H. Vroom. ed. Methods of Organizational Research. Pittsburgh, Pa.: University of Pittsburgh, 1967.
- Bloom, Benjamin S., Hastings, J. Thomas, and Madaus, George F. Handbook on Formative and Summative Evaluation of Student Learning. New York: McGraw-Hill, 1971.
- Blumer, Herbert. Symbolic Interactionism. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1969.
- Campbell, Donald, "Reforms as Experiments," American Psychologist, 1969, XXIV, p. 409-429.
- Carini, Patricia, Engel, Brenda and Hein, George, "Observing and the Art of Seeing, Children's Work as 'Data', Program Documentation: the Emergent Curriculum Program Evaluation," Workshop, Lesley College, Cambridge, Mass.: March 2, 3, 4, 1978.
- Chronbach, L. "Beyond the Two Disciplines of Scientific Psychology," American Psychologist, XXX (2), 1975.
- Cohen, David K. "Politics and Research--Evaluation of Large-scale Programs," Review of Educational Research, 1970, XL (2), p. 213-238.
- Coser, Lewis. The Functions of Social Conflict. New York: The Free Press, 1956.
- Hemphill, John K. "The Relationships between Research and Evaluation Studies," In Ralph W. Tyler, Ed., Educational Evaluation: New Roles, New Means. Chicago: University of Chicago Press, 1969, 68th Yearbook Part II, National Society for the Study of Education, p. 189-220.
- Kaplan, A. The Conduct of Inquiry: Methodology for Behavioral Science. New York: Chandler Publishing Co., 1964.
- Miller, Alden; Ohlin, Lloyd; and Coates, Robert. A Theory of Social Change: Correctional Change Processes in Two States. Cambridge, Mass.: Ballinger Publishing Company, 1977.
- Stake, Robert, "The Case Study Method in Social Inquiry," Educational Researcher, 1978, p. 5-8.
- Stake, Robert, "The Countenance of Educational Evaluation," Teachers College Record (68) (1967): 523-40.
- Thibault, John and Kelley, Harold. Social Psychology of Small Groups. New York: Wiley, 1961.

HOW TO

Affiliated with Education Daily

REPRINTED BY PERMISSION

Emily C. Harris

EVALUATE EDUCATION PROGRAMS

A MONTHLY GUIDE TO METHODS AND IDEAS THAT WORK

Vol. 2, No. 1

January 1978

Highlights In This Issue

- What to consider when you choose an information collection method (see below)
- Should you use ability tests? (see p. 2) How about questionnaires? (see p. 2) Are interviews feasible? (p. 4)
- A quick guide to pros and cons of each method (p. 5)
- News about evaluation activities (see p. 8)

How To Collect Evaluation Information

You can collect evaluation information in many different ways. If you need information to find out whether an innovative elementary school reading center has actually improved student reading ability, for example, you would probably ask this question: How much have students' reading skills improved?

To provide a credible answer to the question, you could use any of the following information collection techniques:

- Give parents a rating scale to assess their children's reading performance.
- Send questionnaires to teachers to get their opinion about student reading performance.
- Interview students to ask their opinion about student reading performance.
- Observe students as they read and rate their reading ability.
- Have students keep a diary of their progress.
- Give students a nationally normed achievement test that assesses reading performance.
- Have students take a teacher-developed test.
- Review student records for achievement test scores, report card grades, and teachers' comments.

This list illustrates six alternative techniques an evaluator might use to answer the question about improved reading skills. They are interviews, questionnaires, rating scales, standard observations, record reviews, and achievement tests.

To choose the best technique for answering a particular question, you should consider four factors. First, the method should be agreeable to your client and colleagues. If you want to use questionnaires, but the district's staff prefers interviews, you must decide just how serious the consequences of imposing your own choice might be.

(more)

How To Collect Evaluation Information (Cont.)

Second, the information collection techniques should be technically sound and the data collected from them should be reliable, valid, and targeted to the evaluation questions. Third, the information collection techniques should provide the best data the evaluation budget can afford, which means that you will have to decide such things as whether to buy or develop your own measures, and whether to use more than one technique for each evaluation question. Fourth, you must be sure that the methods you choose will allow enough time for gathering and analyzing the data.

Should You Use Paper and Pencil Tests Of Ability?

Paper and pencil tests are among the most commonly used measurement techniques in education.

Achievement tests measure competency in a given subject. They can be developed by the program or evaluation staff or you can buy them from publishers. Achievement tests can be used to measure a student's knowledge of basic English usage or a class's ability to solve quadratic equations.

The advantages of achievement tests are that they can be administered to large groups at relatively low cost and that carefully developed and validated tests are available in many subject areas.

One disadvantage is that achievement tests must be properly validated to provide accurate information, and this can be a costly procedure. Another is that having high scores on a test of factual knowledge doesn't always mean that the student can apply that knowledge.

Aptitude tests are measures of potential. The most common measure of aptitude is an IQ test. Aptitude tests have the same advantages and disadvantages as achievement tests.

What About Paper and Pencil Self-Report Measures?

Paper and pencil self-report measures ask people to tell about their attitudes, beliefs, feelings, and perceptions. Questionnaires, rating and ranking scales, the Semantic Differential, the Q-sort, and diaries are among the techniques most frequently used in evaluating education programs.

Questionnaires are self-administered survey forms that consist of a set of questions. Answers to questionnaire items can require free responses (short answers) or they can be structured into "forced" choices (multiple choice items). Questionnaires are frequently used in large scale evaluations to obtain participants' reactions and opinions. They are less expensive to construct than most measures, but the kind of information you can get from them is limited, and people don't always answer the questions truthfully. Don't forget that you will have to follow up on those who don't respond, and that's an expensive undertaking.

Rating scales can be used for self assessment or for appraisals of other people, groups, events, or products. Student attentiveness, for example, can be rated on a five-point scale from 1 for not very attentive to 5 for very attentive. Rating scales are particularly useful when you need to reduce judgmental data

What About Paper and Pencil Self-Report Measures? (Cont.)

to a manageable form. They are relatively easy to complete and they produce objectified data.

Unfortunately, they are subject to many types of bias--some raters are lenient and others are not, and sometimes raters let their personal feelings influence their ratings (a halo effect). Further, the amount of information you can obtain from rating scales is limited because the rating categories are never perfect.

Ranking scales involve putting a set of items into a hierarchy according to some value or preference. Asking a teacher to rank four textbooks from one to four according to their reading difficulty is an example. Like rating scales, ranking scales are easy to complete and produce objectified data. But remember that rank ordering a long list of items is no fun and it takes a lot of time. Ranking scales sometimes ask people to make distinctions among things where they can't really see any difference.

Semantic Differential is used to measure attitudes by relying on the indirect meanings of words.

For example, students might be asked to rate their country using a series of seven point scales like the following:

<u>UNITED STATES</u>						
GOOD	—	—	—	—	—	BAD
PASSIVE *	—	—	—	—	—	ACTIVE
SMALL	—	—	—	—	—	LARGE
DEMOCRATIC	—	—	—	—	—	UNDEMOCRATIC
RICH	—	—	—	—	—	POOR

The Semantic Differential is relatively easy to complete, it produces objectified data, and respondents usually find it harder to choose "socially acceptable" answers than when they use an ordinary rating scale. However, the Differential can be difficult to score.

The Q-sort requires individuals to place a series of items or statements into rating categories so that some minimum number of items is assigned to each category. For example, teachers could be asked to rate ten textbooks as "above average" or "below average" so that at least two texts are assigned to each category (the remaining texts can be rated either way).

Q-sorts produce objectified data and they force respondents to establish priorities among items that are being compared in an evaluation. But the Q-sort requires people to make very difficult distinctions, the directions are often hard to follow, and the resulting data can require complex analysis methods.

Diary techniques ask people to keep daily or weekly accounts of specific behaviors, attitudes, thoughts, or events. The critical incident technique asks people to record only those things that are particularly important, unique, useful, or revelatory. For example, you could ask students to keep a daily diary of the amount of time spent in free reading, or to record the names of any books

What About Paper and Pencil Self-Report Measures? (Cont.)

they liked or disliked.

Diaries and critical incidents permit people to describe unique situations in their own words, but people sometimes forget to maintain them and they are often difficult to score and interpret.

Why Not Try Observations?

Another information collection technique frequently used in evaluations is an eye-witness account of individual behavior or program activities. You could use observations, for example, to find out which visual aids teachers are using. The information collected by observers can be reported by checklists, rating scales, field notes, and summary reports.

Standard observations require careful planning so that the information obtained is accurate. Observations can give information collectors first-hand information about a program, and they are often the only feasible and economical way to gather certain kinds of information. But it is costly to train observers, and several may be needed to get reliable results. Another drawback is that people who know they are being observed may not behave normally.

Time sampling observations involve repeated observations of a given situation. For example, observers may note how many students and which ones ask direct and indirect questions during ten consecutive five-minute intervals.

Time sampling allows first hand observations of a program, and the many observations make it possible to identify unusual events that you might otherwise think were routine occurrences. When all the observations are made one after the other, however, they are likely to depict only one particular situation and not the program as a whole.

Are Interviews Feasible?

An interview is an information collection technique in which one person talks to another or to a group. Interviews can be completely unstructured and spontaneous, or you can decide ahead of time the kinds of questions to ask. If you use multiple choice questions, even the response categories are predetermined.

Face-to-face interviews might be used, for example, to find out why participants dropped out of a program, and might consist of three basic questions with a series of two or more in-depth questions for each basic question asked. The best thing about the face-to-face interview is that it permits you to probe sensitive subjects like attitudes or values. But these interviews are usually time-consuming and expensive, and you will have to give interviewers special training.

Telephone interviews also permit in-depth probing of sensitive issues and are less costly than face-to-face interviews. They are still expensive when compared to questionnaires, however. You should also remember that not everyone has a telephone, and some people are reluctant to reveal their feelings or give personal information over the phone.

Are Performance Tests the Answer?

Performance tests require people to complete a task or make something, and then you assess the quality of the performance or product. One example of a performance test is when you have someone type a letter and then you count the number of words typed correctly in a set amount of time. Another example is when a group of experts observe a teacher instructing a class and then use a rating scale to indicate their appraisal of the teacher's ability.

The major advantage of performance testing is that it relies on tasks that are close to "real world" activities. It is often very time consuming and expensive, however, because performance tests generally have to be administered individually and they sometimes require the use of special equipment.

Would Record Reviews Be Enough?

Record reviews mean that you collect evaluation information by going through program related documents. In a program where older pupils tutor younger ones, for example, you might review attendance records to see if either the tutors or the children they taught came to school more regularly after the program began.

Record reviews are "unobtrusive" in the sense that they do not interfere with the activities of the program being evaluated. They can also be relatively inexpensive because no new data collection is required. One problem is that program documents may be disorganized or unavailable.

INFORMATION COLLECTION ALTERNATIVES		ADVANTAGES	DISADVANTAGES
Paper and Pencil Tests of Ability	Achievement Tests Aptitude Tests	Can be administered to large groups at relatively low costs Many published, standardized tests are available	Expensive to develop and validate High scores do not necessarily imply that the tested knowledge can be applied
Paper and Pencil Self-Report Measures	Questionnaires	Can be administered to large groups at relatively low costs	Can be difficult to obtain sensitive information Respondents may not always be truthful Must follow-up to obtain adequate numbers of respondents

INFORMATION COLLECTION ALTERNATIVES		ADVANTAGES	DISADVANTAGES
Paper and Pencil Self-Report Measures (cont.)	Rating Scales	<p>Easy to complete</p> <p>Produces objectified data</p> <p>Reduces judgmental data into a manageable form</p>	<p>Responses may be biased because some raters are lenient and others are strict</p> <p>Amount of information obtainable is circumscribed by the rating categories</p> <p>Halo effect</p>
	Ranking Scales	<p>Easy to complete</p> <p>Produces objectified data</p>	<p>Difficult to rank a long list of items</p> <p>Distinctions are called for that are not perceived</p>
	Semantic Differentials	<p>Easy to complete</p> <p>Produces objectified data</p> <p>More difficult to give "socially acceptable" responses</p>	<p>Difficult to score</p>
	Q-Sorts	<p>Produces objectified data</p> <p>Forces respondents to establish priorities among items</p>	<p>Distinctions are called for that are not perceived</p> <p>Directions can be too elaborate.</p> <p>Can require complex data analysis methods</p>
	Diaries and Critical Incidents	<p>Permits people to describe unique situations in their own words</p>	<p>People don't maintain them</p> <p>Difficult to score and interpret</p>

(Cont.)

INFORMATION COLLECTION ALTERNATIVES		ADVANTAGES	DISADVANTAGES
Observations	Standard Observations	Can observe events at first hand	Observers can change the environment Inter- and intra-observer reliability can be difficult to obtain
	Time Sampling Observations	Can observe events at first hand More opportunities to observe	Observers can change the environment Inter- and intra-observer reliability can be difficult to obtain
Interviews	Face-to-face Interviews	Permits indepth probing Sensitive issues can be discussed	Costly Inter- and intra-rater reliability can be difficult to obtain
	Telephone Interviews	Permits indepth probing Sensitive issues can be discussed Less costly than face-to-face interviews	Costly Some people may not have telephones More difficult to probe or discuss sensitive issues
Performance Tests		Close to real world situations	Costly Generally must be administered individually Can require special equipment or apparatus
Record Review		Unobtrusive Can be inexpensive No new data collection required	Documents may be disorganized or unavailable 152

News Briefs

The essentials of an adequate evaluation are spelled out in no uncertain terms by G. Kasten Tallmadge of RMC Research Corporation in a new "Ideabook" for the National Institute of Education. Did a change occur? Was the effect consistent enough to be statistically significant? Was the effect educationally significant? Will it work just as well in other places? Are you sure it was the program that made the difference? Is the evidence understandable and believable? To top it off, you'll find a lucid chapter on common evaluation hazards and how to avoid them.

A limited number of copies of Ideabook are available from the National Institute of Education, Education and Work Group, Washington, D.C. 20208.

Nearly 200 programs that measured up to Tallmadge's stringent criteria are described in a 200-page paperback called "Educational Programs That Work." Exemplary projects that have been approved by the Office of Education/National Institute of Education review panel include early childhood education, reading, math, special ed, career ed, environmental ed, alternative schools, and more.

Order by sending prepayment of \$4.95 per copy to Far West Laboratory for Educational Research and Development, 1855 Folsom St., San Francisco, CA 94103, attention Order Department.

The Evaluation Network offers up-to-the-minute information on resources, training programs and materials, new developments in evaluation technology and theory, and practical applications. Membership includes subscription to Evaluation News, a quarterly publication edited by Michael Scriven.

For further information write EN c/o Phi Delta Kappa, Eighth St. and Union Ave., PO Box 789, Bloomington, Indiana 47401. Annual dues: \$5.50.

The Evaluation Improvement Program is a nontechnical inservice effort to train teachers, principals and agency managers in the techniques of education program evaluation. It offers self-instructional material, including a Program Evaluator's Guide, an evaluator's workbook and a trainer's guide, and a catalog of evaluation training materials from the National Institute of Education.

Contact W.W. Walton, Evaluation Improvement Program, PO Box 2845, Educational Testing Service, Princeton, NJ 08540.

Dr. Arlene Fink and Dr. Jacqueline Kosecoff, editors of this newsletter, will conduct two workshops this spring on How To Evaluate Education Programs. The dates are April 24-25 in Washington and May 11-12 in San Francisco. Mark your calendar now and write or call John Ekberg, ADD, Capitol Publications, Inc., 2430 Pennsylvania Ave., NW, Washington, D.C. 20037. Phone (202) 452-1600.

Evaluating an AGE Program

Wayne D. Dvorak and Donald P. Lang

Any evaluation plan designed to evaluate arts in general education programs, if it is to be effective, must reflect a reconsideration of the whole process of education evaluation. AGE programs represent a critical philosophical shift regarding the place of the arts in general education and the values and special contributions the arts can make to a school program. The approach is significant because it is global—the arts are used as one means to teach everything. The evaluation plan, therefore, must be designed to reflect that universality—the philosophy, the aims, the outcomes, and the complexity of the program it purports to evaluate. It must be comprehensible and flexible, yet retain validity.

Education programs, particularly those that are characterized by new approaches, will ultimately be evaluated by everyone—project directors, participants, teachers, students, parents, and administrators. What is not needed is an evaluation plan that reduces complex and unique education programs to esoteric jargon and statistical “simplifiers” and descriptors like “x,” “y,” population scores, and standard deviations.

The evaluation plan proposed here results in an understanding of the program and its effect upon the participants. It can be and has been used effectively to evaluate either arts education or music education programs. Developed by Robert E. Stake, director of the Center for Instructional Research and Curriculum Evaluation at the University of Illinois, Urbana, the plan is called “Responsive Evaluation.” It establishes some important alternatives to the whole process of education evaluation research, particularly of arts in general education programs.

Responsive evaluation

The two major activities of a Responsive

Evaluation are the description and the judgment of the program under examination. The difference between “description” of the program and the use of “descriptors” is an important distinction. Descriptors are simplifiers. Descriptors only tell, for example, that certain “individuals are observed, found to differ, and the distribution of the scores is described. Covariations of various kinds are reported and interpreted.”¹ Description portrays. Description tells us what the program is really like. Responsive Evaluation portrays the program in all its scope and complexity. It is based upon “. . . what people do naturally to evaluate things. They observe and react.”² The approach is not new. What is new is the beginning of a technology developed around this natural behavior, in part to overcome its defects. The accompanying figure is a graphic representation of that technology.

Description

Initially, statements concerning the intended antecedents (from where the program is starting), intended transactions (what classroom procedures will be used), and the intended outcomes (who will benefit and how they will benefit) are gathered by the evaluator from program personnel. Next, he or she arranges for observations of the program by various people, including himself and individuals with no vested interest in the program. He gathers their perceptions and impressions of the observed antecedents, observed procedures, and observed outcomes, remembering to take into consideration the differing value perspectives of each observer.

The diversity of an AGE program demands redundancy in the evaluation process. Repeated encounters by the same observer, by many different observers of the same experience, and a widening perception by all can “discover” significant happenings, either favorable or

Dvorak is assistant professor of music education at Hartt College of Music, University of Hartford, West Hartford, Connecticut. Lang is assistant professor of music education, University of Delaware, Newark.

¹Robert E. Stake, “Responsive Evaluation” (Urbana, Illinois: Center for Instructional Research and Curriculum Evaluation, 1972), p. 4.

²Stake, p. 1.

unfavorable, that may be missed in a single encounter. What is of value within the program and what is not will emerge as a consensus among the observers.

At this point it is important to note the possibility of using appropriate tests as part of the data-gathering process for establishing observed outcomes. The choice of these instruments is made as a result of observing the program in action and of interacting with various groups that have an interest in the program.

Judgment

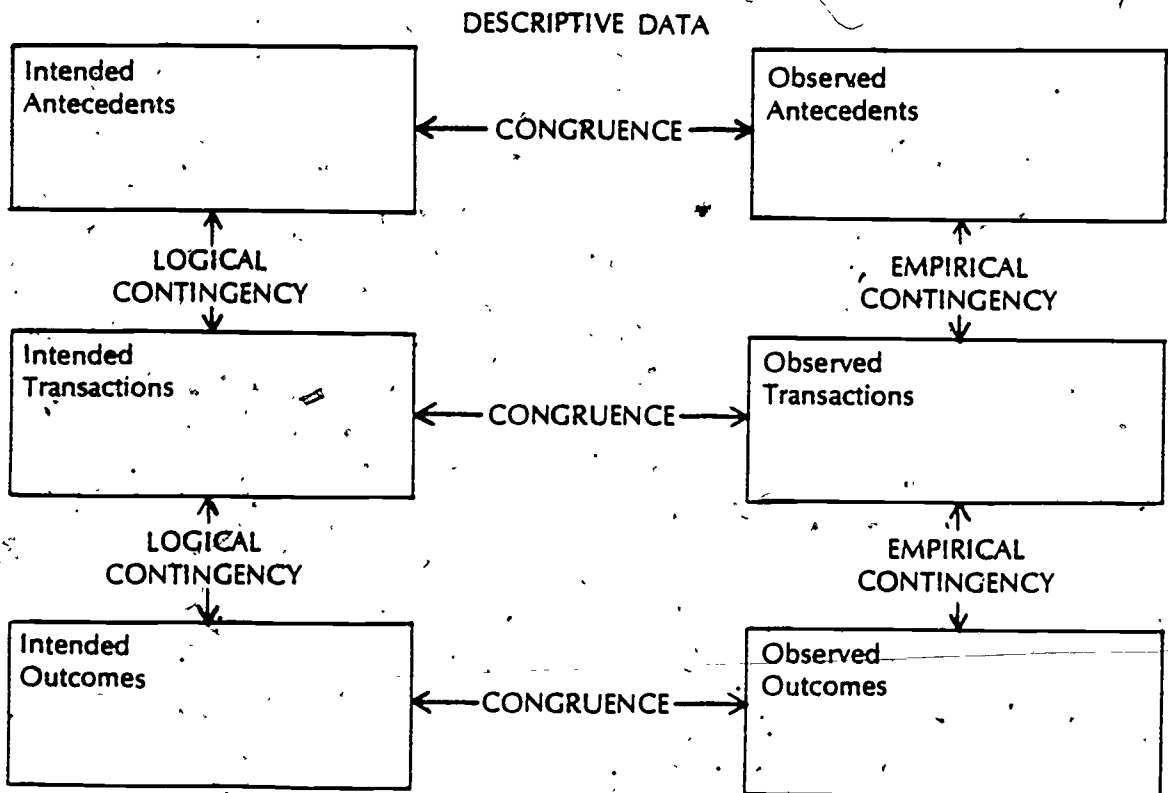
Education programs are rarely simple—most assuredly, arts programs are complex. Each observer and participant of the program brings to his record of the program his own biological inheritance, his past learnings, and his perception of the immediate situation. This becomes the "truth" of that situation for the observer. Since there exists no single, ultimate truth regarding the value or relevance of any specific education program (preordinate evaluators notwithstanding), the gathering of many different versions of the truth according to differing value-perspectives is more effective, revealing, and reliable than the pursuit of a single, ultimate value. Divergent and even contradictory subjective evaluative commentaries can often generate valuable information about how an AGE program really

works, as well as contribute directly and significantly to the evaluative phases of the research project.

If the description is accurate and complete, if the reader of the final report can get into the heart of the program, if the program personnel can be shown the realities of the program, rather than what they believe the program to be, then a basis for critical judgment and meaningful evaluation of the program has been provided that goes beyond the interpretation of test scores.

The total number of value perspectives gathered and reported is limited only by the resources, the time, the priorities, and the interests of the evaluator and his clients. Proper attention to diversity will allow for the greatest service to the greatest number of people. It can also help to counteract the reports of the observers or participants with "an axe to grind," an important consideration in any style of research reporting.

In the judgmental phase of the evaluation, the evaluator notes the discrepancies and congruencies between the intended antecedents, procedures, and outcomes. Either or both of the intended and actual antecedents, procedures, and outcomes may also be compared to an ideal or accepted standard. Similarly, the evaluator may note if the intended antecedents and procedures could be expected to produce the desired outcomes—logical con-



tingency. Also, the evaluator may note the cause-and-effect relationships among observed antecedents, procedures, and outcomes.

Any complex, diverse program that has had repeated viewing will yield ambiguous and contradictory information. Traditionally, this has been the bane of preordinate evaluators—the results do not fit neatly into the paradigm. The Responsive Evaluation approach overcomes this dilemma through the use of what Stake calls "adversarial procedures."

Most common in the work of legal theorists, adversarial procedures call for the reporting of evidence about a given program by two individuals—an advocate and an adversary. The advocate's statement is a summary of the most positive claims that might reasonably be made about the program, while the adversary summarizes the most damaging claims. Neither statement is intended to indicate the personal opinion of its author. Each statement presents, rather, the claims that might be made about the strengths and weaknesses of the program based upon the evidence gathered during the evaluative process. In this way, the value of the program emerge to those people who must assess its worth and must judge it.

Responsive Evaluation has yet another important benefit for the evaluator and his client. Because it is not restricted to elaborate paradigms and a statistical format, the final report can be written in a natural communicative style. There are many examples of the natural style of evaluative reporting that can stimulate the reader's interest in evaluation as well as educate the reader in alternative evaluative process. They include reports by Stake,³ Brauner,⁴ Lang,⁵ and Dvorak.⁶ Each of these examples emphasizes the service function rather than the research function of evaluative reporting. An evaluation must provide useful information to a client. The results may or may not be generalizable to other situations, but they must be valid for the individual program under evaluation.

Summary

Even though the evaluation plans and final reports of the Responsive Evaluation model will vary in direct relationship to the diversity of programs being evaluated, the following

³Robert E. Stake and Craig Gjerde, "An Evaluation of TCITY, the Twin City Institute for Talented Youth, 1971." AERA Monograph Series on Curriculum Evaluation, Vol. 7 (Chicago: Rand McNally, 1974).

⁴Charles Brauner, "The First Probe." AERA Monograph Series on Curriculum Evaluation, Vol. 7 (Chicago: Rand McNally, 1974).

⁵Donald P. Lang, "An Investigation of the Effectiveness of the Methodologies of Music Instruction Used in Selected Schools of the West Hartford, Connecticut, School System." unpublished doctoral dissertation (West Hartford, Connecticut: Hartt College of Music, University of Hartford, 1978).

⁶Wayne D. Dvorak, "Orff Music Education for Disadvantaged Children—Language Arts and Mathematics Through Music." evaluation report (Hartford, Connecticut: Connecticut State Department of Education, 1975).

principles are common to all:

- Parity. The value of an education program is determined by people, not the research design.
- Ubiquity. Evaluation is a partner to all educational endeavors.
- Diversity. There is no single, ultimate truth in the value of any education program.
- Utility. Evaluation must be a service to its clients.
- Redundancy. Understanding of any education program comes with repeated encounters.
- Ambiguity. It is neither necessary nor desirable to force a consensus about the value of any education program.
- Generalizability. Important in the work of preordinate evaluators, generalizability may not be worth the effort. Each program must be evaluated on the basis of its unique aspects.⁷

In his book, *Zen and the Art of Motorcycle Maintenance*, Robert Pirsig reveals a new and thought-provoking way of looking at some of the premises upon which Western thought is based, including those thought processes normally associated with education, research, and especially the scientific method. Stake frequently has said that *Zen and the Art of Motorcycle Maintenance* is the best book on education evaluation that is yet available. The application of some of these ideas to arts education evaluation is not only relevant, it is essential. As Pirsig reminded us:

It's been necessary since before the time of Socrates to reject the passions, the emotions, in order to free the rational mind for an understanding of nature's order which was as yet unknown. Now it's time to further an understanding of nature's order by re-assimilating those passions which were originally fled from. The passions, the emotions, the affective domain of man's consciousness are a part of nature's order too. The central part.⁸

Maybe these ideas only work in literary form (fiction at that) and are not applicable to evaluation of AGE programs. But as Stake himself has said, "... were he to have greater ties with the anthropologist, the journalist, and the poet, the contemporary evaluator might have dealt himself a more responsive assignment."⁹ The analogy is clear. For effective evaluation of AGE programs, the logic is inescapable. □

⁷Robert E. Stake, "The Seven Principal Cardinals of Educational Evaluation," handout for presentation at AERA annual meeting, Chicago, April 1972.

⁸Robert M. Pirsig, *Zen and the Art of Motorcycle Maintenance* (New York: Bantam Books, Inc., 1974), p. 287.

⁹Robert E. Stake, "An Approach to the Evaluation for Instructional Programs (Program Portrayal vs. Analysis)," paper delivered at AERA annual meeting, April 4, 1972, p. 1.

An Evaluation of

TCITY

**the Twin City Institute
for Talented Youth**

1971

**Designed and carried out by
Robert Stake and Craig Gjerde
of CRCE, University of Illinois**

89

TCITY-1971 EVALUATION REPORT: CONTENTS

Page

The Setting and People	Charles Carson	1
Goals		2
Admissions		3
Calendar		4
Class Aims and Accomplishments		5
Field Trips	Craig Gjerde	6
The Open Campus	Terry Penny	7
Student Attitudes		8
A Participant Observer's Perspective	Ben Stake	10
A Student's Perspective	Betsy Marsden	17
Collective Perspective (Students)		18
Collective Perspective (Faculty)		19
Emphasizing Cognitive Talents		20
Emphasizing Creative Talents	Herb Grika	21
The Light-House School	Bob Berry	22
Management	Bob Stake	24
Miscellaneous Items		25
An Advocate's Statement		26
An Adversary's Statement		27
Editorial: Talented Youth in Perspective		28

EVALUATION PLAN

The evaluation plan for this report was derived from a journal article, "The Countenance of Educational Evaluation", by Mr. Stake. It calls for (1) a thorough description and (2) statement of numerous personal judgements of the Institute. The plan emphasizes the collection of data on the background against which the Institute activities take place, the activities themselves, and the results.

Because the Institute is of short duration; because the objectives of the staff are grand, elusive, and diverse; and because our tests are insensitive to many student learnings; the Results reported here were gathered by direct observation. Achievement tests were not used. As a matter of course, it was considered more important to get a picture of the results for a whole class than for the individual members of the class. Student views as well as teacher and evaluator views were collected.

The daily activities of the evaluation people were to be closely correlated with the activities of the teaching and administrative staff, doing something sometimes called "formative-evaluation:" helping the staff raise questions, gather evidence, and solve procedural problems.

For the "outsider" this cooperation raises the question of objectivity of the final report. To offset in part this weakness, this evaluation report features an advocate's report and an adversary's report, one summarizing the most favorable arguments in support of the Institute and the other summarizing the most damaging criticism. The reader is left with responsibility of resolving these conflicting arguments.

Evaluation staff: Bob Stake; Craig Gjerde; Carol Hansen; Cathy Ahl



TCITY-1971 EVALUATION REPORT: THE SETTING AND PEOPLE

The 1971 session of the TWIN CITY INSTITUTE FOR TALENTED YOUTH was held on the campus of Macalester College in St. Paul, Minnesota. Numerous activities extended to the urban areas nearby, out into the state, and even to West Berlin and Mexico City. Students back-packed across Isle Royal, dug for Indian artifacts at Fort Sweeney, visited Twin City and Chicago industries, and talked to officials at City Hall. Still, most Institute activities took place in typical classrooms, at Macalester. See map below.

This was the fifth summer. Previous Institutes had been held at Murray High in Saint Paul and Marshall University High in Minneapolis, and at Macalester, and Augsburg College in Minneapolis.

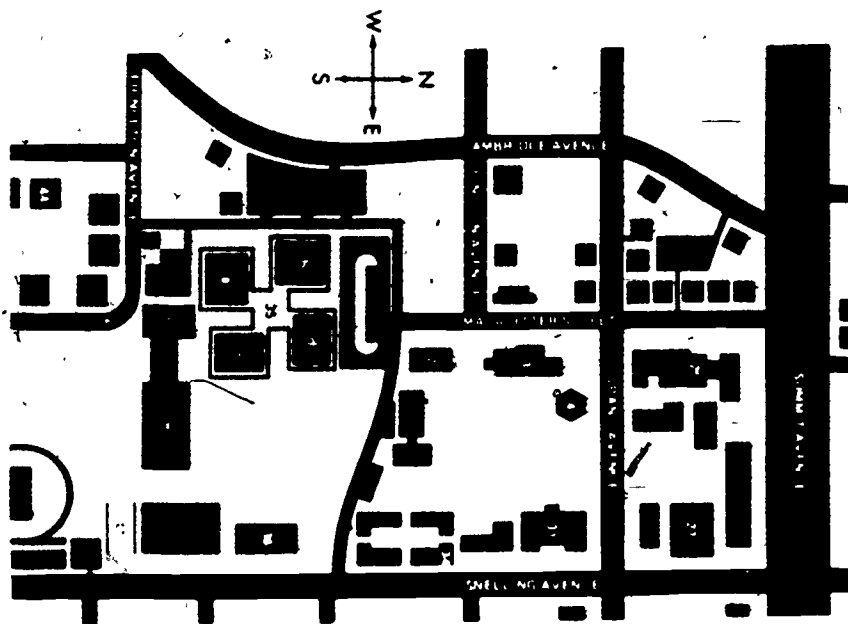
The Institute is a summer school experience for specially talented youngsters in the ninth through twelfth grades of Minneapolis and St. Paul schools. It is supported by these two districts, by a substantial grant from the Harrington Foundation, and by numerous donations. No tuition is charged.

About 820 students attended in 1971. The number has been growing each year. The length of the term has increased too, now up another week to seven weeks. The 1971 term extended from June 15 to July 30.

The teaching staff consisted of 28 master teachers and 28 associate teachers. Master teachers were selected from schools throughout the metropolitan area and beyond. The associates were selected from the two urban districts. Additional teaching help became available from University of Minnesota students seeking experience, stimulation or course projects. Three administrators, two evaluators, and a secretary also were staff members.

Class Spaces

- 35 American Studies
- 38 Archeology
- 39 Astronomy
- 39 Chemical Bonding
- 38 Computers
- 39 Computers in Science
- 36 Dance
- 38 Ecological Biology
- 35 Environmental Accounting
- 25 French
- 25 German
- 2 Graphic Arts
- 39 Mathematics for the Disenchanted
- 35 Music
- 35 Poetry
- 34 Russian
- 2 Sculpture
- 20 Sesame Street
- 4 Spanish
- 35 Theatre
- 35 The Urban Hero
- 2 Wilderness Leadership
- 35 Writers Workshop
- 31 TCITY Office
- 3 Library, AV Ctr



TCITY-1971 EVALUATION REPORT: GOALS

The primary objective of the Twin City Institute is to create an educational program that has strong academic and social appeal for students who possess a variety of artistic, language, scientific and leadership talents. The program is not designed to repeat the regular school experience, nor to repair it for the disenchanting. The Institute is designed to create a special experience, cutting across student interests, group identification, idea exploration, and the traditional school curriculum.

The Institute is willing to take risks. Teachers will develop courses that have their locus in theory or intuition rather than in more conventional curricular constructs. The staff will recognize that any program, traditional or experimental, that chooses to work in an atmosphere of freedom, where trust is extended in social relationships, and where new ideas are encouraged, is going to be vulnerable to charges of aimlessness and confusion.

The Institute staff--particularly teachers involved in such areas as the arts and social sciences, where truths are more subjective and where curricular goals are disputed--will develop programs that emphasize inquiry. To create a compelling and stimulating learning environment, TCI teachers will develop problem-centered courses which encourage thought, inquiry, and creativity.

Finally, teachers will introduce students to a total Institute environment. The Institute campus will be open to the young, sensitive, inquiring mind. Students will share with each other--through conversation and exhibit--their productions and products: telescopes, dances, poetry, music, foreign language fetes, newspapers, etc. Against a background of computer programming, art, pottery-making, dance, athletics, folk dancing, games, canoe building, students will meet and talk and create.

Having come interested in math, science, poetry, or archaeology, students will leave more committed to understanding and appreciating the total concerns of men. Success for the Institute will mean that students and teachers have infected each other with a personal honesty and will have demonstrated that learning can truly be humanizing.

Charles Caruson, Director, TCITY.

Goal Evaluation. Evaluators have an obligation to raise the question, "Were the right goals pursued?" Different people have different ideas, of course, as to what the right goals are. Still, goals and priorities should be evaluated.

Many teachers and curriculum specialists endorse TCITY's increased emphasis on humanization, personal awareness, and problem solving, and decreased emphasis on skills and knowledge. Some teachers and many parents disagree, wanting the school experience to pay off in answers to the classical academic questions--the kind that get students employed, admitted to college, etc. In the eyes of the evaluators, the TCITY goals are worthy goals, suitably discussed and reasonably operationalized.

TCITY-1971 EVALUATION REPORT: /ADMISSIONS

Students residing in Minneapolis or St. Paul during 1970-71, and in grades 8-11, were eligible for the 1971 TCITY Summer Institute.

Early in the year, Associate Director, Robert Rose visited Twin City high schools to promote the Institute and to encourage applications from interested students. As the March 26 deadline approached, it became obvious that some schools had few applicants. Counselors there were asked to encourage students to apply. Some emphasis was placed on recruiting disadvantaged students from inner-city schools.

Students were asked to select a first, second, and third course preference and to ask two teachers for recommendations. School counselors provided academic grades, a recommendation, and the combined verbal and numerical scores from the Differential Aptitude Test.

Some courses, like Wilderness Leadership and Sesame Street, had far too many applicants for the number of openings. A few, like Music, were under-subscribed. Students listing music as a second or third choice were added to rosters.

Mr. Caruson and Mr. Rose examined the forms, seeking especially talented students with DAT scores over 75 and also considering low achieving youngsters with no marked behavioral problems, whose teachers and counselors felt that the Institute would draw out their special talents. About 200 of these "Special Admits" were accepted. Some of the teachers examined applications and made recommendations, but the responsibility and burden of selection fell mainly upon Caruson and Rose.

957 of the 1247 applicants were accepted, with the expectation that some would cancel. Of this number, 771 quickly accepted. Some students accepted later, and a few students who had been in the "alternate" status were added to the accepted list to fill low classes. It was expected that 775-800 would attend.

On Monday, June 14, 838 students appeared at Macalester College to start the Institute. During the course of the 7 weeks about 7 percent of these students disappeared from their classes, leaving approximately 800 on the final days. These were pretty evenly divided between Minneapolis and St. Paul and split 450 to 350 in favor of girls. There were 45 scholarship students. About 16 percent of TCITY-1971 students were repeaters from 1970; 4 percent had also attended in 1969.

A decision to admit eighth graders this year led to problems-- 70 additional schools to contact, new counselors to involve, and an absence of DAT scores. Some teachers thought 8th graders were less able to discipline themselves, and a few students complained about communication difficulties, but the consensus was that the younger students adapted well and caused few problems. Some larger classes felt that difference in experience was a better criterion for sub-grouping than age.

Any program for "talented" youngsters is potentially faced with charges of elitism and racism. Many teachers thought there was less elitism and snobbery at TCITY than in honors classes in regular school. Minority enrollment was, perhaps, 8 percent.

Admissions questions that continue to be discussed by the directors and the TCITY Board: Is there a better way to identify talented students than by test taking and faculty recommendations? How large a percentage of "Special Admits" can be enrolled without endangering the academic spirit of the Institute? Are the admission procedures consistent with the philosophical goals of the Institute?

TCITY-1971 EVALUATION REPORT: CALENDAR

A calendar of events for TCITY-1971

- Fall Course planning by master teachers
- Spring Selection of associate teachers; screening of students
- Apr. 15 Notice to applicants about acceptance
- May 15 Macalester College--classes met, planned their summer
- June 14 TCITY opened with 823 students, 56 teachers
- June 21 All Physical Science classes left for two weeks at Tamarac
- June 22 German class students (12) flew to Germany for one month
- July 1 Dance and Music combined to present a program; Russian classes crowned a czar, other classes joined in.
- July 9 Sesame Class visited a Minneapolis City Council meeting
- July 16 Language picnic; ragtime pianist in music class
- July 22 TCITY Board Luncheon at Macalester; Open house; AIM Indian group halted Archeological dig at Welch, Minnesota
- July 26 Environmental Accounting on a three day trip to Chicago
- July 28 Theatre show "Rip-Off"; Dance production; Poetry reading
- July 29 Art display; Poetry reading; Music on the Mall; Interpretive dance; Evening theatre production for parents.
- July 30 Institute ended; Language groups performed folk dances

Monday, July 19: A typical day at the Institute

Around 8 a.m. students started gathering in the bus area: the Sesame II class and combined French and German classes left for a Cannon River canoe trip with Wilderness class guides. Rest of Wilderness class left for Isle Royale on an 8 day back-packing trip. Biology classes went to St. Paul Ramsey Hospital for a series of physical tests.

About 8:15 or 8:30 other classes started in classrooms on the Macalester campus. Most of these classes took a short break about 10:30 and then continued until 12:00 or 12:30. Fifteen students started on a bike trip along the Mississippi River at 11:30. Some time during the morning, Basic Computer challenged Environmental Accounting to a volleyball game. A few students played tennis, frisbee, softball, and went swimming later in the day.

After classes, opponents in the chess tournament started finding each other. At 1:00 Dr. Mitra demonstrated the use of acrylics, and Math students began constructing a geodesic dome. At 2:00 Avi Davis' Dance class met in the old gym.

Art rooms were busy with painting and pottery; Astronomy students were grinding lenses for their telescopes; a few Science students were finishing their redwood and fiberglass canoes. Poetry students were in the office duplicating their "Broad-sides" for distribution to people in the streets, and there were other activities.

By 1:30 most of the 150 students who stayed for afternoon activities on campus had departed for home, except for the canoe builders who stayed till 6 or 7 in their race against time.

CITY-1971 EVALUATION REPORT: CLASS AIMS AND ACCOMPLISHMENTS

AMERICAN STUDIES 30 students
Master: Gene Lohman; Assoc: Bob Niemela

Aims: To examine "American" experience as revealed through histories and art forms. To relate contemporary problems to American traditions and more.

Sketch: Built a group, let it plan activities. Used thematic rather than chronological study of American events, followed such themes as "the Making of a President", "packing the Supreme Court", "important American writers since 1945". Took 6 field trips.

Comment: Sustained high level of interest and excitement. Students involved more in interpretation--often with too few facts--than in acquisition of facts. Teachers, students were open, honest about their ignorance and bias--a good scene.

CHEMICAL BONDING 38 students
Master: John Edwards; Assoc: Don Land

Aims: To establish electron patterns of atoms. To relate these patterns to the periodic table. To relate the electron distribution pattern to spacial arrangements that can be demonstrated by 3-D figures. To use determined pattern of electron distribution to predict chemical formulas. To demonstrate the shortcomings of having a single model.

Sketch: Topics covered and understandings reached. Lab used to demonstrate the actual making of products discussed.

Comment: Took younger students, showed that 14 year olds--though sometime left breathless--can learn complex chemical concepts.

FRENCH 36 students
Master: Barb Gunderson; Assoc: D. Hopen

Aims: To learn French through simulation of experiences in France, to compare life in France and the USA, to consider the uniqueness of French Canada.

Sketch: A field test for simulation units developed by U of M people, e.g., preparing French bread, preparing crepes, playing Bridge, Cyrano de Bergerac, Weather Expressions.

Comment: Teacher preparation excellent, a fun class.

SCULPTURE, POTTERY 37 students
Master: J. Fontaine; Assoc: P. Fitzgerald

Aims: To create an awareness of the relationships between man, his environment and his art; a greater perception of art and the environment; a greater sensitivity to form, color, and design; an introspection into why men make art--and why each student does.

Sketch: Students at work, creating three dimensional art objects of clay, plaster, paper mache, metal, styrofoam; potting, welding, jewelry making.

Comment: For most students a new experience, a personal expression.

COMPUTERS IN SCIENCE 28 students
Master: John Crocker; Assoc: Paul Gifford

Aims: To solve science problems using the computer. Emphasis on simulation. Examine career opportunities.

Comment: The highly successful mini-course, canoe building, also taught by Crocker and Gifford, almost swamped this course.

ARCHEOLOGY

44 students

Masters: Dan Conrad, Bill Hobson
Assoc: Skip Olson, Dan Willette,
Jan Peterson

Aims: To participate in a real social science field study, relating the methods and concepts and issues of the several sciences. To excavate a village site near Fort Sweeney to professional specifications. To consider the life and culture of the Mississippian Indian.

Sketch: At the instigation last spring of the master teachers, Conrad and Hobson, Mr. Carson arranged with the State Historical Society to excavate portions of an Indian site near Fort Sweeney. The Chief State Archeologist was reluctant to risk a dig by high school students, but his colleague, David Nystuen, saw its educational merit and got the okay. Les Peterson, a younger staff archeologist, accepted responsibility for supervising the dig.

Members of the TCITY staff and a party of students spent a May weekend at the Fort Sweeney site for orientation and training. They discussed the historical, social, religious, and ecological aspects of Mississippian Indian life. They considered the technical and ethical aspects of a 5 week encampment at the Welch Village ski chalet. Planning at that point was careful and detailed.

During the first week of the Institute the "Special Math" class--using precise instruments--surveyed the site. A village was expected there, near the visible burial mounds. The "village" was the target, not the mounds. About a dozen 10x10 foot squares were randomly selected for excavation.

At the beginning of the second week 44 archeological students, 5 TCITY staff members, several Education

students from the University of Minnesota, and Les began digging. They dug with Crowels, carefully screening the dirt, examining the surfaces. Slowly, even painfully, those first weeks, the students lowered the surface 18 inches. Nothing. No artifacts. None of the expected tell-tale posthole discolorations. No village. The fact that they had provided a basis for revision of the archeological map only partly abated their disappointment.

During the evenings they played volleyball, canoed, swam, partied, discussed digging techniques, revised camp rules, (broken first by teachers longing for a beer), made clay pots, took pictures, ate, talked about the similarities and differences between grave-robbing and archeology, ... One evening they invited families from neighboring Welch Village to share a pot luck supper and songfest. The evenings went well.

In the fourth week, digging adjacent to the mounds, they unearthed pot chards, missile points and other artifacts--not of the expected Mississippians but of an earlier Indian, the Woodland Indian. The find, slowly realized, was a basis for revising both the local archeological map and calendar. The archeologists were as pleased as the students.

In the final week on-site, following Les Peterson's evaluations, the diggers moved closer in. The plan was to avoid the middles of mounds, but to dig at their edges so as to determine the age and circumstance of the burials. On the evening of the next to last day on site, the diggers were confronted by a large group of Indians, who identified themselves as members of the American Indian Movement. Outraged at seeing students near the graves, they shoveled dirt back into the excavations, even as the startled students scrambled out. Some students were frightened off; some

pleaded for talking it over. The AIM members went on to burn field notes, pictures, tools. One TCI student was physically shaken, all were emotionally shaken.

Some students and their antagonists returned to the Chalet. More film and diaries were taken, but apparently nothing else. A student asked the AIM people to stay for dinner. They accepted; then the cook announced she could not feed them all. Further talk did no good. The AIM demanded abandonment of the site by 9 p.m. The camp was cleaned up and vacated by 8.

The Twin Cities press carried the story on page one. The Indians called, "sacrilege!" Mr. Caruson answered, "You need to understand a culture in order to honor it." Several days later some of the Indians came to campus to discuss the encounter with the students. To a large degree they were able to reach a reconciliation.

Comment: For these students this was a trying, exhausting, fulfilling summer. The issues and struggles were serious--they learned a great deal about archeology, about communities and about themselves.

RUSSIAN 35 students
Masters: Don Ryberg; Cath. Filipovich

Aims: To experience, to ponder the Russian way of life--communication, incentives, political system, culture. To learn some of the language through informal discussions, role-playing, class projects. To send a group to visit the USSR.

Sketch: Using such activities as a coronation and publication of a newspaper, the students became familiar with some Russian language and culture.

Comment: Teachers competent, Industrious, ingenious, but the class too "academic" for some students.

JRBAN HERO 24 students
Master: S. Sandell; Assoc: R. Pestello

Aims: To search for today's urban heroes, to contrast them to the John Waynes, Charles Lindbergs and George Washingtons; to visit the city, to learn of urban survival, of power, of men.

Sketch: Students first developed a class-community; they became aware of the personal, social, and economic motivations of the city; they developed skills to express their ideas in a creative medium (e.g., photographic essay, guerilla theatre); they explored themes (e.g., the individual acting according to his convictions in a man-made environment, value conflict, and conflict resolution).

Comment: Teachers imaginative, sensitive; class took an incredible number of valuable field trips.

GRAPHIC ART 36 students
Master: Gopal Mitra; Assoc: Robt. Horton

Aims: Painting, drawing, exploring new media. "We will approach art by considering each student's personal talents and interests." Creativity expression to be refined through extensive studio work. Understandings of the freedoms and disciplines of art to be stressed.

Sketch: Each student completed several canvases. Themes such as Oriental religions and self-evaluation were discussed. Personal guidance given.

Comment: Excellent teaching, social experience. Mitra, an excellent teacher, had a bold philosophy of life, not appreciated by all. Nevertheless, as artist-in-residence, as one who demanded serious application of learned skills, he was a major TCI-71 asset.

SESAME STREET I, II 55 students
Masters: Roger Clemence, Tom Walz
Assoc: William Bichel

Aims: An educational experience using the television medium to convey life goals and knowledge of basic social sciences to urban children; with field trips to gather ideas, role playing to learn perspectives, script building to understand communication.

Sketch: Two classes, one oriented to the stimulation of the city, one to on-stage production.

Comment: Course philosophically elegant, operationally a mess. Teachers personally sensitive but pedagogically insensitive; so unwilling to impose (to structure) that learning opportunities slipped away. Students turned on, will continue into fall.

MATH FOR THE DISENCHANTED 22 students
Master: A. Indelicato; Assoc: G. Scheffer

Aims: To give the student who dislikes math an understanding of its utility and a look at some interesting abstract ideas. Sample topics: Simulation, surveying, topology, crystals.

Sketch: Students decided to focus on two topics, surveying and geodesic domes. Computer math was added later. Students surveyed Fort Snelling site to assist archeology class.

Comment: Class turned out to be more "just curious" than "disenchanted" but rapport, motivation, and involvement did grow.

MUSIC 25 students
Master: Johs. Reidel; Assoc: Judy Evans

Aims: To increase awareness of classical, folk, and pop music; to contrast the music of North, South, and Central America. A course for developing understandings of music, not for developing the skills of the musician.

Sketch: Musical sensitivity grew with exposure to ethnic music and basic concepts in musicology. Work with ensembles.

Comment: Course content excellent; studio facilities inadequate. Staff contributed little to other fine arts courses on total Institute; no real ties to Woodstock Nation.

WILDERNESS LEADERSHIP 36 students
Masters: Constance Gore, Bob Tauring, Marc Wanwig; Assoc: Ron Pressley

Aims: To combine the skills of camping, canoeing, backpacking, etc. with the responsibilities of leadership and organization, so that students can guide others into the wilderness safely and with a sensitivity to its social, esthetic, and ecological aspects. Practical experience.

Sketch: Students practiced making plans for trips; assisted groups on Cannon River canoe outings, etc.; learned the problems in looking out for others in ordinary as well as hardship and deprivation conditions, and constructed own pack frames and other items of equipment.

Comment: A very popular and worthwhile educational experience, it should be expanded for forthcoming Institutes, but should be changed to provide more instruction in guiding and wilderness living.

SPANISH 49 students
Master: Ramedo Saucedo; Assoc: Vic Barela

Aims: To experience Latin American life on campus, in Spanish homes in the city, in Mexico City.

Sketch: Activity oriented, to Spanish movies, Spanish homes, Spanish kitchens. 22 students visited Mexico City.

Comment: Well managed; class suffered a bit with the midterm departure of the favored few, but recovered.

GERMAN

40 students

Master: D. Cameron; Assoc: P. Scheppe

Aims: To learn what the German people are like through a study of language and culture. To send a small group to visit Germany.

Sketch: Teacher worked on attitudes toward language, and attitudes toward the work it takes to learn a foreign language. Teacher claimed students developed a more sophisticated perspective of the student's own culture and of the German culture. 12 students to Germany.

Comment: Associate teacher, left in charge, gave us evaluators a hard time.

THEATRE CLASS

27 students

Master: R. Declercq; Assoc: M. Pfeifer

Aims: The students in this class will prepare and participate in at least one dramatic production. They will get students from other classes to share their drama experiences. Story-theatre (Three Bears, Pogo, etc.) will be developed.

Sketch: Students put on RIP-OFF, a collage of story-theatre productions. Worked with poetry groups to contrast and combine media of expression.

Comment: Students worked hard, responded with enthusiasm and teamwork to DeClercq's direction, the strong internal rapport, love and respect for each other.

POETRY

30 students

Master: John Caddy; Assoc: Joyce Thomas

Aim: "Imagine the human dawn. Imagine an ancient form of man, preparing for the hunt..." (So began the catalogue description of this class) The promise contained ways to awaken "the sleeping voice", to express oneself in "the ancient natural way way". The method: establish the group, write what is personally significant, read recent poetry, talk.

Accomplishment: Perhaps 3/4 of the students gained a strong sense of trust in other group members, most did the writing seen needed as a base for "sensory awakening", personal expression.

Comment: A most sensitive, mature teacher; students productive, somewhat too dependent, with a not unexpected fixation on sex-theme poetry.

ENVIRONMENTAL ACCOUNTING 25 students
Master: Dennis Daly; Assoc: Mel Pibal

Aims: To develop interest in accounting and business; to portray their multifaceted, dynamic character and the role they play in our society.

Sketch: Discussions, simulation, role-playing, field trips, a trip to Chicago. Students learned how accounting is used to help individuals and organizations attain their goals.

Comment: Program supported by the Certified Public Accountants, State of Minnesota. Teachers ingenious, one perhaps too verbal.

ECOLOGICAL BIOLOGY

55 students

Masters: Tony Angellar; Harold Strobel
Assoc: Bill Holmson

Aims: Through team teaching, to pose and ponder questions about the basic functions and reactions of animals and man, e.g., organismic learning, body-environment interaction, chronobiology. To measure and plot circadian rhythms, to carry on individual and group projects.

Sketch: Completed units on animal behavior, learning, body-environmental interaction, chronobiology and circadian rhythm. Completed fewer projects than intended.

Comment: Great variety of field events; teacher-talk good but probably too much of it.

DANCE 35 students
Teacher: Mary Rae Josephson
Assoc: Linda Nelson, Avi Davis

Aims: To involve each student in moving, thinking, feeling situations; encouraging her (him) to rely on an expanding movement vocabulary, an awareness of self and sensitivity toward others, an increasing knowledge of dance as an art form and its relation to the other arts, a desire to express herself creatively.

Sketch: Improvement of self-image and group awareness were seen as the two main accomplishments, with variation, of course, across individuals. Time spent in sustained warm-ups, body movement assignments, watching films of master dancers, in theme development, learning concentration, and developing a viable group.

Comment: Teachers highly competent, worked well together and with other teachers. Class needs boys, should be organized as supporting rather than as a principal enrollment.

ASTRONOMY 32 students
Master: Fred Brett; Assoc: Dennis Mallum

Aims: To examine the interdependence of scientific facts as they relate to events of the universe. To develop analytic and inferential skills. Individual projects, particularly building personal telescopes.

Sketch: Large blocks of time spent on grinding, polishing lenses. Used U of Illinois Astronomy Series for background. Camped out to use their telescopes.

Comment: Students took great pride in their work. Sustained involvement.

WRITERS WORKSHOP 24 students
Master: Hu Anderson; Assoc: R. Klepperfeld

Aims: To give students who-want-to-learn-to-write a chance to learn about their own writing. To learn how to search for things worth writing about.

Sketch: A subgroup of the class published "La Bouche", the student newspaper; others in class wrote stories and developed their personal writing styles.

Comment: A strange class, probably the least learning-oriented in the Institute. Students were not pushed to produce; many did not. Over half the students wished they had enrolled in something else--for the rest of the Institute only 15% had that wish.

COMPUTERS 58 students
Masters: Ed Anderson, Fred Blaisdell
Assoc: Jon Gross, Virginia Toms

Aims: To introduce new students to computer uses and programming. To introduce advanced students to FORTRAN. To consider problems in number theory, geometry, economics.

Sketch: After learning the computer language, the students went on to learn many contemporary uses of computers. They investigated the effects of computers on modern living. They learned of new technical developments in the area of on-line usage.

Comment: Facilities crowded, frequent trouble with lines. Strong afternoon following.

TCITY-1971 EVALUATION REPORT: FIELD TRIPS

Field trips were encouraged by being easily arranged: the teacher filled in a bus request form the day before his excursion, and the bus appeared at the appointed hour. Consequently, field trips were often used by teachers to supplement classroom activities and to take students to unique learning sites.

Some classes were absent so much from the Macalester campus that it was difficult to decide whether they were on field trips or whether they were permanently based off-campus. All of the Physical Science classes spent 2 weeks at Camp Tamarac. At camp, there was an integration of class activities into the Tamarac environment with canoe trips, nature walks, innertube floats and visits to Itasca State Park. Teachers felt that two weeks was too long. Students thought it about right, but one pleaded, "let kids go home on the weekend and get clean clothes." The archeology class spent 5 weeks at the dig site. Three days in Chicago ended the Institute for the Environmental Accounting class. The German and Spanish classes had 13 and 22 students and a teacher to Germany and Mexico. Almost all classes spent from 3-5 days canoeing on Cannon River or St. Croix with brigade leaders from the Wilderness class. (One day of an per trip was about average.)

Teachers felt that the wilderness experiences helped to establish class identities. Such field trips heightened awareness of different life styles, stimulated discussion, and provided inspiration.

In retrospect, almost all students felt that these trips were fun and were good learning experiences. Other student comments were: "too much fun squeezed into too few days", "It was fantastic--I loved it", and "It should have come earlier to pull us together mentally". Students evaluated their field trips as "mostly good", "better than time on campus", and "extremely stimulating".

In addition to these longer trips, the class activities calendar was punctuated with one day excursions to places in the Twin Cities. Some of the sites and activities: Metropolitan Minneapolis-St. Paul tour; Jonathan, model housing development; IBM Rochester, computer facility; Pillsbury State Forest, wilderness; Apple River, inner-tubing; Teachers' homes; Metropolitan Stadium, baseball; Como Park, picnic; City Council meetings; Lino Lakes, model city project; General Mills; 3M Company; Univac; Minneapolis Institute of Art; Isle Royale, wilderness, St. Paul Ramsey Hospital, physical tests; New Ulm, "German" town; Duluth, Urban Hero trip.

Buses were expensive. Students generally contributed part of the bus expense and the cost of meals. But students and teachers agreed that excursions off-campus contributed much to TCITY-1971.

Craig Gjerde
Evaluator, CIRCE

TCITY-1971 EVALUATION REPORT: THE OPEN CAMPUS

The TCITY Campus was closed. The TCITY classroom was open. Within the classroom, inter-age grouping was surprisingly successful, within and across sex. Use of student interests, skills and attention to student desires was omnipresent. Students generally felt that they had ample freedom to pursue the subject matter of their course, across disciplines. Most felt free to come and go as they pleased.

Students did not often think of the Campus or the Institute. They had a small-group view and a small-group allegiance. There were exceptions, to be sure. The dance and theater group work was an outstanding example of meaningful cross-course interaction. And there was an infamous band of marauding poets and Sesame Streeters. But most of the students "reached out" only within their own group of thirty.

Some students labored under the mistaken notion that re-opting for another class would be tantamount to dismissal from the Institute ("and the waiting list is long!") Others sagely observed that their group was "the best anyway" so why should they bother learning about other Institute opportunities? The classes were open; the Institute was closed.

T. Denny
Evaluation Specialist, CIRCE

An Adversary's Statement--continued from page 27

*** Few students--or faculty--understand the selection procedures employed to staff the teaching cadre and to fill the student corps. Why should it be a mystery?

The worst has been saved for last. This report concludes with an assertion: the absence of a crucial dimension in the instructional life of T-CITY, that of constructive self-criticism, is a near fatal flaw. The observation and interview notes taken by the adversary evaluator over four days contains but five instances of students engaging in, or faculty helping students to become skillful in, or desirous of, the cultivation of self-criticism. The instances of missed opportunities were excessive in my judgment. Worse: when queried by the writer, faculty and students alike showed little enthusiasms for such fare. Is it too much to expect from Institute participants after but four weeks? Seven may be insufficient. The staff post mortem, "Gleanings", are a start--but it seems odd to start at the end.

The paucity of occurrence is less damning than the absence of manifest, widespread intent. Certain classes accounted for all the instances observed. They did not appear to be accidental. The intent was there. An Institute for talented high school youth cannot justifiably fail to feature individual and group self-criticism.

(Prepared by T. Denny; not to indicate his opinion of TCITY-1971, but as a summary of the most damaging charges that might reasonably be made.)

TCITY-1971 EVALUATION REPORT: STUDENT ATTITUDES

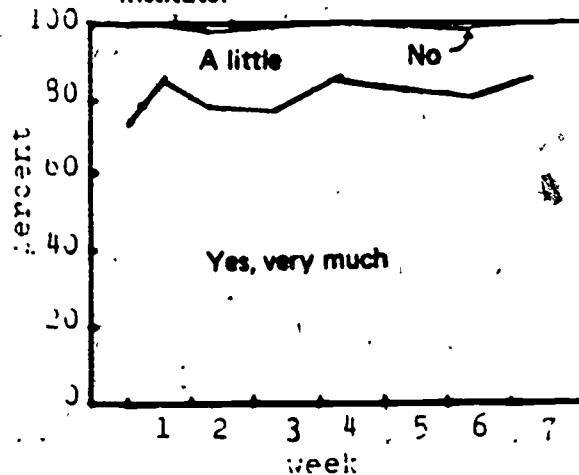
The Director and Evaluator tried to keep in touch with student feelings. Once a week each student was asked three questions taken at random from a pool of 12 attitude questions. Thus each question was answered each week by 75-200 students.

Four questions pertained to student learning. Every week at least 90% said they were learning a lot. The number of students saying TCITY is one of their best learning opportunities increased steadily from 75% to just under 90%. The number who reported themselves involved in creating or developing a project increased from 40% to over 60% by the third week, and leveled off. To the question, "Do you have the feeling that--when this Institute is over--you will say that TCITY has been a very satisfying experience?" about 80% said "yes" until the final week when the positive response rose to 100%.

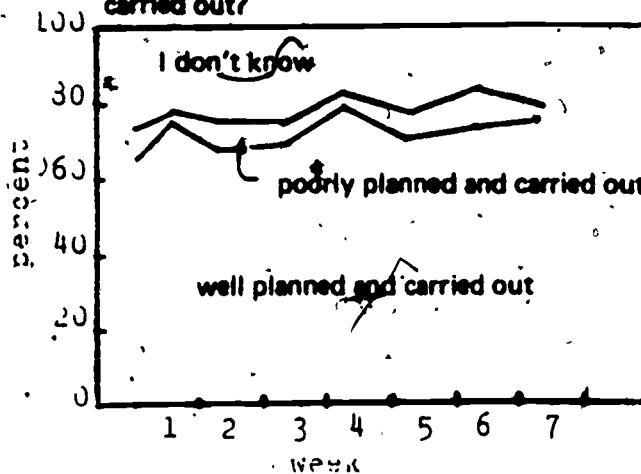
Three questions pertained to how well the student liked the Institute. One is shown at the right. Everyone responded each week that they liked the people at the Institute. About 75% were satisfied with afternoon activities though the percent was dropping toward the end.

At first, about half the students said they were getting too little information about TCITY events; by the seventh week a quarter still said so. Another administrative question is shown below-right. Another brought forth the almost unanimous opinion that to applicants, TCITY offers a selection of challenging, relevant and useful courses.

Are you enjoying yourself at this Institute?



Consider the whole Institute (and not just your class). Have the arrangements been well planned and effectively carried out?



Class _____ Date _____

104

Class _____ Date _____

Last name _____ First _____

04. Do you like the people at this Institute?

Yes _____

TCITY-1971 EVALUATION REPORT: A PARTICIPANT OBSERVER'S PERSPECTIVE

As my part in the evaluation of TCITY, I was a participant observer. I was a member of the Wilderness Leadership class and was involved in the guidance of other classes on canoe and camping trips. Having had a fair amount of previous camping experience I was able to observe carefully the learning experiences of other students.

It is apparent to me that students want to learn. At TCITY, teachers provided the opportunity for an interesting learning experience. Students took the initiative, and "away she goes".

Students will learn, even under adverse conditions. On one occasion after a nearly sleepless night under a non-rainproof shelter, a friend came to sit-out the rain under my dry shelter. He was in low spirits and explained to me that he was not learning what he wanted to. He said that he was learning what was wrong, but not what was right. But we figured out his teachers were not trying to teach him what was right. They were not trying to teach him the answers--they were trying to teach him the questions. And even though he said he didn't want to, he was learning. They kept putting him in situations where he would learn.

I believe the Wilderness Leadership class was a very successful one. The students could have been taught more than they were, and they could have learned more than they did. But they enjoyed it, even with adverse conditions, and learned about handling themselves and others in the outdoors. Lots of time was spent not being students, just being themselves. It seemed they learned just as well that way.

One of the important reasons for the success of the class was the freedom given the students; the freedom to choose whether or not to go to class or--in this class--which trips to go on. With the amount of involvement that was asked, a trip almost every week, it would have been easy for the students to reject it. But it was their choice, and they were glad to do it.

The important thing in all the Institute classes I worked with was the atmosphere that influenced students' attitudes. The atmosphere at TCITY was one that encouraged the student to learn because he wanted to, not because he had to.

The basic element of this atmosphere was freedom: freedom of movement in and out of the classroom, and freedom of choice of subject material. At TCITY I saw students moving, choosing, learning.

Ben Stake, Student
University High School
Urbana, Illinois



Gopal Mitra
TCITY-1971 Art teacher

Welding torches

Our minds are welded into
a useless facts index
that swells in schools.

Offered a torch
we ignite the
stagnated forms
sculpting our minds into
geodesic domes
and polyethelene bubbles.

Teachers/friends light
the torch
but we shape the sculpture.

Ignoring tired eyes
sleepless faces
arrive and awaken.

Life's paintbox
colors our experience
but we find it
here.

Why must we dry up
and falter come fall?

Betsy Marsden
TCITY-1971 Poetry Student

TCITY-1971 EVALUATION REPORT: COLLECTIVE PERSPECTIVE (STUDENTS)

The following was taken from the tally of the Student Final Critique Sheet. About 560 students filled out the sheet during the final week of the Institute.

We would like to know how the Institute class activity differs from the activity of the classes of your regular school.

	No Difference	Regular School	Institute
1. Which holds students more responsible for work?	<u>12%</u>	<u>68%</u>	<u>20%</u>
2. Which encourages students to "show off" more?	<u>36%</u>	<u>42%</u>	<u>22%</u>
3. In which do students try out OWN ideas more?	<u>9%</u>	<u>4%</u>	<u>87%</u>
4. In which is more time wasted?	<u>31%</u>	<u>47%</u>	<u>22%</u>

And how do your teachers compare?

	No Difference	Regular School Teachers	Institute Teachers
5. Which explain things better?	<u>38%</u>	<u>11%</u>	<u>51%</u>
6. Which know the subject matter better?	<u>39%</u>	<u>7%</u>	<u>54%</u>
7. Which understand students better?	<u>28%</u>	<u>6%</u>	<u>66%</u>
8. Which resist the urge to talk all the time?	<u>42%</u>	<u>16%</u>	<u>42%</u>

Please rate the following features of the Institute as to how important they were to you:

	Extremely Important	A little Important	Not Important
27. Close contact with teachers	<u>66%</u>	<u>29%</u>	<u>5%</u>
28. Close contact with students	<u>80%</u>	<u>17%</u>	<u>3%</u>
29. Opportunity to study for extended period each day	<u>46%</u>	<u>40%</u>	<u>14%</u>
30. Trip to camp, canoe trip, etc.	<u>46%</u>	<u>40%</u>	<u>14%</u>
31. Exhibits, performances put on for "outsiders"	<u>36%</u>	<u>30%</u>	<u>33%</u>
32. Diversity of students	<u>60%</u>	<u>35%</u>	<u>5%</u>
33. Afternoon symposia	<u>28%</u>	<u>52%</u>	<u>20%</u>
34. Concern that these people have for human problems	<u>69%</u>	<u>26%</u>	<u>5%</u>
35. Faith that these people have that these problems can be solved	<u>69%</u>	<u>26%</u>	<u>5%</u>
36. Being treated as a mature person	<u>85%</u>	<u>13%</u>	<u>2%</u>

Here are several goals of the Institute. Please rate the Institute on how well it met these goals, even if you only have a vague idea of what the whole Institute was doing.

	Excellent	Barely Passing	Failed	Don't Know
37. Provide an educational program that is challenging, stimulating, relevant	<u>86%</u>	<u>10%</u>	<u>0%</u>	<u>4%</u>
38. Provide master teachers with the highest ability-to-teach	<u>81%</u>	<u>9%</u>	<u>2%</u>	<u>9%</u>
39. Provide younger teachers a good opportunity to learn more about teaching.	<u>69%</u>	<u>14%</u>	<u>1%</u>	<u>16%</u>
40. Develop curriculum ideas that can be used in the regular schools.	<u>67%</u>	<u>13%</u>	<u>3%</u>	<u>17%</u>

TCITY-1971 EVALUATION REPORT: COLLECTIVE PERSPECTIVE (FACULTY)

In the final week of the Institute all master and associate teachers were asked to complete a four page questionnaire. About 3/4 of the teachers did so.

The master teachers listed the major satisfactions in TCITY-71 as the opportunity to teach the way they wanted to; to work with highly motivated, able students; and to make the learning-experience a humanizing encounter. The associate teachers emphasized these things plus the fact that the Institute was a major learning opportunity for them. Both groups were strong in their praise of TCITY.

The least satisfactory aspects of the experience these teachers had were the administrative arrangements (too vague, too little pre-planning) and the workload (too much expected). Some master teachers objected to the large number of students enrolled in their classes. Almost all the associate teachers reported some unpleasant interpersonal experience during the summer.

Special features drawing approval were the May orientation sessions and Wilderness trips. New arrangements that teachers objected to were the admission of junior high students and the foreign language travel alternative.

Among the suggestions for improvement were the following:

- 1) Increase involvement of University students as teaching aides.
- 2) Better communication about TCITY events, deadlines.
- 3) Less involvement in formal evaluation
- 4) Extend the opportunities to suburban and out-state students.

By and large the faculty responses in 1971 were similar to what they had been in 1970. There seemed to be better communication across faculties in 1970 but better communication to and from the Directors in 1971. Communication remains as an important problem.

Most of the master teachers were pleased with what they had done, but still would like to offer a different course or teach the course differently next year. They strongly hoped that TCITY could be repeated at Macalester again in 1972.

To detail their activities for the summer and to react more thoroughly to such evaluation issues, each master teacher submitted a synopsis called by Caruson "The Gleanings". These reports were loaded with recommendations for summer courses and Institutes.

TCITY-1971 EVALUATION REPORT: EMPHASIZING COGNITIVE TALENTS

In a sample of about half the classes, responses to the Class Activities Questionnaire (CAQ) indicated a few things about the individual classes and the Institute as a whole: Every class showed much more emphasis on the higher level thought processes--application, analysis, synthesis, and evaluation--than on the lower level thought processes--memory, translation, and interpretation.

Every class indicated that ideas were valued more than grades, and that there was much humor and enjoyment of ideas. There was very little lecturing or test stress. Enthusiasm, independence and divergence were indicated in each class summary.

Students reported the percentage of teacher talk to be between 10 and 75 percent with the median at 40 percent.

The CAQ profile for the Urban Hero class showed these responses:

Levels of Thinking	
Lower level	
Memory	none
Translation	inconclusive
Interpretation	some
Higher level	
Application	some
Analysis	some
Synthesis	inconclusive
Evaluation	some

Classroom Conditions	
Discussion opportunity	much
Test/grade stress	none
Lecture	none
Enthusiasm	some
Independence	much
Divergence	much
Humor	much
Ideas valued over grades	much
Enjoyment of ideas	much

Percentage of teacher talk was 40% and average preparation time per week was 1 hour.



TCITY-1971 EVALUATION REPORT: EMPHASIZING CREATIVE TALENTS

Question: What at Twin City fosters Creative Expression?

The above and similar queries were asked during two morning periods during the last ten days of class. Primarily, small random groups of kinds were questioned--about 50 total--NOT in the presence of instructors. Their areas included painting, sculpture, poetry, writing, and drama. Also briefly interviewed were two instructors in sculpture and dance. The choice of students was based on availability. Also, two events were witnessed in drama and dance, plus paintings and sculptures were viewed as works in progress.

Answers from students included:

"I really enjoyed the canoe trip."

"I got a chance to make a lot of friends."

"I am having a lot of fun."

"There is freedom here, we can do what we want."

"I learned a lot."

"My regular school categorizes, here we can switch out of areas we don't like, and extend time on projects we do like."

My response to the above is that they have a different set (from mine) of ideas of what constituted Creative Expression. Within their definition they seemed to succeed. They suggest that success of the Twin City environment to foster C.E. was:

- 1) The availability of tools and teachers--NOT the enforcement.
- 2) The availability of manipulatable time--NOT short class periods and semester-long courses.
- 3) Special community experiences, such as the canoe trip.
- 4) Instructors who were their friends.

Other observations revealed that in the areas of Dance and Drama (two rehearsable arts) physical and communal contacts imbued kids with a special sense of C.E. Here, students were given a chance to note day by day growth as well as longer range improvement measurements. They were appreciative of aspects of process and product (learning and final performance.) They were given a chance to interpret and develop (always within community approval). This was also true of Poetry and its readings:

Personal Remarks:

I must state that the student notion of Creative Expression is vastly different than mine in many ways. Their definitions move toward C.E. as THERAPY AND/OR SOCIAL CONTACT (global village concept?) but from my experience Creative Expression is less provincial. For me it involves larger, natural, "grandeur" things: interconnection, growth, mutability, choice, discovery. the fullest multi-level experiences.

I am not sure that this is possible in a seven week program. But, Twin City must keep trying, with a faculty that exudes inspiration as it disseminates facts and trains skills; nurtures imaginations as the hand follows the mind into the cosmos.

TCITY-1971 EVALUATION REPORT: THE LIGHT-HOUSE SCHOOL

The following is a partial consideration of the ways that TCITY has influenced instructional programs in Minneapolis schools. It is a personal statement based on my work as Secondary Social Studies Consultant for the Minneapolis Public Schools and my discussions with Institute teachers and social studies department chairmen.

It is especially appropriate to ask if social studies teachers not directly associated with the Institute are affected by TCITY offerings in social science. This field involves virtually all secondary school students and some 300 teachers in the district. Furthermore, dissatisfaction with current social studies programs has been expressed by students, teachers and administrators. National trends offer a variety of conflicting perspectives and approaches. In short, the need for change is felt, decisions about innovation in this field must be worked out locally, and the Twin Cities Institute has engaged in structuring learning experiences that might serve as models for new social studies courses.

The quality of recent TCITY work in the social sciences suggests that there would be merit in public school consideration of Institute offerings. The strong points in the TCITY effort would appear to be

- Students and teachers have actively explored the urban "environment beyond the classroom.
- Special attention has been given to student interaction and interaction between students and teachers.
- Questions dealing with values have been sensitively and deliberately explored as an integral part of the learning of concepts and skills of social inquiry.

Nevertheless, the impact of these efforts in Minneapolis high schools would appear to have been negligible because of the following factors--

1. The Institute has developed as a unique institution providing for the special needs of a particular group of young people. This goal has resulted in new approaches and roles different from those existing in the schools.
2. Public school personnel acknowledge the "special" and valuable quality of Institute offerings. However, administrators and teachers have not seen the Institute as providing models of what might be done in their classrooms. Teacher knowledge of Institute offerings is minimal and interest in further information has not been expressed, even though there is interest in programs from curriculum development centers.
3. Institute teachers tend to view their TCITY experience positively, and they thoughtfully explore the implications of this experience for their work in the schools. The autonomy and isolation of their

regular teaching situations does not provide the exchanges that would encourage them to assume leadership in the reconstruction of departmental programs. Several of these teachers are concerned about their inability to communicate their greater awareness of the complexity of teaching-learning processes to colleagues in ways that would further their own reflections and growth. The resulting frustration would appear in a few instances to have contributed to decisions to leave the regular classroom for graduate study and other positions in education.

4. Students who have been involved in the Institute apparently do not return to their schools as agitators for change. Students, like their regular teachers, apparently see the Institute as a unique experience separate from the standard school program. It might be hypothesized that these talented students have learned how to be successful in the public school environment and they resume their previous student roles when they return to the settings in which these behavior patterns have proven successful.

It would appear that the culture of the schools and the unique features of the Institute tend to promote the separation of these experiences. Special programs like TCITY are isolated from the on-going business of the schools because of expectations that students, school personnel and Institute faculty have for both the public schools and the Institute. Explicit efforts to use the Institute's experiences as models for improving the public schools must be devised if a development and demonstration function is desired. Even then success cannot be guaranteed. Efforts to use the regular Minneapolis summer school as a place to test new ideas and to serve as an inservice laboratory for curriculum development and the adaptation of materials from national social studies projects have met with very limited success.

If the Institute is to serve a more effective role in dealing with the urgent need for innovation in the social studies and in other curricular areas, detailed planning involving a range of persons--students, teachers, administrators, chairmen, and consultants--must take place. It must be clear that any effort to use the Institute as a development and demonstration center for curricular revision might very well compromise many of the unique qualities now included in the program as it serves the special needs of talented students.

Again, it must be stressed that these are personal observations focusing on social science programs. In those curricular areas such as computer mathematics the TCITY experience may be a significant element in a broader innovative effort. The Institute teachers of computer math also teach these courses in their schools and they instruct other teachers in staff development courses aimed at establishing similar courses. Also, in foreign language and in the schools involved in TCITY instruction and their work has continued into the school year in ways that have significantly influenced their departments. However, the required social studies courses involving large numbers of teachers and all students in each school have not felt the impact of the TCITY experience.

Robert Berry, Curriculum Supervisor
Rochester Public Schools

TCITY-1971 EVALUATION REPORT: MANAGEMENT

If they changed TCITY to an All Minnesota Institute for Talented Youth--as some have suggested--they could use that well-known slogan, "You're in good hands with All-State." TCITY is shaped by able hands.

The authority of the Institute is highly centralized. Charles Caruson makes the decisions. The teachers have a great deal of leeway; they are expected to take bold steps. But "the Director Man" is in close touch with what they do. He's the gate keeper.

Charles Caruson talks a lot. And he listens a lot. And he asks a lot of questions. He strides across campus, abruptly stops, flashes a big smile, and asks a student a pointed, personal question. "Are you going to the Batique Demonstration this afternoon?" "Are you still riding the bus to campus?" "Did you have a good time at Tamarac?"

One little girl coming into his office early in the Institute was surprised to see him behind the desk. She said "You, the Director Man? I thought the Director Man drove around in a big black Cadillac." Caruson does defy the image.

Bob Rose is second in command, the rest of the command. He makes many of those sweeping rearrangements resulting from Caruson's negotiations. Rose is a hard, productive worker; his track record is impressive. He and Caruson work together sensitively, spending little time together, each knowing his responsibility, doubling up when the situation warrants.

Miss Stepanick handles the secretarial chores in a no-nonsense, get-the-job done way. She gets something reasonably productive out of her summer staff of amateur office workers. There is no obvious substantial fault in her office operation. (No effort was made by this evaluation team to check on fiscal integrity. Also, by appearances, Caruson and Rose do a good job of making contacts for financial support--but this too was not examined.)

The style of management of the total Institute is more intuitive than it is deliberative. This helped make the 1971 Institute more immediate, sensitive, and flexible but less exportable (to other districts) and less of a firm structure against which individual teachers can plan and operate. For example, if the teacher doesn't know the special all-Institute events for the next week it is difficult to plan special class activities.

The director and associate director deal with staff, students, and visitors in a casual manner. Caruson's door is open for people to walk in. Many do. There are bound to be some who wish he and Rose would manage things in more conventional ways. Such people are more likely to remain silent; we did not find any.

The most impressive ability of TCI management is its ability to provide outstanding master teachers.

Overall the TCITY management gets a high rating for employing high quality teachers, keeping red-tape at a minimum, and being sensitive to student needs. It rates a low mark on utilizing the potential of the Institute for district "staff development" and on keeping Institute people informed as to what others there are doing. To the extent that these latter are outside the purpose of TCITY, the management is seen by the evaluation staff to be outstanding.

TCITY-1971 EVALUATION REPORT: MISCELLANEOUS ITEMS

Information Flow. "Getting the word" continued to be a problem, but not as much as it was in 1970. Bulletin boards, news sheets, memos and grapevine carried a substantial information load--but many students and teachers--perhaps a third--felt uninformed about TCITY happenings.

Productivity. TCITY teachers appeared to emphasize a project or performance more in 1971 than 1970. Still, only about two thirds of the 1971 students got involved in something they would call a project.

Teacher Training. TCITY offers the observer an excellent array of effective teaching styles. Trainees can here find opportunities to assume teaching responsibilities. (55 U of M preservice students made daily observations; some took as much of an assignment with a class as the Associate Instructor.) Still, teacher training must be considered a minor payoff. Too few teachers are involved. Many more would probably contaminate the student learning experience. If teacher training were to be considered a major TCITY goal it would require substantial allocation of resources to this purpose. If TCITY were to be used in any substantial way by the two districts for staff development, a better way of rewarding experienced teachers for participation would have to be found.

Afternoon program. The afternoon program was better in 1971 than 1970--in both quality and quantity. However, less than 30% of the students were involved, even on one of the better days, in a TCITY-1971 afternoon activity. The students were not upset--they just had other things to do.

Hostility. Students seemed to be less hostile toward staff and each other this summer. The pleasantness of the campus and the mild weather may have made a difference. There also may be a lessening of the aftermath of Kent State.

Canoe Building. Science students were expected to enroll in a second subject, a mini-course. Most mini-courses were science courses, but the most spectacular was a course in canoe-building. About a dozen students built handsome redwood canoes for themselves.

Recommendation for 1972 Evaluation. The questions raised by the 1971 evaluation team are likely to continue to influence the TCITY staff in 1972. The staff would probably benefit by engaging an evaluation staff with a rather different orientation in 1972, so that different problems will be identified and studied. The cost of evaluation should remain at (or drop below) 2% of the total operating budget. The next evaluators should anticipate continued high resistance from many teachers and students.

TCITY-1971 EVALUATION REPORT: AN ADVOCATE'S STATEMENT

No visitor who took a long, hard look at TCITY-71 kept his skepticism. A young visitor knows how precious it is to discover, to be heard, to belong. An older visitor knows the rarity of a classroom where teachers and students perceive each other as real people. To the non-visitor it doesn't seem possible that a summer school program can deliver on all these promises to over 800 kids, but TCITY-71 did.

Every curriculum specialist fears that by relaxing conduct rules and encouraging student independence they may be saying goodbye to the hard work and hard thinking that education requires. TCITY-71 teachers and students made learning so attractive, so purposive, that free-ranging thought returned again and again to curricular themes: awareness of the human condition, obstacles to communication, ecological interactions, etc.

TCITY excels because of its staff. Its students give it movement. Its directors give it nurture. Its teachers give it movement, nurture, and direction. It would be incorrect to say that Mr. Caruson, Mr. Rose, and the teachers think alike as to the prime goals and methods of education, but collectively, they create a dynamic, humanistically-bent, academically-based curriculum.

The quality of teaching this summer was consistently high, from day to day, from class to class. Some of the teachers chose to be casual, to offer "opportunities", to share a meaningful experience. Others were more intense, more intent upon sharing information and problem solving methods. Both kinds were there, doing it well.

The quality of the learning also was high. The students were tuned in. They were busy. They responded to the moves of their teachers. They improvised, they carried ideas and arguments, indignations and admirations, to the volleyball court, to the Commons, to the shade of campus elms and Cannon River oaks. The youngsters took a long step towards maturity.

True, it was a costly step. Thousands of hours, thousands of dollars, and at least a few hundred aggravations. But fit to a scale of public school budgets--and budgets for parks, interstate highways, and weapons of war--TCITY-71 rates as a BEST BUY. 800 kids, give or take a few, took home a new talent, a new line of thinking, a new awareness--a good purchase.

It cannot be denied that other youngsters in Minneapolis and St. Paul deserve an experience like this. They should have it. Some say, "TCITY is bad because it caters to the elite." But a greater wisdom says "Any effort fixated on giving an equal share of good things to all groups is destined to share nothing of value." For less advantaged youth, a more equitable share of educational opportunities should be guaranteed. But even in times of economic recession, opportunities for the talented should be protected.

TCITY-71 has succeeded. It is even a best buy. It satisfies a social obligation to specially educate some of those who will lead--in the arts, in business, in government, in life. The teachers of TCITY-71 have blended a summer of caring, caprice, openness, and intellectual struggle to give potential leaders a summer of challenge.

(Prepared by R. Stake, not to indicate his opinion of the Institute, but as a summary of the most positive claims that might reasonably be made.)

T-CITY-1971 EVALUATION REPORT: AN ADVERSARY'S STATEMENT

T-CITY is not a scandalum magnatum. But it is both less than it pretends to be and more than it wishes to be. There is enough evidence at least to question certain facets of the Institute--if not to return a true bill against it. Costly, enlarging, innovative; exemplary: these Institute attributes are worthy of critical examination.

How costly is this Institute? Dollar costs are sufficient to give each group of six students \$1,000 to design and conduct their own summer experience. Over 100 Upward Bound students could be readied for their college careers at Macalester. About twenty five expert curriculum specialists could be supported for half a year to design and develop new curricula for the high school.

What is the cost of removing 800 talented leaders from the local youth culture? What is the cost of widening the experience gap between Institute students and their parents?...and their teachers in "regular" high school?...and their non-Institute friends? Not enough here to charge neo-Facist elitism. Enough to warrant discussion.

Institute abounds with self-named innovators and innovations, alternatives to the business-as-usual education of high schoolers. Note that the Institute is not promoted as an exemplary alternative to schooling. It seeks to promote the development of alternative forms of education for schools. And it is failing to do even that job. What is T-CITY doing to demonstrate that the T-CITY style of life could be lived in schools as we know them? Where in the regular school is the staff so crucial to the life of the Institute?...the money?...the administrative leadership? Where are the opportunities for the teachers, principals, superintendents to come and live that life that they might come to share in the vision?...and where are the parents? T-CITY should be getting poor grades on affecting the regular school program.

There are other dimensions of T-CITY that puzzle the non-believer:

- *** How long can in-class "rapping" continue and still qualify as educative self-exploration? Are there quality control procedures in effect during the summer program: For example: when one-third to one-half a class is absent from a scheduled meeting, should not that be seen as an educational crisis by the instructor?
- *** What does T-CITY do to help students realize that the Institute standards are necessarily high; that the regular school norms and expectations do not count; that a heretofore "best" becomes just a "so-so"? There are unnecessarily disheartened students in T-CITY.
- *** Is it unreasonable to expect that more than two of twenty-two teachers or associate teachers would have some clear idea or plan for utilizing T-CITY approaches in their regular classrooms next fall?

4A

Saturday, July 31, 1971

Minneapolis Tribune

Established 1867

Bower Hawthorne Editor
Wallace Allen Managing Editor
Leonard Inskip Editorial Editor

Volume CV Number 68

'Talented youth' in perspective

The controversy over the archaeological excavations undertaken at Welch, Minn., by students from the Twin City Institute for Talented Youth (TCITY) has focused attention on one part of its curriculum, but has overshadowed its total program. And that's too bad, because TCITY is considered by many in education to be one of the best and most innovative summer programs in the country, and it points toward ways to make regular schools more effective in meeting students' needs.

The institute was begun in 1967 in an effort to get talented students from Minneapolis and St. Paul schools together so that they could challenge and learn from one another and develop their abilities beyond the limited scope of the normal classroom. The curriculum included advanced courses not ordinarily found in school systems, but the approach to them was along the usual lines at first, with each student sticking to his own subject area. Since then, it has evolved, both in form and content. Last year, TCITY organized all students' programs around a theme, Man and his Environment, and encouraged an interdisciplinary approach to that general topic.

This year, the theme is Man and Society. The archaeological dig at Welch was one approach to that theme, and the institute's catalog describes its premise this way: "By working with a culture outside your own, you

should gain a different viewpoint from which to not only view your own culture, but gain an insight into the uniformity of man, both temporally and geographically."

The rest of the curriculum, whether in the fine arts, sciences, languages or social studies, has taken that same human-centered approach. Students have been encouraged and helped to become aware of and to learn from the people and places in the community around them. One course, for example, is called "The Urban Hero," and the catalog points out that "the hero is happening in the city now" as people learn to survive in the urban world. "This summer will be spent finding the urban heroes—the artists, the activists, humanists, philosophers, urban guerrillas—and learning the politics of survival."

Does it work? That question is perhaps best answered by two participants in last summer's program, quoted by Tribune staff writer Catherine Watson in an article in *American Education*, a magazine published by the U.S. Department of Health, Education and Welfare's Office of Education:

"This is the only school I know of," said one girl, "where you get dropins instead of dropouts." And another said, "When you see what school can be like here at the institute, you feel an almost desperate need to change the schools."